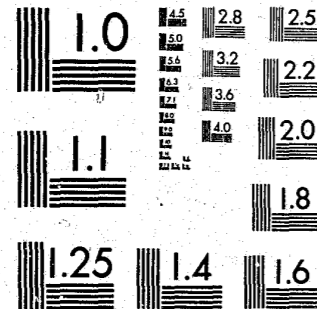


National Criminal Justice Reference Service



This microfiche was produced from documents received for inclusion in the NCJRS data base. Since NCJRS cannot exercise control over the physical condition of the documents submitted, the individual frame quality will vary. The resolution chart on this frame may be used to evaluate the document quality.



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Microfilming procedures used to create this fiche comply with the standards set forth in 41CFR 101-11.504.

Points of view or opinions stated in this document are those of the author(s) and do not represent the official position or policies of the U. S. Department of Justice.

National Institute of Justice
United States Department of Justice
Washington, D. C. 20531

7/1/83



UNIVERSITY OF SOUTHERN CALIFORNIA

social science research institute

LONGITUDINAL RESEARCH IN THE UNITED STATES:
RELEVANCE TO PRIMARY PREVENTION OF DELINQUENCY

By
SARNOFF A. EDNICK

AND
MICHELE HARWAY

VOLUME II-A

DESCRIPTIONS OF LONGITUDINAL RESEARCH

Prepared for the United States Department of Justice,
Law Enforcement Assistance Administration under contract
No. 78-NI-AX-0088. Points of view or opinions stated
in this document are those of the authors and do not
necessarily represent the official position or policies
of the United States Department of Justice.

84660, 92

TABLE OF CONTENTS

DESCRIPTIONS OF LONGITUDINAL RESEARCH

VOL. II A

1. The California Child Health and Development Studies
by B.J. van der Berg 1 - 13
2. The St. Louis Baby Study: Theory, Practice and Findings
by T.E. Jordan 14 - 28
3. A Neuropsychiatric Follow-up of Children in the Collaborative
Perinatal Project Population
by P.A. O'Connor, D. Shaffer, S. Shafer & C. Stokman 29 - 50
4. The Philadelphia Collaborative Perinatal Study
by J. Marecek 51 - 66
5. The Psychosocial Development of Aggressive Behavior
by L.D. Eron, L.R. Huesmann, L.O. Walder & M.M. Lefkowitz 67 - 81
6. ETS Study of Academic Prediction and Growth
by T.L. Hilton 82 - 96
7. The Youth in Transition Project
by J.G. Bachman & P.M. O'Malley 97 - 118

U.S. Department of Justice
National Institute of Justice

This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated in this document are those of the authors and do not necessarily represent the official position or policies of the National Institute of Justice.

Permission to reproduce this copyrighted material has been granted by

Public Domain/LEAA
U.S. Dept. of Justice

to the National Criminal Justice Reference Service (NCJRS).

Further reproduction outside of the NCJRS system requires permission of the copyright owner.

NCJRS

FEB 4 1982

ACQUISITIONS

9. An Epidemiological Study of the Behavioral Problems of Young Children by H.Z. Reinherz, et al.	119 - 131
10. Self-Attitudes and Deviant Responses: Toward the Validation of a General Theory of Deviant Behavior by H.B. Kaplan	132 - 140
11. Family and Peer Processes in Adolescent Drug Use by D.B. Kandel	141 - 154
12. Adolescence to Young Adulthood : A Twelve-year Prospective Study of Problem Behavior and Psychosocial Development by R. Jessor & S.L. Jessor	155 - 168
13. Using Longitudinal Data to Study College Impact by A.W. Astin	169 - 184
14. Project TALENT: A Longitudinal Study of the Development and Utilization of Individuals' Capabilities by L. Steel, L.L. Wise & R.P. Abeles	185 - 208
15. Early Signs of Delinquency and Related Behavioral Difficulties: A Longitudinal Study by G. Spivack	209 - 228
16. The National Longitudinal Surveys of Labor Force Behavior by H. Parnes & K. Sproat	229 - 254
17. The Southern Youth Study: Longitudinal Survey Research in the Deep South by J.S. Picou, J.K. Thomas & A.G. Cosby	255 - 264
18. The Florida Longitudinal Project : A Review by J.M. Fletcher, P. Staz & R. Morris	265 - 284

19. Longitudinal Health Research in the U.S. Navy by E. Gunderson, R.E. Mitchell & R.J. Biersner	285 - 300
20. The Veterans Administration Normative Aging Study by R. Bosse, D.J. Ekerdt & J.E. Silbert	305 - 322
21. Longitudinal Studies of Impaired Competence in the Community: A Description of the Riverside Community Research Project (1963 - 1979) by E.W. Butler & S.S. Friedman	325 - 344
22. The Nebraska Annual Social Indicator Surveys by D.R. Johnson & L.L. White	343 - 356
23. The Woodlawn Mental Health Longitudinal Community Epidemiological Project by S.G. Kellam, M.E. Ensminger, J. Branch, C.H. Brown & J.P. Fleming	354 - 367
24. The Midtown Manhattan Longitudinal Study by L. Srole & A.K. Fischer	368 - 382
25. City Children: The Family Research Project by T.S. Langner	383 - 404
26. Life Events, Role Conditions and Adaptation: A Panel Study of Social Stress by E.G. Menaghan, L.I. Pearlin, M.A. Lieberman & J.T. Mullan	405 - 415
27. The New York Longitudinal Study (NYLS) by A. Thomas & S. Chess	416 - 429
28. Longitudinal Predictive Study of Success or Failure in Birth-Control Programs by S. Oskamp & B. Mindick	430 - 454

A Longitudinal Study of Personality and Cognitive Development by J.H. Block & J. Block	455 - 465
A Longitudinal Study of Children from Kindergarten into the Adult Years by P.E. Kraus	469 - 480
The Child's Political World by J. Lare, S.W. More & K.A. Wagner	481 - 491
Schooling and Young Children by D.R. Entwistle & L.A. Hayduk	492 - 503

The California Child Health and Development Studies *

Bea J. van den Berg, M. D.

* From the Child Health and Development Studies, School of Public Health University of California, Berkeley, California; before 1972 of the Kaiser Foundation Research Institute, Oakland, California. From its inception CHDS has been supported by grants from the National Institute of Child Health and Human Development.

Associates in CHDS research projects are Roberta E. Christianson, Lucille Milkovich and Frank W. Oechsli.

The California Child Health and Developmental Studies

Bea J. van den Berg, M. D.

Introduction

The overall objectives of the Child Health and Development Studies (CHDS) were formulated as follows: To investigate the relationships of biological, genetic, medical and environmental factors in the parents--including events in pregnancy, labor and delivery--to the normal and abnormal development of offspring; to investigate the relationships of these factors to pregnancy loss in the form of early fetal death, perinatal mortality, infant and childhood mortality, and to incidence of congenital anomalies; to provide detailed growth curves for children suitable for the study of differences in growth rate according to the above factors; to provide estimates of illness and injury rates in infants and children; and to investigate the relationships of physical and mental development in early childhood to those during the period of puberty and adolescence.

The Population

Following a pilot study in 1959, it was planned to include approximately 20,000 pregnancies in the studies. This number was considered large enough to study most events occurring during pregnancy and delivery, including conditions and diseases, and prescribed drugs, to study fetal and neonatal deaths, and to identify severe and not severe congenital anomalies and developmental aberrations. On the other hand, the number of 20,000 was considered not too big to allow the accurate accumulation of data essential for carrying out of the main objectives.

The studies are conducted among families, residing in the Oakland area of California, who were members of the Kaiser Foundation Health Plan, a prepaid medical care program. This plan provides comprehensive medical care to the families of members in the Kaiser Hospital facilities. The characteristics inherent in this population and advantageous for a

longitudinal research survey are: the population receives all medical care at the Kaiser facilities; medical care is equally available for all members; reports of medical services rendered by physicians in the various specialties are assembled in a single file for each patient; and the population is socioeconomically broad based, only the extremes--the very indigent and the very affluent--are not represented.

Enrollment of pregnant women started August 1959 and ended September 1966. In all, some 20,500 pregnancies and some 19,000 live-born children constituted the basic study population. The distribution by ethnic background is as follows: 65% of the women are white, 24% are black, 4% of Asian background, 3% Mexican-American, and 4% are of other or mixed ethnic background.

The Assembled Data

Besides the data obtained specifically for study purposes, information available from physicians' observations whenever the gravida or the child received out-patient and in-patient services in any of the Kaiser Foundation medical facilities was utilized.

The assembled data are briefly summarized in four sections:

Medical records on pregnancy, labor and delivery

Pregnancy interview

Medical records of the children

Developmental examinations on subcohorts of the children

Medical Records on Pregnancy, Labor and Delivery

The study utilizes the unique opportunity provided by the data available in a prepaid medical care program to obtain virtually complete information on the medical history of the Study population. All medical information on visits to any of the clinics of Kaiser Foundation Hospitals, prenatal clinics as well as others, has been assembled in one file for each patient. From these records, observations during pregnancy, including all routine body weight and blood pressure measurements, laboratory results, comprehensive data on labor and delivery (including time, dose and type of obstetric analgesics and anesthetics) have been abstracted and coded. A gross description of the placenta was obtained. Furthermore, all medically

attended intercurrent diseases and drug prescriptions were abstracted and coded and used for analyses. This information is available on all 20,530 gravidas delivered of 19,044 live-born children, of whom 18,751 survived the neonatal period.

Pregnancy Interview

A subcohort of 15,865 women participated in an extensive interview early in pregnancy. The other women received prenatal care in one of the outlying Kaiser clinics, where no full-time interviewer for the research project was available.

The extensive CHDS interview covered detailed information on the women's reproductive history, dates and outcomes of previous pregnancies, birth weight and duration of pregnancy of her live-born children, and marital history. Also, information was obtained on gravida's health history and that of her husband and children, data on gravida's and husband's education, occupation and family income. The gravidas responded to questions regarding their family of orientation: On their parents' health, on their fathers' occupation, number of brothers and sisters, their parents' marital state when they were growing up, state (or country) of birth, etc. Further, questions were asked relating to the gravida's present pregnancy, on the exact date of her last menstrual period, on usage and type of contraceptives, planning of pregnancy, and on her psychological orientations toward the pregnancy.

Medical Records of the Children

The medical records of the children's visits to the pediatric clinics and all other clinics (or hospitalizations) are assembled in one record file for each child. From these records we abstracted the medical data for each visit, including diagnosis and treatment, physical growth measurement, and measurements of blood pressure when recorded. These abstracted records--covering the child's health history from birth to the age of at least 5 years, and for many children until 10 years--represent the CHDS basic health file on the children. A large proportion of the data have been coded, are computer-accessible, and have been used for various analyses and studies. Information on congenital anomalies, severe and nonsevere, has been completely coded. Birth weight and length, as well as all subsequent measurements of height and weight are available in coded form.

A certain attrition rate is to be expected from a large cohort study. However, extensive follow-up activities have kept losses low; e.g., at the age of 1 year we did not know the status of only 42 children out of a

total of 19,044 live-born infants, an attrition rate of 0.2%. The procedures were such that regular contact with the mothers was maintained to know the whereabouts of the families. By the time the children reached their fifth birthday, 89.4% were still observed (68.2% were still Kaiser Health Plan members and 21.2% were CHDS follow-up contacts); 3.2% died, were adopted or institutionalized, and 7.4% were lost from observation.

Developmental Examinations on Subcohorts of the Children

Besides the information which has been abstracted from the medical records for all of the children, substantial additional information has been assembled from CHDS developmental examinations for certain subcohorts of the children. Eligibility was defined by birth date within specified limits and residence in the San Francisco East Bay Area.

At their fifth birthday, children born between June 1960 and April 1963, and those born between April 1964 and April 1967, who were still living in the East Bay Area, were invited for these special developmental examinations and follow-up interviews. In the earlier examination period, some 4,000 children were included. Over 3,400 children participated in the later examination with a modified and extended protocol. Both examinations included a general physical examination and tests of acuity of vision, hearing, speech, and cognitive ability, and in the second examination, the mothers responded to a 42-item inventory of the child's behavior and mother-child relationship.

From April 1971 until April 1972, some 3,600 children who were born between April 1960 and April 1963 were examined at their ninth, tenth or eleventh birthday (with funds from Kaiser Foundation Research Institute). This examination included anthropometrics (weight, standing and sitting height, biacromial and bi-iliac distance, head circumference, chest depth and breadth, and triceps and subscapular skinfolds), and cognitive ability tests (Peabody Picture Vocabulary Test and Raven Progressive Matrices). In addition to an interview on the health and well-being of the family, the mothers of the 9-11-year old children responded to a 100-item inventory relating to the child's behavior and mother-child relationship. A cognitive ability test for the mother (Peabody Picture Vocabulary Test) was also included. Blood pressure measurements and screening of vision, hearing and speech was done for a separate sample of 698 children at 9 years of age.

About 10% of the families that we contacted and were eligible for one of the developmental examinations were not examined. In these cases, we arranged for a short telephone interview to establish reason for non-participation. Our longitudinal observations, that are available for participants as well as for refusals, allows us to detect the possibility of bias in the examined groups.

All data required for this project are assembled and accessible on magnetic disk computer files. Data from the various files can be combined quite flexibly for analysis. The project maintains its own small computer (IBM model 1130) and has its own library of research-oriented programs and subprograms, fully documented and catalogued.

We recently finished examining adolescents 15-17 years old and their mothers. Eligibility of the adolescents for this study was based on earlier participation in the developmental examinations at 5 years and 9-11 years; 2,350 were still living in the Oakland area and 85% or 2,000 adolescents were examined. The adolescent study was a part of a larger project with the objective to identify precursors of high blood pressure in young persons. The examination of the adolescents and of their mothers included several blood pressure measurements (under basal conditions and after muscular stress), spirometry, and anthropometric measurements. Interviews and questionnaires covered a health history and a wide range of life-style items.

Research Activities, Summary

CHDS extensive data on longitudinal observations of maternal and child health covering a broad spectrum of medical, biological and social factors made it possible to study a wide variety of subjects. As a consequence, researchers of different biomedical and behavioral disciplines have participated in the research, using various approaches and methodologies. These activities have resulted in a large number of published papers and doctoral dissertations, divided into eleven subject areas. They include:

1. Physiology and Pathology of Pregnancy
2. Drugs and Fetal Development
3. Genetic Studies, Primarily Relating to Blood Groupings
4. Cigarette Smoking During Pregnancy
5. Prematurity and Low Birth Weight
6. Fetal Death and Child Mortality
7. Congenital Anomalies
8. Childhood Morbidity

9. Growth Studies
10. Studies Related to Family Planning
11. The Adolescent Study

Other projects, based on the data assembled in the Adolescent Study are being planned. These include studies on development of pre- and post-pubertal obesity, lung function in relation to smoking and history of respiratory diseases, as well as studies on educational and vocational expectations and on family relationships of teenagers.

References

- Allen, F. H., Jr., Milkovich, L., & Corcoran, P. A. A pedigree showing that Y^T is not in the Lutheran blood-group system. Vox Sanguinis, 1963, 8, 376-377.
- Allen, F. H., Jr., Madden, H. J. & King, R. W. The MN gene MU, which produces M and U but no N, S, or x. Vox Sanguinis, 1963, 8, 549-556.
- Beck, G. J. & van den Berg, B. J. The relationship of the rate of intrauterine growth of low-birth-weight infants to later growth. Journal of Pediatrics, 1975, 66(4), 504-511.
- Carlson, R. & Welch, Q. B. A new model for measuring breeding genetic distance. Annals of Human Genetics, 1977, 40, 455-461.
- Carlson, R. & Welch, Q. B. A new model for measuring breeding genetic distance. II. Polygenic Traits. Annals of Human Genetics, 1978, 41, 491-500.
- Christianson, R. E. Down syndrome and maternal age. The Lancet, 1976, II(7996), 1198.
- Christianson, R. E. Studies on blood pressure during pregnancy.
1. Influence of parity and age. American Journal of Obstetrics and Gynecology, 1976, 125(4), 509-513.
- Christianson, R. E. Gross differences observed in the placentas of smokers and nonsmokers. American Journal of Epidemiology, 1979, 110(2), 178-187.
- Christianson, R. E. & Page, E. W. Diuretic drugs and pregnancy. Obstetrics and Gynecology, 1976, 48(6), 647-652.
- Cohen, B. H. ABO and Rh incompatibility. I. Fetal and neonatal mortality with ABO and Rh incompatibility? Some new interpretations. American Journal of Human Genetics, 1970, 22(4), 412-440.
- Cohen, B. H. ABO and Rh incompatibility. II. Is there a dual interaction in combined ABO and Rh incompatibility? American Journal of Human Genetics, 1970, 22(4), 441-452.
- Food and Drug Administration. Meprobamate and chlordiazepoxide, drugs for human use: Drug efficacy study implementation; amended notice. Federal Register, 1976, 41(143).

- Hoffman, J. I. E. & Christianson, R. E. Congenital heart disease in a cohort of 19,502 births with long-term follow-up. American Journal of Cardiology, 1978, 42(4), 641-647.
- Horowitz, D. R. & Oechsli, F. W. Contraceptive risk taking in a population with limited access to abortion. Journal of Biosocial Science, in press.
- Laukaran, V. H. The relationship of maternal attitude to pregnancy outcome and child health: a cohort study of unwanted pregnancy. Unpublished doctoral dissertation, University of California, Los Angeles, 1978.
- Laukaran, V. H. & van den Berg, B. J. The relationship of maternal attitude to pregnancy outcome and obstetrical complications. American Journal of Obstetrics and Gynecology, 1980, 136(3), 374-379.
- Lewak, N., van den Berg, B. J. & Beckwith, J. B. Sudden infant death syndrome risk factors: prospective data reviewed. Clinical Pediatrics, 1979, 18(7), 404-411.
- Madden, H. J., Cleghorn, T. E., Allen, F. H., Jr., Rosenfield, R. E., & Mackeprang, M. A note on the relatively high frequency of St^a on the red blood cells of Orientals, and report of a third example of anti-St^a. Vox Sanguinis, 1964, 9, 502-504.
- Milkovich, L. & van den Berg, B. J. Effects on prenatal meprobamate and chlordiazepoxide hydrochloride on human embryonic and fetal development. New England Journal of Medicine, 1974, 291(24), 1268-1271.
- Milkovich, L. & van den Berg, B. J. An evaluation of the teratogenicity of certain anti-nauseant drugs. American Journal of Obstetrics and Gynecology, 1976, 125(2), 244-248.
- Milkovich, L. & van den Berg, B. J. Effects of antenatal exposure to anorectic drugs. American Journal of Obstetrics and Gynecology, 1977, 129(6), 637-642.
- Oechsli, F. W. Prepaid medical care: Ethnic and social differentials in infant and child mortality. Paper presented at the American Public Health Association Annual Meeting, New Orleans, October 1974.
- Oechsli, F. W. A population model based on a life table that includes marriage and parity. Theoretical Population Biology, 1975, 7(2), 229-245.

- Oechsli, F. W. A general method for constructing increment-decrement life tables that agree with the data. Theoretical Population Biology, 1979, 16(1), 13-24.
- Page, E. W. & Christianson, R. E. The impact of mean arterial pressure in the middle trimester upon pregnancy outcome. American Journal of Obstetrics and Gynecology, 1976, 125(6), 740-746.
- Page, E. W. & Christianson, R. E. Influence of blood pressure changes with and without proteinuria upon outcome of pregnancy. American Journal of Obstetrics and Gynecology, 1976, 126(7), 821-833.
- Peckham, C. H. Uterine bleeding during pregnancy. I. When not followed by immediate termination of pregnancy. Obstetrics and Gynecology, 1970, 35(6), 937-941.
- Peckham, C. H. Uterine bleeding during pregnancy. II. Factors relating to its incidence. Obstetrics and Gynecology, 1972, 39(1), 48-51.
- Peckham, C. H. & Christianson, R. E. The relationship between pre-pregnancy weight and certain obstetric factors. American Journal of Obstetrics and Gynecology, 1971, 111(1), 1-7.
- Peckham, C. H. & King, R. W. A study of intercurrent conditions observed during pregnancy. American Journal of Obstetrics and Gynecology, 1963, 87(5), 609-620.
- Peckham, C. H. & Yerushalmy, J. Aplasia of one umbilical artery: Incidence by race and certain obstetric factors. Obstetrics and Gynecology, 1965, 26(3), 359-366.
- Peritz, E. A statistical study of intrauterine selection factors related to the ABO system. I. The analysis of data on liveborn children. Annals of Human Genetics, 1967, 30, 259-271.
- Peritz, E. A statistical study of the intrauterine selection factors related to the ABO system. II. The analysis of fetal mortality data. Annals of Human Genetics, 1971, 34(4), 389-394.
- Peritz, E. & Rust, P. F. On the estimation of the nonpaternity rate using more than one blood-group system. American Journal of Human Genetics, 1972, 24(1), 46-53.
- Peritz, E. & Sproul, A. Some aspects of the analysis of hand-wrist bone-age readings. American Journal of Physical Anthropology, 1971, 35(3), 441-447.

Reed, T. E. Research on blood groups and selection from the Child Health and Development Studies, Oakland, California. I. Infant birth measurements. American Journal of Human Genetics, 1967, 19(6), 732-746.

Reed, T. E. Research on blood groups and selection from the Child Health and Development Studies, Oakland, California. II. Gravidae's characteristics. American Journal of Human Genetics, 1968, 20(2), 119-128.

Reed, T. E. Research on blood groups and selection from the Child Health and Development Studies, Oakland, California. III. Couple mating type and reproductive performance. American Journal of Human Genetics, 1968, 20(2), 129-141.

Rust, P. F. On the probability of detecting nonpaternity through the Rh blood-group system. American Journal of Human Genetics, 1972, 24(1), 54-57.

Schoen, E. J., Solomon, I. L. & Milkovich, L. Bone age in short boys and tall girls. American Journal of Diseases of Children, 1970, 119, 252-254.

Shaw, E. B. Sudden unexpected death in infancy syndrome. American Journal of Diseases of Children, 1968, 116, 115-119.

Shaw, E. B. Sudden unexpected death in infancy syndrome. American Journal of Diseases of Children, 1970, 119, 416-418.

Sproul, A. & Peritz, E. Assessment of skeletal age in short and tall children. American Journal of Physical Anthropology, 1971, 35(3), 433-439.

Starfield, B., van den Berg, B. J., Steinwachs, D. M., Katz, H. P. & Horn, S. D. Variations in utilization of health services by children. Pediatrics, 1979, 63, 633-641.

Taylor, W. F. On the methodology of measuring the probability of fetal death in a prospective study. Human Biology, 1964, 36(2), 86-103.

Taylor, W. F. The probability of fetal death. Congenital Malformations, Proceedings of the Third International Conference, The Hague, Excerpta Medica, International Congress Series No. 204, 1970, 307-320.

Tuddenham, R. D., Brooks, J. & Milkovich, L. Mothers' reports of behavior of ten-year-olds: relationships with sex ethnicity, and mother's education. Developmental Psychology, 1974, 10(6), 959-995.

van den Berg, B. J. Idiopathic respiratory distress syndrome (IRDS) and other forms of dyspnea (D) in a large cohort of births. I. Incidence by birthweight and gestation in relation to antepartum bleeding, cesarean section, and breech presentation. Proceedings of the International Congress of Pediatrics, August 1971, XVII, 171-176.

van den Berg, B. J. Studies on convulsive disorders in young children. III. Recurrence of febrile convulsions. Epilepsia, 1974, 15, 177-190.

van den Berg, B. J. Studies on convulsive disorders in young children. IV. Incidence of convulsions among siblings. Developmental Medicine and Child Neurology, 1974, 16(4), 457-464.

van den Berg, B. J. Epidemiologic observations of prematurity: Effects of tobacco, coffee and alcohol. In D. Reed & F. Stanley (Eds.), The Epidemiology of Prematurity. Baltimore: Urban & Schwarzenberg, 1977.

van den Berg, B. J. & Yerushalmy, J. The relationship of the rate of intrauterine growth of infants of low birth weight to mortality, morbidity, and congenital anomalies. Journal of Pediatrics, 1966, 69(4), 531-545.

van den Berg, B. J. & Yerushalmy, J. Studies on convulsive disorders in young children. I. Incidence of febrile and nonfebrile convulsions by age and other factors. Pediatric Research, 1969, 3, 298-304.

van den Berg, B. J. & Yerushalmy, J. Studies on convulsive disorders in young children. II. Intermittent phenobarbital prophylaxis and recurrence of febrile convulsions. Journal of Pediatrics, 1971, 78(6), 1004-1012.

van den Berg, B. J. & Yerushalmy, J. Studies on convulsive disorders in young children. V. Excess of early fetal deaths among pregnancies preceding the birth of children with febrile or nonfebrile convulsions. Journal of Pediatrics, 1974, 84(6), 837-840.

Wingerd, J. The relation of growth from birth to 2 years to sex, parental size and other factors, using Rao's method of the transformed time scale. Human Biology, 1970, 42(1), 105-131.

Wingerd, J., Christianson, R., Lovitt, W. V. & Schoen, E. J. Placental ratio in white and black women: relation to smoking and anemia. American Journal of Obstetrics and Gynecology, 1976, 124(7), 671-675.

Wingerd, J., Peritz, E. & Sproul, A. Race and stature differences in the skeletal maturation of the hand and wrist. Annals of Human Biology, 1974, 1(2), 201-209.

Wingerd, J. & Schoen, E. J. Factors influencing length at birth and height at five years. Pediatrics, 1974, 53(5), 737-741.

Wingerd, J., Schoen, E. J. & Solomon, I. L. Growth standards in the first two years of life based on measurements of white and black children in a pre-paid health care program. Pediatrics, 1971, 47(5), 818-825.

Wingerd, J., Solomon, I. L. & Schoen, E. J. Parent-specific height standards for preadolescent children of three racial groups, with method for rapid determination. Pediatrics, 1973, 52(4), 555-560.

Yerushalmy, J. Statistical considerations and evaluations of epidemiological evidence. In G. James and T. Rosenthal (Eds.), Tobacco and Health. Springfield: Charles G. Thomas, 1962.

Yerushalmy, J. Mother's cigarette smoking and survival of infant. American Journal of Obstetrics and Gynecology, 1964, 88(4), 505-518.

Yerushalmy, J. Religious, educational, and socio-economic factors associated with different methods of fertility control. Proceedings of U. N. World Population Conference, August 1965, Belgrade, Yugoslavia.

Yerushalmy, J. The classification of newborn infants by birth weight and gestational age. Journal of Pediatrics, 1967, 71(2), 164-172.

Yerushalmy, J., The California Child Health and Developmental Studies. Study design, and some illustrative findings on congenital heart disease. In Congenital Malformations, Proceedings of the Third International Conference, The Hague, Excerpta Medica International Congress Series No. 204, 1969, 299-306.

Yerushalmy, J. Self-selection--A major problem in observational studies. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, 1970, 1971; Vol. IV: Biology and Health, 329-342.

Yerushalmy, J. The relationship of parents' cigarette smoking to outcome of pregnancy--implications as to the problem of inferring causation from observed associations. American Journal of Epidemiology, 1971, 93(6), 443-456.

Yerushalmy, J. Infants with low birth weight born before their mothers started to smoke cigarettes. American Journal of Obstetrics and Gynecology, 1972, 112(2), 277-284.

Yerushalmy, J. Methodologic problems encountered in investigating the possible teratogenic effects of drugs. In M. A. Klingberg, A. Abramovici & J. Chemke (Eds.), Drugs and Fetal Development. New York: Plenum Publishing Corp., 1972.

Yerushalmy, J. & Milkovich, L. Evaluation of the teratogenic effect of meclizine in man. American Journal of Obstetrics and Gynecology, 1965, 93(4), 553-562.

Yerushalmy, J., van den Berg, B. J., Erhardt, C. L. & Jacobziner, H. Birth weight and gestation as indices of 'immaturity.' American Journal of Diseases of Children, 1965, 109, 43-57.

The St. Louis Baby Study:
Theory, Practice, and Findings

Thomas E. Jordan
University of Missouri - St. Louis

The St. Louis Baby Study:
Theory, Practice, and Findings

Thomas E. Jordan
University of Missouri - St. Louis

Scope

Our inquiry explores several quite distinct domains. Longitudinal study with its techniques and challenges creates a nexus of information which varies in relevance as new data appear. With deliberation, we set out to generate a corpus of data on a representative North American cohort of newborns and their families. This discrete objective was instrumental to contributing to formation on policy for childhood, our ultimate objective.

Design

The program of studies since approximately 1966 has consisted of prospective longitudinal studies of individual children. The inquiry is naturalistic in the sense of an absence of interventions and treatments planned by the investigator. It is also clinical, in the sense that each child has been seen by the same caseworker whenever possible, a relationship in which each successful contact maximizes the chances of another one. Of course, some families have chosen to expose their children to educational experiences, and there is variation in patterns of television viewing and preschool attendance. In addition, each home is a mini-culture and children have various configurations of kin. The presence or absence of male heads-of-household, occurrence of divorce or separation, and death are examples of Nature assigning children to treatment, apart from assigning a genetic endowment to each child.

Relevant Studies

The reader may wish to judge how similar the St. Louis Baby Study is to other studies. We view it, essentially, as having the following characteristics:

1. Prospective case studies of individual children in their homes.
2. Cohort drawn from a representative North American metropolitan area.
3. Probands representing the spectrum of the community,
 - rich, poor
 - black, white
 - inner city, suburbs
 - major religions
 - none excluded on grounds of illegitimacy, twinning, degree of access, etc.
4. Multivariate analyses of data.
5. Computer storage, retrieval, and analysis.
6. Criteria in several domains of child development.
7. Data on social and family variables.

Major Findings

Findings to date are limited to the period from birth to school entry, since analyses of data always lag behind acquisition in prospective studies. However, the span of years examined in one book (Jordan, 1980), several monographs, and many reports, contain several of the stages we consider important in our model of development.

We have established, through analyses of both continuous and categorical data, that biological risk in the perinatal phase is replaced in significance by environmental factors at about age two years (Jordan, 1971, 1979).

Among environmental influences the most consistent for all phases of child development is the aggregate of sex, perinatal SES, ethnic group, and home stimulation (STIM). Among single predictors, within appropriate clusters of independent variables, the most influential in the preschool years by domain appear to be: Physical criteria - sex: Cognitive criteria - ethnic group and SES.

Maternal traits which influence the non-somatic aspects of development are years of schooling, and authoritarian-liberal child rearing ideology especially.

In the matter of other influences race is barely perceptible and is washed out by social class data (Jordan, 1974). This does not mean that black-white differences do not exist; they do, but low attainment scores in black preschoolers relate in a statistically significant way to low SES scores, but not to race, when both variables are present.

In studies of handicapped children (Jordan, 1963, 1976a, 1976c, 1979) we have found developmental histories to be useful tools in the study of bio-social antecedents to physical, mental, and learning disabilities.

Expectations of high reliability of child measures are sometimes unreasonably high. In the case of somatic measurements, physical structures, e.g. the ulna, sometimes fluctuate in circadian fashion (Valk and Van Den Busch, 1978). Temporal instability is the reality, and psychometric reliability is a proportionately less relevant construct.

Some Considerations

Cohorts can be formed relatively easily by seeking subjects who share a common social identity such as attending the same church, or to better effect, are in the same year of schooling. Such groups are captive, in the sense that they are not at-large. Our families enjoy no common element and do not know each other. This is an important element in planning since there are no short cuts; each family is contacted at the home address and has been visited many times by the same caseworker. In all cases workers and families have been matched by race, and occasionally by sex.

Data in child development studies take many forms. From the beginning we have taken data by non-invasive means. There have been no blood samples, and no devices, e.g., skin-fold calipers, have been applied. On one occasion, small hair samples were taken for lead-poison studies which proved impractical; the procedures were probably counter-productive in the long run.

As with other studies (Tanner and Israelsohn, 1963) we have had problems with fathers. Our data on fathers are limited to items we obtain from their wives, which has enabled us to avoid some terminations of cooperation. The risk to cooperation posed by fathers is also reported from New Zealand (Silva, personal communication).

Because we have employed non-captive cohorts and work in politically and socially volatile segments of the community we have taken great pains to suppress publicity (Cromwell, Vaughn, and Mindel, 1975).

A final consideration of interest is the challenge of interacting with relevant agencies. Schools have been the major challenge. On the one hand individual teachers have been very cooperative; on the other, some middle level bureaucrats in one school system have been uncooperative, while others have been very helpful. Parents' requests in writing for test scores have been refused, despite their legal rights to such information under Federal laws, on the grounds that local school board policy does not permit release of information to parents.

Events As Independent Variables

Cohort research is not without sensitivity to treatments; it is that the investigator accepts Nature as the source of differential experiences affecting development. Not the least of the interventions by Mother Nature is the year by which a cohort is defined. Research indicates that certain years, especially 1926-27 and 1935-36 in Gt. Britain (Wilkins, 1960), and 1968 in the United States (Richman, 1975), are critical for study of selected aspects of crime and addiction. At the time of writing, a teachers' strike has destroyed the instructional year of a local school system. The set of children entering first and second grades now constitute a cohort distinguished by disruption of schooling at its most sensitive stage. We would not want to consider these youngsters to be just another age/school year group, but only the passage of time will tell the hypothesis of their vulnerability. In the case of our 1966-67 cohort a rather different example of the intrusion of events into field research occurred. The murder of Dr. Martin Luther King had two effects. First, it closed the black community to research for a period of time, and second, it demoralized some workers on the research team. Time passed, and the work continued. Perhaps the best example of this susceptibility to events in the real world is the occlusion of research by an outbreak of disease, such as rubella in Baltimore (Bordlay, et al., 1968).

Three Cohorts

Our program began in the late fifties and is not really the consequence of the movements of the sixties. The first cohort was used for methodological studies and exploration of procedures. A major outcome was rejection of hospital data gathered by non-specialists and the decision to take perinatal data through research-oriented people stationed in hospitals solely for the purpose. The second, 1966 cohort, was planned to fall after the date for local school entry and so was structured for December and later. Also, the probands were recruited week by week so that subsequent data-taking would be spread out over four months. This contrasts with other patterns of cohort formation such as all cases born on one day (Prokopec, 1960), or within one week (Douglas, Ross, and Simpson, 1971). No children were excluded, except by non-cooperation, and may be

contrasted with exclusion of Francophones (Hopkins, 1947) illegitimates (Neligan, Prudham and Steiner, 1974), children without a male head-of-house (Fogelman and Goldstein, 1976), and boys (Svalastoga, 1976). The 1966 cohort covers the entire spectrum of social classes, blacks and whites, degrees of biological risk, and residents of the inner city and suburbs, and major religious groups. It is normally distributed in terms of SES scores. The third cohort is a small group of infants on whom extensive perinatal data (e.g. Brazelton tests) were taken (Lyon-MacNeil, 1974).

The 1966-Data Set

As a whole, the data set on the 1966 cohort consists of about 900,000 pieces of data stored on tape. More analytically, the data consist of a series of individual case histories for whom information has been acquired through individualized testing and interviewing. In the last few years lack of funds has led to use of correspondence, questionnaires, and school records--techniques which work well, but are no substitute for data-taking within the families. We continue to observe selected children very closely, however. Examples are a precocious boy, and retarded twins. In the case of the boy, we have learned that Bill now attends a local school conducted by the Benedictines of Ampleforth, where his academic progress will undoubtedly accelerate still more. As balance, he plays a good game of tennis, practices the violin, and has published several articles. Recently, Bill won third prize in a national contest on Living History; all this at age twelve. The scope of our attention to the 1966 data set ranges from the statistical to the personal, with most analyses using medium-size data sets of approximately three hundred cases.

Variable Domains

Data are available in several domains. At all ages measures of height and weight are available; dummy variables and various indices have been calculated in order to describe morphology. Intellectual, and linguistic measures are major items, followed at school age by achievement and classroom data. Lesser domains are behavioral, motoric, personality and social measures. In addition, domains of data exist on home, family, parents, and siblings. In longitudinal analyses criteria at early ages become predictors at later ages. Some domains extinguish as sources of information, e.g. biological risk data from the perinatal period, and others rise, e.g. measures of home environments and readiness for school. Table 1 presents representative variables at several child-ages in five domains.

TABLE 1
REPRESENTATIVE VARIABLES BY DOMAIN AND CHILD AGE

Variable Number	Variable	Domain	Child Age
4	Weight (lb.)	Physical Development	Delivery
90	Weight (lb.)		42 Months
738	Weight (lb.)		10 Years
3	Height (in.)		Delivery
89	Height (in.)	Parental Traits	42 Months
737	Height (in.)		10 Years
117	Copy Forms		48 Months
163-170	Various Conditions		54 Months
305-307	Eye Dominance	Preschool Experiences	72 Months
347	Resting Pulse		78 Months
58	Ponderal Index		30 Months
338	N. Pregnancies		72 Months
424	N. Achievement Score	Life Factors	84 Months
80	Maternal Employment-Site		36 Months
83-86	Family Problems Scale		36 Months
243	Father Present in Family		60 Months
742	Parents' Occupation	Child Attainment	10 Years
690	Marital Status		10 Years
140	Type of Preschool		48 Months
222	STIM Score		60 Months
237-242	Television Viewing Habits	Child Attainment	60 Months
232	Type of Dwelling		60 Months
319	Life Changes Scale		72 Months
231	N. Rooms in Home		60 Months
693	Religious Affiliation	Child Attainment	10 Years
118	Boehm Concepts Score		48 Months
300	Wepman Auditory Score		72 Months
407	WRAT Reading Score		84 Months
628-682	School Behavior Profile (items)	Child Attainment	9 Years
624	Special Curricular Services		10 Years
1169-1205	Iowa Test of Basic Skills (Scores)		12 Years

Taking Data

The differences between project and programmatic work are evident in the decision to take data in a fashion which increases rather than decreases the possibilities of taking data successfully on subsequent occasions. This proposition has led to strategies of data-taking; for example, all testing has been done with a span of time consonant with the attention span of a child at given ages. Data have been taken in the home in order to minimize obstacles to cooperation, and by appointment. In several thousand interviews in the home we have had only one complaint from a parent, and that was due to a misunderstanding. Rapport between family and project staff has been extraordinary, and in many instances the same caseworker has seen a child pass several milestones of development. Tragedy as well as happiness has been shared in the case of illness, and even death. In all instances caseworkers have been matched with families by race, and also by sex when possible. In a few instances the writer directed caseworkers to cease entering homes because of the possibility of violence within the home or within the neighborhood.

It is important to note that we have exchanged information with families, not simply taken it. Routinely, we have distributed information on child care, Headstart, and free immunization. The writer continues to write up developmental histories when children have school or medical problems. While not providing experimental treatments, no field-centered research program can avoid an obligation to be a good citizen within its community.

Procedures

Data-taking was most intense in the period of semi-annual testing. At all ages it was preceded by careful choice of data domains; once the relevance of domains was established as a priority for a given period of development, measures were applied to all children at their birthday, or as members of the other half of the cohort, in the summer, six months later. Caseworkers were trained to give tests on children not in the cohort; as an example, our use of Piagetian tests failed to survive the practice period since the tasks seemed quite ambiguous to the children. Cases were assigned to two teams each containing three or four caseworkers. Within each group caseworkers dealt with families they had studied previously. Supervising caseworkers checked data, and all staff met once a week. Scores were transferred to magnetic tape and are kept under security in another city. All procedures have been scrutinized by the University's Committee on the Rights of Human Subjects in Experimentation. An interesting aspect of procedures has been norming. Typically, we have had seven hundred scores on any subtest, a number well beyond that available when published norms were generated; consequently, we have been self-norming. In the case of e.g. WPPSI Vocabulary our scoring has been much more sophisticated than the published guidelines due to the greater range of responses.

In order to keep up a good relationship, investigators may gather identical descriptive material, administer three scales to all subjects, but give as the fourth measure one or the other of two scales each of which is at a lower level of priority than the other three. Such scales may not be given at all to children who tire easily. The issue becomes, accordingly, not how many complete records, "pure" sets as they are sometimes called (Tanner and Israelsohn, 1963) but how many cases have a particular configuration of variables. In the St. Louis Baby Study the number of probands studied annually has fluctuated. At birth, six, twelve and twenty-four months data were taken on the number shown in Table 2. At age three the cohort was split into two groups, with the second tested in the summers six months later, until age eight when annual study was resumed at the birthday and, subsequently in late Spring at the end of the school year.

In the several years since major funding ended we have obtained limited data on as many as three hundred and fifty children. The overall picture is one of rise and fall in success at contacting children and their families, not in their cooperativeness, which has been uniform in rich and poor.

Mobility as a source of attrition is quite interesting. We have followed families to Europe and Asia with no loss of data. The real challenge has been the poor family which covers its tracks when moving across the street, and is protected by neighbors. Some of our best family contacts have been made through the corner bar and grocery store. In other cases contact has been terminated by the Study because homes were too unsafe for caseworkers--even husky six-foot males! In this regard, we add danger to Eichorn's (1973) list of "...apocalyptic horsemen of Longitudinal studies--death, disability, distance, and disinclination" (p. 303).

Unexplored Problems

Our research to date has given us a picture of research as a complex of data elements subventing a variety of narrow, but inter-related research projects. Obviously, some problems are more salient than others, and topics are left unstudied for reasons such as time and resources. It may be interesting to know about questions not explored, and the reasons. Nutrition is a topic of great interest in childhood. On close inspection, study of nutrition turned out to have enormous problems of procedure and validity. Much of the literature in this area deals with starvation, while the problem in St. Louis is occasional mis-nutrition rather than the degree of malnutrition leading to marasmus and kwashiorkor. In addition, the methodologies for studying nutrition leave much to be desired. Records of what is on the table at meal times appear to be retrospective, and also tell us nothing about food ingested away from the table or on the street. Even blood studies, whose invasive nature we have precluded, are quite time-dependent, and are less valuable than one might imagine. Somatic measures, other than caliper studies which require some removal of clothes and touching, are weak as measures of nutrition, although there are promising innovations

TABLE 2
SIZE OF THE COHORT THROUGH THE PHASE OF SEMI-ANNUAL TESTING

TIME	N	COMMENTS
Birth (1966-67)	1008	(originally larger, but some withdrawal of cooperation while probands still in hospital, due to fathers' resistance)
6 Months	693	(ineffective tracing plus loss of false cooperatives)
12 Months	580	(followed by small study N=81 at 18 months)
24 Months	580	(better tracing procedures staff replacements)
30 Months	223	(N=223, largely due to financing crisis)
36 Months	380	
42 Months	376	(cohort split into birthday and summer groups)
48 Months	421	
54 Months	404	(total at age four years = 825)
60 Months	414	
66 Months	392	
72 Months	396	(total at age six years = 776)
78 Months	370	
84 Months	284	(reduced funding, children in school, complicated by uncooperative school officials)
↓		
13 Years	223	(typical annual sample of accessible and cooperative families. Large amount of educational and aptitude data retrieved from school archives on another subsample of children in the same span of years.)

(Babu and Chuttani, 1979; Waterlow et al., 1977).

We have a good deal of data on preschool programs and television habits (Jordan, 1972). Professor Steven Spaner (1976) has explored these data to some extent, but not since data on elementary school performance were gathered. With time we hope to explore this corpus of data.

Future Possibilities

At the time of writing the cohort stands on the threshold of adolescence. In a number of cases girls have reached menarche, frequently introduced by saying that Mary is "...quite a little lady..." Obviously, adolescence itself is a major possibility for research. Within the development of adolescents there are the shoals which threaten young people in the relatively brief span of transition from childhood to adulthood. Delinquency is a phenomenon which will soon become discernible; it is also a problem on which cohort research can shed light in a unique way (Wolfgang, Figlio, Sellin, 1972; Wadsworth, 1979). It would be naive to say that drugs will arise as a problem, since they sometimes impinge on the lives of children before adolescence. However, it seems likely that the drug scene will be more salient in adolescence. Academic achievement and its antecedents in home and school, and social background will increase in importance. Our data on preschool status as bright and dull, rich and poor, will pose interesting opportunities for research. Undoubtedly, some of our teenagers will become parents, and the appraisal of that predictable event will merit a re-visit, recalling that some members of the 1966 cohort are themselves the product of teen pregnancies. That unhappy state of affairs will have the effect of creating a second generation of babies for the St. Louis Baby Study and would provide data on a third generation. Our data set, ad hoc, constitutes a basis for generating hypotheses about adolescent conception as a dependent rather than independent variable. Finally, the teen years will see the end of formal schooling. Both patterns of school learning and of achievement will be important topics for study in light of the antecedent personal histories.

Our comments thus far have emerged from the data set on St. Louis Children. In an effort to provide a basis for comparisons we have established a World Data Set. It consists of data from a variety of sources on 147 countries around the world. Data elements describe demographic, health, economic, and educational data. Studies have started on the comparative circumstances of childhood with particular attention to the Third World.

References

- Anderson, W. W. & Grant, R. W. Political socialization among adolescents in school--a comparative study of Barbados, Guyana, and Trinidad. Social and Economic Studies. 1977, 26, 217-233.
- Babu, D. S. & Chuttani, C. S. Anthropometric indices independent of age for nutritional assessment in children. Journal of Epidemiology and Community Health. 1979, 33, 177-179.
- Bordlay, J. E., Brookhouser, T. E., Hardy, J., & Hardy, W. G. Prenatal Rubella. Acta Oto-Laryngologica. 1968, 66, 1-9.
- Bottenberg, R. & Ward, J. Applied multiple linear regression, Report AD 413128 Aerospace Medical Division, Lackland Air Force Base, Texas, 1963.
- Broman, S. H., Nichols, P. L. & Kennedy, W. A. Preschool IQ: prenatal and early developmental correlates. Hillsdale, New Jersey, Lawrence Erlbaum Associates, 1975.
- Bryant, F. B., & Wortman, P. M. Secondary analysis: the case for data archives. American Psychologist. 1978, 33, 381-387.
- Cohen, J. Multiple regression as a general data - analytic system. Psychological Bulletin. 1968, 70, 426-433.
- Cromwell, R. E., Vaughn, C. B., & Mindel, C. E. Ethnic minority research in an urban setting: a process of exchange. American Sociologist. 1975, 10, 141-150.
- Douglas, J. W. B., Ross, J. M. & Simpson, H. R. All our future. Panther Books, 1971.
- Eichorn, D. H. The Berkeley longitudinal studies: continuities and correlates of behavior. Canadian Journal of Behavioral Science. 1973, 5, 297-320.
- Fogelman, K. Educational and career aspirations of sixteen-year-olds. British Journal of Guidance and Counselling. 1979, 1, 42-56.
- Fogelman, K. & Goldstein, H. Social factors associated with changes in educational attainment between 7 and 11 years of age. Educational Statistics. 1976, 2, 95-109.
- Franden, O. Research report #11: further notes on young leftists. Project Metropolitan, Stockholm University, 1979.

Goldstein, H. The design and analysis of longitudinal studies. New York and London: Academic Press, 1979.

Gourlay, R. J. & Ashcroft, M. T. Heights and weights of Bermudian school children of African and European origins. West Indian Medical Journal. 1969, 18, 40-46.

Harnischfeger, A., Mullan, J. T. & Wiley, D. E. Early childhood socialization and social class environment: final report. NIE-C-74-0032, University of Chicago, 1975.

Janson, C. C. Research report #9: the longitudinal approach. Project Metropolitan, Stockholm University, 1978.

Jordan, T. E. The developmental basis of learning disability, final report: U. S. P. H. S. Grant M-6797 (A) 1963.

Jordan, T. E. Retrospective versus prospective technique in research on learning disability. Journal of Special Education. 1967, 1, 257-265.

Jordan, T. E. Early developmental adversity and the first two years of life. Multivariate Behavioral Research Monographs. 1971, 6, 1.

Jordan, T. E. Who watches Sesame Street? In Lindgren, H. C. (Ed.), Readings in Child Development. New York: Wiley, 1972.

Jordan, T. E. The natural history of 1008 infants in the readiness years: A final report to the National Institute of Education, Contract OEG-O-70-1204(607) 1974.

Jordan, T. E. Developmental factors influencing exceptional status at age six years. Contemporary Educational Psychology. 1976, 1, 4-16. (a)

Jordan, T. E. Influences on preschool cognitive attainment. Multiple Linear Regression Monographs. 1976, 6, 1. (b)

Jordan, T. E. The mentally retarded. (4th ed.), Columbus, Ohio: Merrill Books, 1976 (c).

Jordan, T. E. On the comparability of multiple linear and interaction regression techniques. Multiple Linear Regression Viewpoints. 1978, 8, 81-89.

Jordan, T. E. Prospective longitudinal studies of retarded children. Technical report #37, St. Louis Baby Study, 1979.

Jordan, T. E. Development in the preschool years. New York: Academic Press, 1980.

Jordan, T. E. Development and Achievement in the Elementary School Years. (in preparation).

Jordan, T. E. & de Charms, R. The achievement motive in normal and mentally retarded children. American Journal of Mental Deficiency. 1959, 64, 457-466.

Jordan, T. E. and Spaner, S. D. Biological and ecological influences on development at twelve months of age. Human Development. 1970, 13, 178-187.

Jordan, T. E. & Spaner, S. D. Biological and ecological influences on development at 24 and 36 months of age. Psychological Reports, 1973, 31, 319-332.

Jordan, T. E. & Spaner, S. D. Biological and ecological influences on development at 48 and 60 months of age. Psychological Reports. 1974, 34, 119-126.

Jordan, T. E. & Spaner, S. D. Effects of age at delivery in juveniles and older females on child development. Monograph #6, St. Louis Baby Study, 1979.

Joreskog, K. Statistical models and methods for analysis of longitudinal data. In Joreskog, K. & Sorbom, D. Advances in factor analysis and structural equation models. Cambridge, Mass.: Abt Books, Inc., 1979.

Koplyay, J., Gott, C. D. & Elton, J. H. Automatic interaction detector-version 4 reference manual, Report, Air Force Systems Command, Brooks Air Force Base, Texas, 1973.

Littleton, A. A comparative analysis of the cognitive strategies employed by learners of different ethnic membership and socio-economic status. Unpublished dissertation, St. Louis University, 1969.

Lyon-McNeil, J. The refinement and prediction of a measure of infant development. Unpublished dissertation, Southern Illinois University, 1974.

McNeil, K. A., Kelley, F. J. & McNeil, J. T. Testing research hypotheses using multiple linear regression. Carbondale, Illinois: Southern Illinois University Press, 1979.

Mendez-Castellano, H. Estudio nacional de Crecimiento y Desarrollo Humanos de la Republica de Venezuela. Fundacredesa, Caracas, Venezuela, 1978.

Miller, F. J. W., Court, S. D. & Walton, W. S. The school years in Newcastle upon Tyne, 1952-1962. London: Oxford University Press, 1974.

Neligan, G., Prudham, D. & Steiner, H. The formative years: birth, family and development in Newcastle upon Tyne. London: Oxford University Press, 1974.

Owens, J. The effects of three intervention programs on the intellectual functioning of four-year old, black preschool children. Unpublished dissertation, University of Illinois, 1972.

Porter, W. T. The growth of St. Louis children. Transactions of the Academy of Science of St. Louis. 1893, 6, 263-380.

Prahl-Anderson, B. A mixed longitudinal interdisciplinary study of growth and development. New York: Academic Press, 1978.

Prokopec, M. Report on the longitudinal follow-up of Prague children (based on the method of the International Children's Centre in Paris). Proceedings of the Sixth International Congress of Anthropological and Ethnological Science. 519-527, 1960.

Quereshi, M. Y. "Patterns of Intellectual Development During Childhood and Adolescence," Genetic Psychology Monographs. 1973, 87, 313-334.

Ramachandran, K. V. et al. A survey of school children in Bombay City. London: Asia Publishing House, 1968.

Richman, A. Peak year of onset of heroin use: epidemiological fact or mystic statistic? In Senay, E., Shorty, V., & Alksne, H. (Eds.), Developments in the field of drug abuse. New York: Schenkman Publishing Company, 1975.

Rubin, R. A. & Balow, B. Measures of infant development and socio-economic status as predictors of later intelligence and school achievement. Developmental Psychology. 1979, 15, 225-227.

Salib, S. B. Life change events and adjustment in children. Unpublished Ph.D. dissertation, St. Louis University, 1975.

Silva, P. A. Learning difficulties in the preschooler: the Dunedin (N. Z.) multidisciplinary child development study. Technical Report, University of Otago, Dunedin, N. Z., 1978.

Silva, P. A. The Dunedin multidisciplinary child development study: a brief description, 1979 Staff List, Papers and Publications... University of Otago, New Zealand, 1980.

Silva, P. A. & Fergusson, D. M. Socio-economics status, maternal characteristics, child experience, and intelligence in pre-school children: a path analytic model. New Zealand Journal of Educational Studies. 1976, 11, 255-264.

Spaner, S. D. Application of multivariate techniques to developmental data: the derivation verification, and validation of a predictive model of twenty-four month cognitive development. Unpublished Ph.D. dissertation, Southern Illinois University, 1970.

Spaner, S. D. Co-variates of Sesame Street viewed by pre-schoolers, a paper presented to the American Educational Research Association Convention, San Francisco, CA, 1976.

Spaner, S. D. & Jordan, T. E. Locus of control, paper presented to the American Psychological Association, Montreal, Quebec, 1973.

Svalastoga, K. Analytic strategy in sequential research: project metropolitan revisited. Report #6, Stockholm University, 1976.

Tanner, J. & Israelsohn, W. Parent-child correlations for body measurements of children between the ages of one month and seven years. Annals of Human Genetics. 1963, 26, 245-259.

Valk, I. M. & Van Den Busch, J. S. G. Intradaily variation of the human ulnar length and short-term growth--a longitudinal study in eleven boys. Growth. 1978, 42, 107-111.

Wadsworth, M. E. J. Delinquency prediction and its uses: the experience of a 21-year follow-up study. International Journal of Mental Health, 1979, 7, 43-62.

Waterlow, J. C., Buzina, R., Keller, W., Lane, J. M., Nichaman, M. Z., & Tanner, J. M. The presentation and use of height and weight data for comparing the nutritional status of groups of children under the age of ten years. Bulletin of the World Health Organization. 1978, 55, 489-498.

Werner, E. E. & Smith, R. S. An epidemiological perspective on some antecedents and consequences of childhood mental health problems and learning disabilities. Journal of Child Psychiatry. 1979, 18, 292-306.

White, B. L., Kaban, B. T., & Attanucci, J. S. The origins of human competence: the final report of the Harvard Preschool Project. New York: Lexington Books, 1979.

Wilkins, L. T. Delinquent generations: home office research unit report 3. London: Her Majesty's Stationery Office, 1960.

Wolfgang, M., Figlio, R. M., & Sellin, T. Delinquency in a birth cohort. University of Chicago Press, 1972.

A Neuropsychiatric Follow-Up of Children in
The Collaborative Perinatal Project Population*

Patricia A. O'Connor, Ph.D.
David Shaffer, M.D.
Stephen Shafer, M.D.
Cornelis Stokman, Ph.D.
Columbia University and
New York State Psychiatric Institute

*This work was made possible by NIMH Center Grant #MH30906-01A and NIMH Education Grant #2T01 MH07715-17 and by the Research Associate Program created by the New York State Legislature, Office of Mental Health, Division of Research.

A Neuropsychiatric Follow-Up of Children in
The Collaborative Perinatal Project Population

Patricia A. O'Connor, Ph.D.
David Shaffer, M.D.
Stephen Shafer, M.D.
Cornelis Stokman, Ph.D.
Columbia University and
New York State Psychiatric Institute

Introduction

The burdens of prospective longitudinal research are well recognized. They exact a cost both in dollars and cents and in time taken from the professional careers of researchers. All too often a study may be locked into hypotheses and/or techniques which lose favor or value with time. Given the importance of longitudinal research, a number of alternative strategies have been used. One of these is the so called "piggyback" method in which a selective followup is carried out on individuals who have previously been studied in a standardized, systematic and relevant fashion. This approach may yield longitudinal data at a reasonable cost while avoiding the pitfalls of retrospection and assignment bias.

This chapter sets out to describe such a study and in particular the characteristics of the sample frame (the Collaborative Perinatal Project) upon which it is based, in the expectation that this will be useful to others with overlapping research goals. Also provided are details of our own methods.

The Collaborative Perinatal Project

Original Objectives

The Collaborative Perinatal Project (CPP) was initiated in 1950 by the National Institute of Neurological and Communicative Disorders and Stroke (NINCDS) (known then as the National Institute of Neurological Diseases and Blindness) for the purpose of investigating pregnancy wastage, defined as any abnormal outcome of pregnancy from death to minimal brain dysfunction syndromes (Berendes, 1966). The rarity of some abnormal outcomes from pregnancy, combined with the lack of specific hypotheses (Berendes, 1966), necessitated the collection of data from a large sample. This in turn required the use of a number of facilities to ensure adequate data collection within a reasonable time period. To that end, various hospitals with either a medical school or university affiliation were asked to participate in the CPP. It was decided that the venture would be collaborative, i.e., all institutions participating equally in the decision making, rather than cooperative, i.e., a single directing institution with others simply cooperating in the collection of data. It was felt that if each institution could provide substantive input, it would be more motivated to continue participation.

Institutional Participation

The 12 participating universities and their 14 affiliated hospitals are listed in Table 1, along with the number of subjects and the number of years during which subjects were followed.

Insert Table 1 About Here

Organization of the CPP

The coordinating activities of the Project were handled through NINCDS, through its Perinatal Research Branch. Decision making, however, was to lie in the hands of the Project Directors' Committee. Policy and procedural problems within a specific area were the concern of Special Problem Committees, which consisted of representatives from the discipline of concern from each institution and other appropriate professionals appointed by the Chief of the Perinatal Research Branch. The recommendations of these subcommittees were then communicated to the Project Directors' Committee which made a final decision. Specific technical questions were referred to committees of experts set up for that purpose (Berendes, 1966; Gates, 1973).

Sample Selection

Each participating center developed its own random sampling procedures and with some limitations established its own criteria for eligibility for participation in the CPP. Two criteria were used to exclude potential subjects at all centers: declared adoption donors and women who presented too late for repeated prenatal care at the respective centers (Niswander et al., 1972).

The institutions based the study in their hospital clinics with the exception of Buffalo Children's Hospital which drew its sample from private patients. Given this constitution, the CPP sample was clearly not representative of the population of the United States, a point raised by critics of the CPP (Weiss, 1966). This important issue requires further discussion.

The virtue of a "representative" sample is that it allows for more confident generalization of findings. However, no sample can be totally "representative" of all people living in all communities. The goal of generalizability can also be served by adequately describing the population and then representing findings in terms of specific population descriptors. The CPP sample lends itself to this because of the wealth of descriptors available on each subject. The sample characteristics have been defined and tests of performance recorded on standard protocols covering a wide range of demographic and constitutional variables. There is also exhaustive information available to facilitate tracing of subjects.

Standardization of Data Collection

Forms were developed by committees established for that purpose and pretesting was begun in 1957 (Berendes, 1966; Gates, 1973). Following the inauguration of formal data collection in January 1959, instances arose which ultimately necessitated some revision of protocols. This was done by committee and where possible new forms were made to be compatible with the old (Weiss, 1966).

To ensure that data would be collected in a uniform fashion, testing personnel were given special training sessions with periodic workshops and onsite visits by CPP staff (Broman, Nicholl & Kennedy, 1975). There were occasional exchanges of personnel between centers and the protocols were carefully standardized to encourage uniform data collection. The Project Directors also created a Performance Committee for the periodic review of the work of each collaborating institution (Berendes, 1966).

Data Storage

The original records remain at the participating institutions. The raw data have been recorded on microfilm and the coded data on computer tapes. Both are stored at NINCDS.

Time of Assessments

Data collection began with the initial registration of the mother in the CPP at the time of her first visit to the participating medical center.

Information was obtained at each subsequent prenatal visit with repeated physical and laboratory examinations. Labor and delivery were carefully monitored with a trained observer in attendance at all times. Information was collected by the appropriately trained person, e.g., nursery nurse, obstetrician, neurologist, psychologist, on the child at birth, in the nursery and then at subsequent, regularly specified intervals through age 8 (although only 6 of the original 14 institutions conducted the 8 year examination: Buffalo Children's Hospital, Children's Medical Center in Boston, Johns Hopkins University, University of Tennessee, University of Minnesota and University of Oregon (Broman et al., 1975)). The timing and type of assessments are given in Table 2.

Insert Table 2 About Here

Assessment of the Parent

Demographic and employment information about both parents was collected from the mother during a prenatal visit shortly after her initial prenatal registration in the CPP. Family health information was also obtained and included mother's reproductive history and medical histories of the mother, father and their respective families. Mother's IQ was measured using a specially abbreviated form of the SRA (Science Research Associates) test at the child's four year examination. Marital status, residential, employment, family income and other demographic changes were updated during the seven year examination of the child (see Table 2).

Assessment of the Child

Detailed assessment of the child began with the delivery room observation of the neonate. Apgar scores and an initial physical examination were recorded around birth. Extensive information was obtained about the delivery itself, including the use of drugs. A standard physical and neurological examination was conducted and nursery

information, such as feeding methods, medications given, respiratory abnormalities and activity level, was collected during the hospital stay. Subsequent examinations consisted of standard physical (including neurological), psychological and educational examinations. The only behavioral measures taken were those of behavior during the psychological tests given at ages four and seven (see Table 2).

Sample: The CPP at Columbia Presbyterian Hospital

Inclusion Criteria

The sample frame at Columbia Presbyterian Hospital consisted of all women admitted to the prenatal clinic between January 1959 and April 1963, except for declared adoption donors and women who had received no more than one prenatal examination. After April 1962, only patients residing in Manhattan and the Bronx were included in the sample frame to reduce the costs, as well as difficulties, in followup of the more distant registrants. At Columbia, the ratio of participants to registrants, initially 1:6, was later reduced so that the average sampling ratio was 1:4.4, yielding a total of 2,235 prenatal clinic registrants (Niswander et al., 1972).

Characteristics of the Columbia Sample

There were 2,067 live CPP births at Columbia. At the time of the seven year examination, 2,019 children had survived and 83.5% of those attended for reexamination (Gates, 1973). Overall, the loss and attrition rate for the Columbia sample was small and the rates compare favorably with loss at other centers (generally 70-80% examined).

The characteristics of the Columbia sample differed somewhat from those of the general population from which the sample was drawn (New York City and New York State). High income and high education level whites were underrepresented, as were black mothers under age 18 and black mothers with low income and educational levels. The sample included a larger proportion of blacks with a higher education and higher income than were found in the general population. However, the subjects chosen for the current study were generally comparable to subjects in the sample frame with regard to marital status, mother's age at registration and time of registration (Niswander et al., 1972).

Sample Selection

The current study is a followup of English speaking male and female CPP children born in 1962 and 1963, who are now adolescents. Of the 568 black and white children born in 1962-63 and known to have survived until 1970, a total of 471 attended for the seven year examination. This included 88 percent of the black and 77 percent of the white group.

Spanish speaking families were excluded because of their reportedly high mobility and because of noncomparability of the early psychometric measures, which were given in English.

This cohort, the youngest in the Columbia sample, was chosen because first, children born in those years would still be in school. Their teachers would thus be able to provide us with important information about their behavior and academic achievements. Second, the school register would facilitate subject tracing because the name of the child's school at age seven is recorded and can be used as a starting point to trace an individual through the school system with relatively little difficulty.

Criteria for Inclusion

Subjects were selected for inclusion in the experimental group (N=63) if they had been noted to have any one of 18 neurological signs, broadly grouped within the constructs of movements, coordination and sensory integration, at age seven. Controls (N=63) were taken from the same subject list and were matched for sex, race, closest birthdate and the absence of any neurological signs.

Instruments

Psychiatric Evaluation

Goal

The primary purpose of the assessment is to ascertain the level of functioning of each adolescent and where appropriate to assign a psychiatric diagnosis, using DSM III criteria (American Psychiatric Association, 1980). To that end, three sources of information about the adolescent have been used. These are the adolescent, one parent (preferably the one with whom the adolescent lives and usually the mother) and several school teachers.

The assessment also seeks to obtain a picture of the pattern of the adolescent's relationships, an interval medical history (since age seven) and an account of a variety of social and family variables which might be significant in determining psychiatric dysfunction.

Direct Evaluation of the Adolescent

A semistructured interview format is used to assess the current mental status of the adolescent and to explore the range and adequacy of current social relationships. The interviewers, as well as the cognitive and neurological testers, are blind with respect to subject status, i.e., experimental versus control group. The interview, the Columbia Psychiatric Interview for Children and Adolescents, Adolescent Form (Shaffer & O'Connor, 1980), assesses through direct questions the presence and severity of psychosomatic, affective and psychotic symptoms, phobias and antisocial behavior and delinquency. Detailed information is obtained about peer relationships, both same and other sex friendships and sexual relationships, and about the intensity and quality of certain family relationships. The instrument has been developed from other instruments. Summary ratings of school behavior, peer relationships (hostility to and from peers and isolation) and quality of adolescent to parent relationships along the dimensions of warmth, intimacy and distance have also been developed for the interview.

The interview for adolescents takes approximately one hour and a half to administer. Subject compliance has been good; in all cases interviewees have willingly answered all parts of the interview.

Evaluation of the Adolescent through Informant (Parent)

A semistructured interview format is also used in obtaining information from informants. The interview, the Columbia Psychiatric Interview for Children and Adolescents, Parent Form, is designed to mirror the information obtained from the adolescent with the following additions and expansions:

1. An interval history section has been included to obtain information comparable to the health and socioeconomic variables that were part of the CPP protocols during the first seven years of the developmental study.
2. The section on substance and alcohol use has been expanded to include an inquiry into the substance and alcohol use of parents.
3. Additional information is also obtained on intrafamily relationships and the marital relationship is explored in some detail, based on a modified version of the Maudsley Hospital interview (Rutter & Brown, 1966).

4. Brief social and medical histories are obtained for each of the siblings and for the parent not being interviewed. Demographic information is also obtained from the parent.

The interview with the informant takes approximately two hours to administer. Subject compliance with the interview has been good; none of the interviewees have refused to answer any parts of the interview.

Assessment of the Parent

Overall psychiatric functioning in the parent being interviewed is assessed using the General Well Being Scale (Dupuy, 1974). This 26 item index gives a measure of general functioning but does not allow for an assignment of diagnosis. However, cutoff points for the presence of pathology have been established (Dupuy, 1974). Information is also obtained about medical and/or psychiatric contacts of each of the parents during the previous five year period, and whole life police contacts.

Interviewer Training

It has been similar for both adolescent and informant interviewers. Training begins with meetings for each group of interviewers (adolescent and informant). The respective instruments are discussed in detail, using videotapes of pilot interviews to illustrate interviewing techniques. Rating definitions are discussed extensively during these meetings so that all interviewers, regardless of training and experience, will use the same criteria in giving any behavior a positive score.

Each interviewer is expected to do three pilot interviews which are then viewed by the training supervisor (POC), the Principal Investigator (DS) or both with the entire group of interviewers (adolescent or informant). Pilot interviews are continued until the DS and/or training supervisor feel that he/she is sufficiently prepared to begin experimental interviewing.

Weekly meetings of the DS and the interviewers, adolescent and informant separately, are held for the purpose of reviewing all positive or difficult ratings.

Assessment of Reliability

The reliability of the assessment of psychiatric functioning can be examined in two ways: intrarater reliability and intrasubject or interrater reliability. Our method of assessing both of these has been to have interviewers of each of the informants (adolescent and parent) watch and rate videotapes of their respective interviews (each routinely tapes every fifth interview for reliability purposes) so that all

interviewers watch interviewer X with subject A to obtain ratings on the measure used. This procedure has encompassed the assessment of reliability (by providing intrasubject and interrater measures), as well as additional training (by encouraging consistency among interviewers through the experiencing of each other's techniques), which should also serve to increase reliability.

Evaluation of the Adolescent Through Teachers

The Conners Questionnaire (Conners, 1980) has been widely used to provide teacher ratings of school behavior (although typically for children of elementary school age). It was decided that since adolescents participating in the study would no longer be in self contained classrooms with a single teacher, three teachers in primary subject areas (English, math and social studies) would be asked to rate the student. The method of teacher recruitment was an initial contact of a researcher on the Project with the principal of the school where the adolescent was in attendance. The principal then arranged for the questionnaire packets to be distributed to three teachers and requested their cooperation. In many cases, two or even three contacts with the school have been necessary to ensure teacher compliance.

The Conners Scale was amplified with seven additional "positive" (or "desirable behavior") items, considered to be logically incompatible with seven items already present in the scale. Responses from approximately 20 adults were used to determine which pairs of items could be considered logically inconsistent. The rationale behind the inclusion of these items was, somewhat in the manner of a "lie" scale, that it might then be possible to assess the care or consistency with which teachers were filling out the Conners Scale. One would expect a negative correlation between the 39 original Conners items and the seven additional "positive" items. It was decided to insert the additional items adjacent to their paired items to focus the teachers' attention in such a way that they would be encouraged to evaluate each item and give as thoughtful a response as possible.

A Supplemental Teacher Questionnaire, devised by the Project staff, was sent to each teacher along with the Conners Teacher Questionnaire. It provides information on how long and in what context(s) the teacher has known the pupil being rated; it asks for the teacher's opinion as to whether or not the pupil has any academic or social/behavioral problems for which the teacher feels special help is needed; it asks whether the teacher or anyone else to the teacher's knowledge is providing special help, and if so what kind; it asks the teacher to rate, on a continuous scale from 0 to 100, the pupil's overall adjustment compared to other pupils of comparable age in the teacher's experience; and it requests information about the teacher, including teaching experience.

As its name implies, this questionnaire is intended primarily to supplement information collected from other sources, including the interviews, and is expected to aid in interpreting the data from the Connors.

Finally, the New York City Board of Education has given the Project a letter of consent permitting access to the Cumulative Pupil Record of each adolescent. This letter is sent, along with signed consents from the adolescent and parent, to the principal of the school in which the adolescent is a student. A form is being developed which will be used to abstract relevant information from the Cumulative Pupil Record, e.g., number of absences, grades repeated, standardized achievement test scores, rapid versus steady decline in school achievement.

Neurological Assessment

Goals

The neurological assessment is designed to fulfill two purposes. The first is a simple reevaluation of the neurological status as determined at the age seven examination, and the second is the richer measurement of the constituents of that examination. In the forms used at age seven, items were rated as "definitely present," "suspect" or "absent" in some cases, and only dichotomously as "normal" or "abnormal" in others. The goal in the present study was to elaborate the "definitely present" and "abnormal" ratings to allow for more exact determinations of what constitutes soft neurological signs and for greater variance in determining the clinical significance and applicability of such signs.

Instrument

The neurological examination consists of an extensive standardized assessment of those neurological signs measured at age seven and subsequently used as the criteria for selection of our experimental subjects. Criteria for measuring each sign are set forth as mentioned above with the aim of developing scales to quantify the various soft signs. For example, dysdiadochokinesis can be assessed with smoothness and rhythmicity ratings on each of the extremities, permitting the scale value for an individual subject to represent more precisely the extent to which the sign is present.

In addition, the examination includes the assessment of minor physical anomalies, cranial nerves, laterality, musculoskeletal system, reflexes and interval medical history (since age seven) and standard physical measurements (height, weight, head circumference). At the end of the examination, the subject is given an overall rating of neurological function based on the clinical judgment of the tester. The function of this rating is merely to permit a global assessment and it does not presume to replace the more exact measurement of each individual sign.

The order of items in the neurological examination is the same for all subjects with the exception that the presentation of certain rapid movement items is counterbalanced.

Assessment of Reliability and Validity

The reliability of the neurological examination was tested on three groups other than the study population: (1) 10 adolescent inpatients at a nearby psychiatric center; (2) young adult staff on the Project; and (3) 15 normal adolescents. The 10 inpatients were given parts of the examination twice in a four week interval. Simultaneous ratings were obtained by two or three of the neurological testers. Comparisons were made inter and intrarater. Rank correlation coefficients (r_s) for total number of positive ratings revealed a high degree of agreement between raters for both test ($r_s=.77$) and retest ($r_s=.93$) as well as intrarater test-retest reliability (rater 1: $r_s=.92$ and rater 2: $r_s=.77$; rater 3 did not test enough subjects for analysis).

Cognitive Assessments

Goals

The primary purpose of the cognitive assessment is to establish the current full scale IQ and the level of competence attained by each subject in the areas of reading, spelling and math in order to determine which adolescents have learning difficulties. Measures of attention have increasingly gained recognition as correlates of and possibly intervening variables in pathology and are therefore included in the cognitive assessment. Unfortunately, no measures of attention are available from the age seven examination.

Instruments

The cognitive battery is comprised of the following tests, most of which have been standardized on other populations:

1. IQ: Wechsler Adult Intelligence Scale (Wechsler, 1955). Where the subject was less than 15 years, nine months at the time of testing, the Wechsler Intelligence Scale for Children -- Revised Form was administered with administration of the WAIS in a retest session following the subject's 16th birthday.

2. Reading and Achievement:

- a. Wide Range Achievement Test (Jastak & Jastak, 1965), Reading Recognition subtest.

b. Peabody Individual Achievement Tests (Dunn & Markwardt, 1970): Mathematics, Reading Comprehension and Spelling subtests.

3. The Continuous Performance Test: the Continuous Performance Test (CPT) has two modes of operation: (1) the nature of stimulus presentation and (2) the rate of stimulus presentation. In the first, the subject is required to press a button, either (a) when the letter "X" appears in a series of nine other letters, approximately one per second, or (b) when the letter "X" appears preceded by the letter "A". In the second mode, the rate of presentation may be either (a) static so that the interstimulus level (ISI) remains the same throughout the testing session or (b) dynamic so that the ISI is a function of the proportion of correct responses of the subject. In the case of successive correct responses, the ISI decreases by a small amount (becomes more difficult); for errors of omission, the ISI increases by a small amount (becomes easier).

Previous research with adolescents (e.g., Crosby, 1972; Kornetsky et al., 1959; Mirsky & Orren, 1977; and Rosvold et al., 1956) and our pilot studies confirmed our initial judgment that the static mode, regardless of the nature of the stimulus presentation, would be too easy for our sample. Thus the dynamic mode was selected for the rate of stimulus presentation.

Our pilot studies also showed that the AX stimulus presentation was a more difficult task than the single "X" presentation and thus would be more likely to provide a better discrimination between normal and clinical groups (also Buchsbaum et al., 1978).

Errors of commission provide an additional measure which again discriminated between clinical and normal groups in our pilot studies (a finding also reported by Buchsbaum et al., 1978).

Thus, the dynamic mode with the "AX" presentation is used for the Continuous Performance Test, giving the individual and mean interstimulus intervals and number of errors of commission as well as the number of correct responses as measures of attention with this task.

4. Reaction Time Paradigm: the Sutton-Hakerem-Zubin-Portnoy (1961) reaction time paradigm is built into the design of the Sunrise CPT equipment, giving a tone or light signal to which the subject must respond by removing the finger as quickly as possible from a capacitance switch. The testing begins with a warning tone (1.9 kHz) at which time the subject places the index finger of the preferred hand on the switch. The subject responds to an imperative signal (either light or tone of 2.9 kHz) so that the pattern of presentation is warning signal, preparatory interval, imperative signal. The preparatory intervals, .5, 1.0,

2.0, 5.0 or 10 seconds are varied randomly. There are six blocks of 30 trials each in which the preparatory intervals and the imperative signals have been systematically varied. The data are output as individual reaction times for each presented signal and as means and standard deviations for each block of responses, separately and combined, for the visual and auditory signals.

5. Trail-making Tests: (Reitan, undated) also known as the Spoke Test, this measure of attention first requires the subject to connect circles containing consecutive numbers (A Form) and then circles containing consecutive numbers alternative with consecutive letters (B Form). The final sum is the total number of errors or misordered circles for both forms.

The trail-making test is in principle a very useful attention measure since it requires the subject to override pressures from two highly overlearned skills. It has also been shown to differentiate between children at risk for schizophrenia and their controls (Asarnow et al, 1978), and between a small group of adolescent inpatients in a psychiatric facility and controls in our piloting studies.

6. Cancellation Tests: (Diller & Weinberg, 1970; Weinberg et al., 1977) the Cancellation Test requires the subject to cross out each of 140 triads on a single page that matches the target triad at the top of the page. The triads consist of letters, numbers and geometric figures and a subject is presented with eight pages of each kind of triad (total pages per subject = six). Ten percent (n=14) of the triads per page match the target.

Initial piloting established that the triads of geometric shapes required the longest time for completion and resulted in the greatest number of errors, both of omission and commission. Thus the order of presentation of the pages of triads was modified to equalize the average position of the three different types of triads (letters 1, numbers 1, geometrics 1, geometrics 2, numbers 2, letters 2) and fixed for all subjects.

Errors of commission are summed with errors of omission across the six pages of triads to yield a final score representing the total number of errors. Total time may also be used as a measure of attention in this task.

7. Paired Associates Task: (Wechsler, 1945/1973; Stone & Wechsler, 1946/1974) the Paired Associates Task is included in the cognitive battery as a measure of learning and recall. The Wechsler Memory subtest with one modification, as described below, was selected for use with the classical mode of administration, i.e., 10 pairs of words are read to

the subject once, followed by repeated reading of the word of each pair until the subject can give the correct match to each presented word. Criterion is two perfect recitations of the entire list and all subjects are tested until criterion, with a maximum of 30 trials.

In piloting this measure, the lists as constructed by Wechsler were found to be too easy for adolescents. Therefore one of Wechsler's easy pairs (normally five easy pairs) was replaced with a hard pair (normally five hard pairs) from an equivalent list.

The initial administration of this task repeated, presentations until criterion is reached, provide a measure of learning. Recall is assessed by the readministration of the pairs, again to a criterion of two perfect recitations, approximately three to three and a half hours after the first testing session. The data collected are the number of trials to criterion for learning and for recall and the time required for each testing session.

Conclusions

The CPP represents a massive investment in the documentation of early childhood neurological and psychological development in a large sample unselected for any particular outcome.

Although it is not a "representative" sample, the initial selection procedures and the very extensive documentation of the individual characteristics, medical and social experiences and developmental progress of its subjects make the CPP an unequalled resource from which to study a number of problems. A large amount of literature already exists which illustrates the range of the study's usefulness in natural history research. The effects of in utero exposure to drugs have been studied by Erhardt et al., 1970; Hartz et al., 1974; Shapiro et al., 1976; and others. The experience of medical complications at delivery and illness in the neonatal period have also been extensively examined (see Naeye, 1975; Niswander et al., 1975; Fisch et al., 1975; Gordon et al., 1973; Hardy, 1973), as has illness or exposure to lead in early childhood (de la Burde & Choate, 1975), and the antecedents of developmental characteristics and disorders (Torry et al., 1975; Broman et al., 1975; Fisch et al., 1976; Colligan, 1974; Babson & Henderson, 1974). Another purpose for which the study population can be used is to examine the natural history or stability of a particular characteristic (Rosenblith, 1973; Shaffer et al., 1980), or as we are doing in the current Columbia study, to examine the relationship between such a condition or characteristic and later functioning. The documentation of parental characteristics allows for its use as a sample frame from which to study the offspring of patients with particular characteristics such as the Rieder et al. study (1975) of the offspring of schizophrenics.

The suitability of a population for such piggyback studies depends to a large extent on the ease with which the subjects can be traced. We have been able to demonstrate that given the very extensive identifiers available in the CPP study records, the current whereabouts of subjects -- at least those who are still of school age -- can be carried out without undue expense. In our case, we were able to trace the whereabouts of 97 percent of a sample of 180 children (including 54 black females) using the part time services of a series of graduate students over a two year period.

Our own study has added yet a further body of information to the life records of this cohort of individuals. In this chapter, we have gone to some length to document the data that we have collected on the sample, in the event that it may stimulate yet further research into the development of the natural history of psychiatric disorders.

References

- Adams, R.M., Kocsis, J.J. & Estes, R.E. Soft neurological signs in learning disabled children and controls. American Journal of Diseases of Children, 1974, 128: 614-618.
- American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders. Washington D.C., 3rd edition, 1980.
- Asarnow, R.F., Steffy, R., MacCrimmon, D.J. & Cleghorn, M.D. An attentional assessment of foster children at risk for schizophrenia. In L.C. Wynne, R.L. Cromwell, S. Matthysse (Eds.), The Nature of Schizophrenia. New York: John Wiley and Sons, 1978.
- Babson, S.G. & Henderson, N.B. Fetal undergrowth: Relation of head growth to later intellectual performance. Pediatrics, 1974, 53: 890-894.
- Berendes, H.W. The structure and scope of the Collaborative Project on cerebral palsy, mental retardation, and other neurological and sensory disorders of infancy and childhood. In S.S. Chipman, A.M. Lilienfeld, B.G. Greenberg & J.F. Donnelly, Research Methodology and Needs in Perinatal Studies. Springfield, Ill.: Charles C. Thomas, 1966.
- Boshes, B. & Myklebust, H.R. A neurological and behavioral study of children with learning disorders. Neurology, 1964, 14: 7-12.
- Broman, S.H., Nichols, P. & Kennedy, W.A. Preschool I.Q.: Prenatal and Early Developmental Correlates. Hillside, N.J.: Erlbaum Pub., 1975.
- Buchsbaum, M.S., Murphy, D.L., Coursey, R.D., Lake, C.R. & Ziegler, M.G. Beta-hydroxylase and attention in a biochemical high risk sample. Journal of Psychiatric Research, 1978, 14: 215-244.
- Colligan, R.C. Psychometric deficits related to perinatal stress. Journal of Learning Disabilities, 1974, 7: 154-160.
- Conners, C. A teacher rating scale for use in drug studies with children. American Journal of Psychiatry, 1969, 126: 152-156.
- Crosby, K.G. Attention and distractibility in mentally retarded and intellectually average children. American Journal of Mental Deficiency, 1972, 77(1): 46-53.
- De La Burde, B. & Choate, M.L. Early asymptomatic lead exposure and development at school age. Journal of Pediatrics, 1975, 87: 638-642.
- Diller, L. & Weinberg, J. Evidence for accident-prone behavior in hemiplegic patients. Archives of Physical Medicine, 1970, 51: 358-363.
- Dunn, L.M. & Markwardt, F.C. Peabody Individual Achievement Test, Circle Pines, Minn.: American Guidance Service, Inc., 1970.
- Dupuy, H.J. Utility of the National Center for Health Statistics' General Well-Being Schedule in the assessment of self-representations of subjective well-being and distress. Paper presented at the National Conference on the Evaluation of Drug, Alcohol and Mental Health Programs, 1974.
- Endicott, J.S., Spitzer, R.L., Fleiss, J. & Cohen, J. The Global Assessment Scale. Archives of General Psychiatry, 1976, 33: 766-771.
- Erhardt, C.L., Abramson, H., Pakter, J. & Nelson. An epidemiological approach to infant mortality. Archives of Environmental Health, 1970, 20: 743-757.
- Fisch, R.O., Bilek, M.K., Horrobin, J.M. & Chang, P.N. Children with superior intelligence at 7 years of age: A prospective study of the influence of perinatal medical and socioeconomic factors. American Journal of Diseases of Children, 1976, 130: 481-487.
- Fisch, R.O., Bilek, M.K. & Ulstrom, R. Obesity and leanness at birth and their relationship to body habitus in later childhood. Pediatrics, 1975, 56: 521-528.
- Gates, M.J. Final report: Collaborative perinatal study of National Institute of Neurological Diseases and Stroke. New York: Columbia University College of Physicians and Surgeons, 1973.
- Gordon, M., Rich, H., Deutschberger, J. & Green, M. The immediate and long-term outcome of obstetric birth trauma. American Journal of Obstetrics and Gynecology, 1973, 117: 51-56.
- Gould, M.S., Wunsch-Hitzig, R., Dohrenwend, B.P. Formulation of hypotheses about the prevalence, treatment, and prognostic significance of psychiatric disorders in children in the United States. In B.P. Dohrenwend & B.S. Dohrenwend, M.S. Gould, B. Link, R. Neugebauer & R. Wunsch-Hitzig, Mental Illness in the United States: Epidemiological Estimates, New York: Praeger Publications, 1980.
- Hardy, J.B. Birth weight and subsequent physical and intellectual development. The New England Journal of Medicine, 1973, 289: 973-974.

- Hartz, S., Rosenberg, L., Heinonen, O.P., Slone, D. & Shapiro, S. Epidemiological screening of multiple risk factors in relation to multiple congenital malformations: A statistical approach. Teratology, 1974, 10: 311.
- Hertzog, M.A. & Birch, H.G. Neurologic organization in psychiatrically disturbed adolescent girls. Archives of General Psychiatry, 1966, 15: 590-599.
- Hertzog, M.A. & Birch, H.G. Neurologic organization in psychiatrically disturbed adolescents. Archives of General Psychiatry, 1968, 19: 528-537.
- Jastak, J.F. & Jastak, S.R. The Wide Range Achievement Test. Guidance Associates, 1526 Gilpin Avenue, Wilmington, Del., 1965.
- Kandel, D., Singer, E. & Kessler, R. The epidemiology of drug use among New York State high school students: Distribution, trends and change in rates of use. American Journal of Public Health, 1976, 66: 43-53.
- Kenny, T.J. & Clemmens, R.L. Medical and psychological correlates in children with learning disability. The Journal of Pediatrics, 1971, 78(2): 273-277.
- Kestenbaum, C.J. & Bird, H.R. A reliability study of the Mental Health Assessment Form for school age children. Journal of the American Academy of Child Psychiatry, 1978, 17(2): 338-347.
- Kornetsky, C., Mirsky, A.F., Kessler, E.K. & Dorff, J.E. The effects of dextro-amphetamine on behavioral deficits produced by sleep loss in humans. The Journal of Pharmacology and Experimental Therapeutics, 1959, 127: 46-50.
- Marcus, J. Neurological and Physiological characteristics of the children of schizophrenic parents. Paper presented at the Seventh Congress of the International Association of Child Psychiatry and Allied Professions. Jerusalem, 1970.
- Mirsky, A.F. & Orren, M.M. Attention. In L.H. Miller, C.A. Sandman & A.J. Kasten (Eds.), Neuropeptide Influences on the Brain and Behavior, New York: Raven Press, 1977.
- Mosher, L.R., Pollin, W. & Stabenau, J.R. Identical twins discordant for schizophrenia. Archives of General Psychiatry, 1971, 24: 422-430.

- Naeye, R. Amniotic fluid infections and premature rupture of the membranes, epidemiologic differences. Pediatric Research, 1975, 9: 298.
- Niswander, K.R., Gordon, M., Berendes, H.W., Blac, W.A., Clifford, S.H., Douglas, R.G., Drage, J.S., Hardy, J.B., Jackson, E.C., Lipsitt, L.P., Shapiro, S. & Weiss, W. The Women and Their Pregnancies. The Collaborative Perinatal Study of the National Institute of Neurological Diseases and Stroke, DHEW Publication No. (NIH) 73-379, 1972.
- Niswander, K.R., Gordon, M. & Drage, J.S. The effect of intrauterine hypoxia on the child surviving to 4 years. American Journal of Obstetrics and Gynecology, 1975, 121: 892-899.
- Peters, J.E., Roming, J.W. & Dykman, R.A. A special neurological examination of children with learning disabilities. Developmental Medicine and Child Neurology, 1975, 17: 63-75.
- Quitkin, F., Rifkin, A. & Klein, D.F. Neurologic soft signs in schizophrenia and character disorder. Archives of General Psychiatry, 1976, 33: 845-853.
- Reitan, R.M. Trail Making Test: Manual of Administration and Scoring. Neuropsychology Laboratory, Tucson, Ariz., undated.
- Rie, E.D., Rie, H.E., Stewart, S. & Rettemnier, S.C. An analysis of neurological soft signs in children with learning problems. Brain and Language, 1978, 6: 32-46.
- Rieder, R.O., Rosenthal, D., Wender, P. & Blumenthal, H. The offspring of schizophrenics: Fetal and neonatal deaths. Archives of General Psychiatry, 1975, 32: 200-211.
- Rochford, M.M., Detre, T., Tucker, G.J. & Harrow, M. Neuropsychological impairments in functional psychiatric disease. Archives of General Psychiatry, 1970, 22: 114-119.
- Rosenblith, J.F. Prognostic value of neonatal behavioral tests. Early Child Development and Care, 1973, 3: 31-50.
- Rosvold, H.E., Mirsky, A.F., Sarason, I., Bransome, E.D. & Beck, L.H. A continuous performance test of brain damage. Journal of Consulting Psychology, 1956, 20: 343-350.
- Routh, D.K. & Roberts, R.D. Minimal brain dysfunction in children: Failure to find evidence for a behavioral syndrome. Psychological Reports, 1972, 31: 307-314.

Rutter, M. & Brown, G.W. The reliability and validity of measures of family life and relationships in families containing a psychiatric patient. Social Psychiatry, 1966, 1: 28-53.

Rutter, M. & Graham, P. The reliability and validity of the psychiatric assessment of the child: I. Interview with the child. British Journal of Psychiatry, 1968, 114: 563-579.

Shaffer, D. "Soft" neurological signs and later psychiatric disorder -- A review. Journal of Child Psychology and Psychiatry, 1978, 19: 63-65.

Shaffer, D. & O'Connor, P.A. Columbia Psychiatric Interview for Children: Parent Form and Adolescent Form. New York Psychiatric Institute, New York, 1980.

Shaffer, D., Stonman, C.S., O'Connor, P.A., Shafer, S., Barmack, J.E., Hess, S. & Spalten, D. Early soft neurological signs and later psychopathology. In N. Erlenmeyer-Kimling & B.S. Dohrenwend (Eds.), Lifespan Research on the Prediction of Psychopathology. (tentative) (in press).

Shapiro, S., Hartz, S.C., Siskind, V., Mitchell, A.A.

Spitzer, R.L. & Endicott, J. Schedule of Affective Disorders and Schizophrenia (SADS). New York: Biometrics Research, New York State Psychiatric Institute, 3rd edition, 1977.

Stine, O.C., Saratsiotis, J.M. & Mosser, R.S. Relationships between neurological findings and classroom behavior. American Journal of Diseases of Children, 1975, 129: 1036-1040.

Stone, C.P. & Wechsler, D. Wechsler Memory Scale Form II. In D. Wechsler & C.P. Stone, Wechsler Memory Scale: Manual. New York: The Psychological Corporation. (Published originally as an alternate form of the Wechsler Memory Scale by C.P. Stone, J. Girdner & R. Albrecht, The Journal of Psychology, 1946, 22: 199-206.

Sutton, S., Hakerem, G., Zubin, J. & Portnoy, M. The effect of shift of sensory modality on serial reaction time: A comparison of schizophrenics and normals. American Journal of Psychology, 1961, 74: 224-232.

Torrey, E.G., Hersh, S.P. & McCabe, K.D. Early childhood psychosis and bleeding during pregnancy: A prospective study of gravid women and their offspring. Journal of Autism and Childhood Schizophrenia, 1975, 5: 287-297.

Wechsler, D. A standardized memory scale for clinical use. In D. Wechsler & C.P. Stone, Wechsler Memory Scale: Manual. New York: The Psychological Corporation. (Originally in The Journal of Psychology, 1945, 19: 87-95.)

Wechsler, D. Wechsler Adult Intelligence Scale. New York: Psychological Corporation, 1955.

Wineberg, J., Diller, L., Gordon, W.A., Gerstman, L.J., Lieberman, A., Lakin, P., Hodges, G. & Ezrachi, O. Visual scanning training effect on reading related tasks in acquired right brain damage. Archives of Physical Medicine and Rehabilitation, 1977, 58(11): 479-486.

Weiss, W. The analysis of data in the Collaborative Project of the National Institute of Neurological Diseases and Blindness. In S.S. Chipman, A.M. Lilienfeld, B.G. Greenberg & J.F. Donnelly, Research Methodology and Needs in Perinatal Studies, Springfield, Ill.: Charles C. Thomas, 1966.

Table 1*

The Collaborating Institutions & Number of Registrants

<u>Collaborating Institutions</u>	<u>Years of Registration</u>	<u>Collaborating Institutions</u>	<u>Years of Registration</u>
Boston Hospital for Women (Lying-In Divison) and Children's Hospital Medical Center Boston, Massachusetts	13,137 (1/2/59 - 12/31/65)	University of Minnesota Hospitals Health Sciences Center Minneapolis, Minnesota	3,275 (1/20/59 - 12/31/65)
Children's Hospital State University of New York at Buffalo Buffalo, New York	2,964 (10/7/60 - 12/31/65)	Metropolitan Hospital New York Medical College New York, New York	4,709 (2/2/59 - 12/31/65)
Charity Hospital Tulane University School of Medicine & Medical Center, Louisiana State University	2,590 (3/15/60 - 12/31/65)	University of Oregon Medical School Portland, Oregon	3,055 (3/2/59 - 12/31/65)
Columbia-Presbyterian Medical Center Columbia University College of Physicians and Surgeons New York, New York	2,235 (1/16/59 - 4/30/63)	Pennsylvania Hospital and the Children's Hospital of Philadelphia University of Pennsylvania Philadelphia, Pennsylvania	10,315 (1/8/59 - 12/31/65)
The Johns Hopkins Hospital The Johns Hopkins University School of Medicine Baltimore, Maryland	3,774 (1/6/59 - 12/31/64)	Child Study Center Brown University Providence, Rhode Island	2,851 (3/2/60 - 12/31/65)
Medical College of Virginia Virginia Commonwealth University Richmond, Virginia	3,250 (1/2/59 - 12/31/65)	Gailor Hospital University of Tennessee College of Medicine Memphis, Tennessee	3,553 (10/16/59 - 12/21/65)

*Source: Niswander et al., The Women and Their Pregnancies.
DHEW Publication No. (NIH) 73-379, 1972.

Total Live Births: 45,717

Total Twin Births: 569

The Philadelphia Collaborative Perinatal Study

Jeanne Marecek
Swarthmore College

The W.T. Grant Foundation provided initial support for the establishment of the Institute for the Continuous Study of Man. Space for the medical records library is provided by the Eastern Pennsylvania Psychiatric Institute, Philadelphia, PA.

The study of behavioral deviations of childhood was carried out by L. Isobel Rigg, M.D. and T. McNair Scott, M.D. The interview protocol was designed by the late John Rose, M.D. of the Philadelphia Child Guidance Clinic. Donald E. Goldstein, now at the University of Washington, Seattle, designed the computerized master file of the Philadelphia CPP data. The studies of physical growth and development and of the effects of exposure to lead are being carried out at the W.M. Krogman Center for Research in Child Growth and Development, Children's Hospital of Philadelphia. Research on early family formation is being carried out in collaboration with the Philadelphia Health Management Corporation. The research program on delinquent behavior is being carried out at the Center for the Study of Criminology and Criminal Law, University of Pennsylvania, Philadelphia, PA.

The Philadelphia Collaborative Perinatal Study

Jeanne Marecek
Swarthmore College

This chapter describes the Philadelphia Collaborative Perinatal Project (CPP) carried out from 1959 to 1973 at Pennsylvania Hospital and Children's Hospital of Philadelphia.

In this chapter, I shall briefly describe the purpose of the CPP and the sampling design used to enroll participants in the project. Next, a detailed description of the sample will be provided and ways in which this sample diverges from samples collected at other CPP sites will be highlighted. In addition, rates of attrition will be discussed. The third section will describe briefly the design of the data collection procedures, the measures taken, and the timing of the exams. Because the Philadelphia CPP followed the standardized national protocol for data collection, only a summary will be included here. Again, divergencies between the Philadelphia CPP and other sites will be noted; in particular, we shall discuss a sub-project in which additional developmental data were collected in parallel to the main study. Finally, I shall conclude with a view toward the future -- prospects for further investigations and some problems.

Overview of the Philadelphia Collaborative Perinatal Project

The National Collaborative Perinatal Project was begun in 1959 to investigate factors leading to abnormal outcomes of pregnancy. Factors under investigation included parents' genetic backgrounds and biological characteristics; environmental events such as socioeconomic stress; and complications of pregnancy, labor, and delivery. Children's physical health at birth and throughout their infancy and early childhood was studied, with special emphasis on neurological and sensory impairment.

The CPP enrolled its first participants in 1959 and continued to register participants until 1965. The last births occurred in 1966. Families remained in the project for the first seven years of the index child's life. Thus, data collection persisted until 1973.

The Philadelphia Collaborative Perinatal Project

Two Philadelphia hospitals jointly participated in the CPP. The Pennsylvania Hospital was responsible for enrolling pregnant women in the study and for collecting obstetrical information from the time of the woman's first prenatal visit to the time when she and her newborn were discharged from the hospital. Children's Hospital of Philadelphia assumed responsibility for collecting pediatric information from infancy to the child's seventh year of life.

Characteristics of the Philadelphia CPP Sample

The characteristics of the sample were largely determined by the location of Pennsylvania Hospital and the sampling frame. The hospital is adjacent to several lower-income neighborhoods in inner-city Philadelphia and many of its clinic patients were drawn from these neighborhoods. The sampling frame consisted of all pregnant women entering the obstetric clinic except the following: 1) women who were planning to deliver elsewhere; 2) women who were planning to give up their infants for adoption; 3) women who were transients planning to leave the area before delivering; and, 4) women in labor at their first visit to the clinic. The ratio of women actually enrolled in the CPP to those in the sampling frame was nearly 100% (Niswander & Gordon, 1972).

Overall, the Philadelphia CPP registered 9,792 pregnancies or roughly 19 percent of the national CPP sample. These pregnancies resulted in 9,636 liveborn children. In accord with the protocol of the national study, only families of liveborn children were followed. Thus, for most purposes, this group constitutes the sample of interest. Therefore, the descriptions of sample characteristics and participation rates that follow pertain only to liveborn children and their mothers.

Table 1 shows the numbers of liveborn children in each birth cohort according to their ethnic group. The CPP protocol allowed women to include as many pregnancies in the study as they wished. In the national sample, the rate of repeat pregnancies was 19%; however, in Philadelphia, it was considerably higher. Twenty-seven percent of the study children

were siblings of other study children. Table 2 shows the number of women who participated more than once in the study.

It should be noted that women with repeat pregnancies did not necessarily register all subsequent pregnancies in the CPP. Thus, information from the birth registry alone is not sufficient for studies of child-spacing or family size. Nonetheless, these repeat births are of considerable interest because they allow researchers to compare pregnancies or childhood experiences of siblings while holding constant some characteristics of the mother and her environment.

Characteristics of Mothers Enrolled in the Philadelphia CPP. The focus of the CPP was on early childhood outcomes among the children born in the study. Nonetheless, the data base provides considerable information on the mothers who participated in the study. This information can be used in studies relating maternal characteristics to developmental outcomes among the childbearing period. The number of registrants who bore at least one liveborn child is 7112. Table 3 shows the distribution of registrants' ages at the birth of the child. The sample of black women shows a somewhat wider age range than the samples of women from other ethnic groups, probably reflecting the tendency for blacks to have somewhat larger families. The figures in Table 3 present a somewhat misleading view of the women in the Philadelphia CPP because they are not corrected for women who participated with repeated pregnancies in the study. (Such women are counted every time they give birth.)

By enrolling all eligible pregnant women into the study, the Philadelphia CPP obtained a broad sample of women with respect to age and parity. However, in comparison with the national CPP, young women and women giving birth for the first time were overrepresented (Niswander & Gordon, 1972). Statistics for the city of Philadelphia during the period when the CPP was ongoing show that women under 18 years of age were overrepresented in the CPP relative to the city at large. In the city, 3.3% of registered births were to women under 18 years of age; in the CPP sample, 7% of births to whites and 17.2% of births to blacks were to women under eighteen (Marecek, 1979).

It should be noted that the socioeconomic index used in this study bases its calculations on the characteristics of the designated head of household. These designations vary to some degree with the race and age of the participant. Black women were somewhat more likely than other women to be living in single-adult households of which they were the heads. Younger women (especially teenagers) were more likely than older women to designate a parent as the head of the household. This was especially true for young black women, who commonly resided with their families of origin at the time of their pregnancies. Thus, caution must be exercised in

drawing comparisons across ethnic groups. Comparison of the household incomes of CPP participants with those of Pennsylvania residents in general offer another means of assessing the economic status of the sample. At registration into the Philadelphia CPP 16.8% of white households and 20.7% of black households reported incomes under \$2,000. Comparable statistics for the state of Pennsylvania were 3.6% for whites and 19.6% for blacks. Only 0.8% of white households and 1.0% of black households in the CPP had incomes of \$8,000 or higher, as compared to 33.6% and 10.4% respectively in the state of Pennsylvania. These data suggest that black CPP registrants were somewhat economically-deprived relative to black Pennsylvanians in general and that white CPP registrants were considerably less well off than white Pennsylvanians (Marecek, 1979).

The educational attainment of the participating women themselves bears further evidence of their low socioeconomic status. Among the white registrants, 23.1 percent completed less than nine years of education and only 3.2 percent completed more than 12 years of education. The comparable rates for white women in Pennsylvania were 17.4% and 10.0%. Black CPP registrants were in the midrange of black Pennsylvania women in terms of educational attainment. Only 12.7% of black CPP registrants completed less than nine years of education, but also only 2.6% completed more than 12 years of education; comparable rates for black women in the state were 26.1% and 6.3% respectively (Marecek, 1979).

At registration, roughly 65 percent of the women in the Philadelphia CPP were married, 25% were never married and 10% were separated, widowed, or divorced (Niswander & Gordon, 1972). Niswander and Gordon's statistics count all registrations and, thus, count some women more than once. If we consider the marital status at registration of women bearing their first child, the proportion of married women among the black women is 38 percent; among the white women, it is 94 percent.

In summary, the sample of women who enrolled in the Philadelphia CPP can be characterized as members of the more deprived segments of society. This is especially true of the white women participating in the study who appear distinctly less advantaged than women in the population at large. Black women also are disadvantaged; however, their status is more normative relative to that of blacks in the population at large.

Characteristics of the Philadelphia CPP children. Table 1 provides the race and year of birth of children born in the CPP. As of this writing, the oldest children are reaching their twenty-first birthdays and the youngest are turning fourteen. There were 117 multiple births in the sample, including one set of triplets. Detailed information on the neonatal status of the children in the sample can be found in Niswander and Gordon (1972).

Participation rates. In any longitudinal study, maintaining participants' willingness to cooperate is essential. In the CPP, incentives to participate included free obstetric and pediatric care, free transportation to the hospital for mothers and their children, and small gifts for the mothers. In addition to providing these material incentives, the project staff made strenuous efforts to remain in contact with participating families. These efforts were more successful for black families than for white or Puerto Rican families, as Table 4 reveals.

Data Collected in the Philadelphia CPP

In general, the procedures for collecting and scoring data, the type of information collected, and the timing of various exams and interviews was carried out according to the national protocol. Data collection commenced when a registrant made her first prenatal visit to the obstetric clinic at Pennsylvania Hospital. The timing of this visit was, of course, chosen by the registrant and thus, it varies somewhat across the sample. Nonetheless, over 80% of the sample had enrolled in the study by the 28th week of pregnancy. The battery of interviews and examinations given at registration provides demographic and socioeconomic information on the woman and the baby's father and personal and family medical histories. Data on the progress of the pregnancy were collected and extensive observations of the labor and delivery were recorded. Additional data on the mothers consisted of two Maternal Behavior Profiles completed by an examiner at the child's 4-month and 8-month exams. Finally, a update of the family's socioeconomic status and a review of the family's physical health was conducted at the 7-year exams. The national protocol called for the administration of a nonverbal intelligence test to CPP mothers at the 4-year exam. However, less than 30 percent of the Philadelphia participants have data available on this test.

Major examinations administered to the child are listed by chronological order in Table 5. The amount of information available on the fathers of children born in the CPP is slight. It consists mostly of some demographic information (name, race, age, religion, birthplace, occupation) and a family medical history directed toward identifying the presence of hereditary abnormalities. Follow-up information is not available on the fathers unless they were married to the child's mother at the time of the 7-year follow-up.

Detailed information on the data collection protocols and on the coding procedures can be obtained by consulting the manuals compiled by the national project staff (The Collaborative Study on Cerebral Palsy, Mental Retardation, and other Neurological and Sensory Disorders of Infancy and Childhood, Parts I-IV, 1966 and 1974; Definition of Codes for the

Collaborative Perinatal Project, 1969).

Subproject on Behavioral Deviations of Childhood. Concurrent with the CPP, project staff in Philadelphia collected additional information on the presence of behavioral abnormalities in a subsample of children. All children in the 1960, 1961 and 1962 birth cohorts were included in this study. Data were collected by means of semi-structured interviews conducted at 6-month intervals from birth to the child's seventh birthday. The respondent was the child's mother or primary caregiver. Table 6 lists the categories of behavior that were the focus of the interviews.

An exploratory study was carried out on the interviews conducted at the child's fourth and seventh birthdays (Rigg & Scott, no date). Interview protocols were evaluated as to the presence of each behavioral deviation by trained psychiatric judges. In order for a deviation to be scored as present, it had to be described in the caregiver's account as a regularly-occurring or established behavior. As Rigg and Scott note, several categories of deviation showed strong differences across gender and across ethnic group.

For several categories of behavior deviations, the frequency of occurrence is high enough to permit further study. Furthermore, full sets of raw interview protocols (up to 14 interviews) are available in the project files. In addition, the data from Rigg and Scott's pilot study are entered in the computerized data files.

The Institute for the Continuous Study of Man

When the period of active data collection for the national CPP ended, each local site assumed the task of preserving and using its own data. In Philadelphia, a nonprofit corporation, the Institute for the Continuous Study of Man, was established in 1974 to carry out the task.

The major goal of the Institute is the utilization of the Philadelphia CPP data base and the continuation of research studies on the population. Research currently conducted under the Institute's auspices comprises four program areas.

A Look Into the Future

Members of the CPP population are presently making the transition from adolescence to young adulthood. This is a crucial developmental

period, especially for young people from urban, lower-income backgrounds, who are at elevated risk for school drop-out, unemployment, unintended pregnancy, and delinquent behavior. Studies of CPP participants and their families could shed light on the processes involved in making successful transitions to adulthood and on ways of preventing deleterious social behavior from occurring in ways of minimizing its long-term effects.

A second area of inquiry to which the CPP lends itself is the investigation of the antecedents of health conditions of adulthood and of health threatening behaviors such as smoking, overeating, and substance abuse. The data collected by the CPP provide some information on childhood health conditions and on maternal health-related behaviors. This information could be used in prospective studies on the acquisition or establishment of beneficial or detrimental health-related behaviors.

Finally, it should be noted that a number of young women in the Philadelphia CPP have begun childbearing. Thus, three-generational studies of the CPP mothers, their children, and their grandchildren will soon be possible. Furthermore, as the CPP "children" become parents, it will be possible to study the antecedents of parenting and other family processes using the early psychosocial data.

Three major issues need to be addressed if the Philadelphia CPP data base is to be used to its fullest extent. The first is the need to maintain contact with as full a complement of study participants as possible. The near future will be a crucial time for maintaining contact with participants. As they finish their schooling, it will no longer be possible to locate them through the School District's records. Furthermore, many will be establishing their own households and a number of young women may relinquish their family names when they marry. Thus, we need to devise strategies for augmenting the directory system with information that will enable us to find participants even if their names, addresses, and telephone numbers change.

A second issue concerns the regulation of use of research data by Federal and local agencies. The data collected by the CPP fall under the scope of the Privacy Act of 1974 (P.L. 93-579), which defines conditions of disclosure, participants' rights of access, and requirements for maintaining confidentiality. Generally speaking, the Institute's researchers have found the guidelines provided by the Privacy Act to be both adequate and reasonable. Recently, however, more stringent requirements for consent procedures and more limitations on release of information have been proposed in some circles. Were such requirements to become law, future use of the CPP data and, indeed, of most other prospective, longitudinal data bases would be jeopardized (cf. Gordis & Gold, 1980). Should the adoption of more stringent guidelines become imminent, scientists committed to the longitudinal approach may need to inform the public and various legislative bodies of the special requirements of such research and of its

unique payoff in scientific knowledge.

The third issue confronting users of the Philadelphia CPP data base concerns funding for operating expenses associated with maintaining the data archives, the sample, and general administration. A portion of these expenses can be defrayed through specific research projects, but experience has shown that this approach will not solve the problems of maintenance entirely. For example, piecemeal efforts at sample maintenance will lead to retention of only a highly-skewed and unrepresentative sample. This would greatly diminish the usefulness of the Philadelphia CPP as a prospective, longitudinal data base. Although efforts are underway to generate revenue for general support, there is also a need for a broader effort to inform the public and the scientific community about the unique ways in which longitudinal studies can contribute to our knowledge about human development and about the research possibilities of extent sources of longitudinal data. We hope that this chapter contributes to that end.

Table 1

Number of Children in Each Birth Cohort in the Philadelphia CPP, by Ethnic Group

Year of Birth	Race				Total
	Black	White	Hispanic	Other	
1959	746	103	11	1	861
1960	1108	145	15	0	1268
1961	1069	131	45	5	1250
1962	1205	134	42	3	1384
1963	1335	150	54	1	1539
1964	1470	113	61	1	1645
1965	1209	88	64	1	1362
1966	298	12	15	1	326
Unknown	1	0	0	0	1
TOTAL	8440	876	307	13	9636

Table 2

Number of Philadelphia CPP Registrants Who Had Repeat Births.^a

Number of Births	Black	White	Hispanic	Other	TOTAL
1	4403	597	160	9	5169
2	1212	103	49	2	1366
3	377	23	11	0	411
4	90	1	4	0	95
5	22	0	0	0	22
6	2	0	0	0	2
TOTALS	6106	724	224	11	7065

^aThese numbers include only births resulting in liveborn children. Miscarriages and stillbirths are excluded.

Table 3

Age of Registrants at the Birth of the Child, by Ethnic Group

Mother's Age	Black		White		Hispanic		Other		Total N
	N	%	N	%	N	%	N	%	
<15	101	1.2	1	0.1	0	0	0	0	102
15-19	2373	28.1	154	17.6	78	25.4	2	15.4	2607
20-25	3259	38.6	389	44.4	140	45.6	7	53.8	3795
26-30	1459	17.3	146	16.7	53	17.3	2	15.4	1660
31-35	846	10.0	108	12.3	20	6.5	0	0	974
36-40	345	4.1	64	7.3	15	4.9	2	15.4	426
41+	57	0.7	14	1.6	1	0.3	0	0	72

Table 4

Participation Rates in the Philadelphia CPP by Ethnic Group

Time of Exam	BLACK		WHITE		OTHER	
	N	%	N	%	N	%
Registration Interview	8440	100	876	100	320	100
8-month Neurological Exam	6290	74.5	517	59.0	207	64.7
4-year Psychological Exam	5799	68.7	475	54.2	178	52.2
7-year Psychological Exam	6756	80.0	516	58.9	178	55.6
7-year Socioeconomic Review	5526	78.5	529	60.4	169	52.8

Table 5

Major Exams Administered to Children in the Philadelphia CPP,
by Time of Administration

<u>Within 48 Hours of Birth</u>
Neonatal Physical Exam
Neonatal Neurological Exam
<u>Four Months</u>
Pediatric Exam
Record of Physical Growth Measurements
<u>Eight Months</u>
Bayley Scales of Mental and Motor Development ^a
Infant Behavior Profile
<u>One Year</u>
Neurological Exam
<u>Three Years</u>
Speech, Language, and Hearing Exam ^b
<u>Four Years</u>
Stanford-Binet Intelligence Scale, Form L-M
Graham-Ernhart Block Sort Test
Motor Test (Gross Motor, Fine Motor, and Dominance subtests)
Child Behavior Profile
<u>Seven Years</u>
Pediatric-Neurological Exam
Visual Screening
Auditory-Vocal Association Test
Bender-Gestalt Test (with Koppitz scoring)
Wechsler Intelligence Scale for Children
Goodenough-Harris Draw-a-Person Test
Tactile Finger Recognition Test
Wide Range Achievement Test ^a (Spelling, Arithmetic, Reading)
<u>Eight Years</u>
Final Speech, Language, and Hearing Exam
School Follow-up Record

^aThese tests were adapted for use in the CPP.

^bIn Philadelphia, the full Speech, Language, and Hearing Exam was given only to Children (N=543) who showed signs of abnormality in a preliminary screening.

Table 6

Definitions of Behavior Deviations

1. Abnormality of Behavior Control

Both overcontrol and undercontrol of behavior are assessed. Overcontrol of behavior includes inhibition of approval-seeking behavior, fearfulness, and shyness. Constipation in the absence of acute physical illness is also included. Undercontrol of behavior includes disorganization of behavior under stress. Habitual temper tantrums, hitting, biting, and breath-holding are examples of undercontrol of behavior. At age 7, low attention span and inability to persist in tasks requiring adherence to rules and behavior conformity are also instances of undercontrol of behavior.

2. Bedwetting

In both the 4-year old child and the 7-year old child, bedwetting is scored as a behavior deviation if it occurs once a week or more.

3. Disruption in Conduct

Disturbances in social conduct at home, in school, or in the community are scored as conduct disruptions. These include stealing, truancy, and destructiveness.

4. Delay in the Development of Self-Care

At age 4, children who have failed to accomplish the following are scored as having delayed development: are unable to dress themselves (except for tying shoes); are able to feed themselves; are toilet-trained by day; are unable to sleep alone.

At age 7, children are scored as having delayed development if they have not accomplished all of the 4-year skills and the following: are able to tie shoes and are able to play outdoors with adult supervision.

5. Nail-Biting

Nail-biting is scored as a deviation if it is habitual.

6. Pica

Habitual ingestion of non-food substances (paper, paint, dirt, etc.) is scored as a deviation.

7. Phobic Responses

Strong fear reactions to common environmental stimuli (e.g. bugs, dogs, noise of vacuum cleaner) are scored as deviant behaviors. Fearful reactions to new situations are not scored as phobic responses.

Table 6 (Continued)

8. Sleep Disturbance
Chronic wakefulness, nightmares, and sleep-walking are scored as behavior deviations.

9. Thumb-Sucking

Thumb-sucking is scored as a behavior deviation if it is habitual.

10. Social Maladjustment

Both active and passive maladjusted behaviors are scored. The former include difficulties in interpersonal relations with peers, fighting, and taking toys away from playmates. The latter include withdrawal from other children and failure to protect self or toys from other children.

11. Mechanical Speech Defects

Difficulty with specific sounds or lisping are recorded as deviant behaviors. A history of attending speech classes is taken to indicate the presence of speech defects.

12. Defect in Speech Construction

Inadequate sentence structure and omission of pronouns, verbs, etc. are scored as deviant behaviors.

13. Stuttering

An established pattern of stuttering is scored as a deviant behavior.

14. Speech Disturbance Summary

If any of the three speech problems listed above are observed, the child is scored as having a speech disturbance.

References

- Broman, S.H., Nichols, P.L. & Kennedy, W.A. Preschool IQ: Prenatal and early development correlates. New York: Wiley, 1975
- Definition of codes for the Collaborative Perinatal Project: Perinatal Research Branch National Institute of Neurological Diseases and Stroke, Bethesda, MD, 1969.
- Gordis, L. & Gold, E. Privacy, confidentiality and the use of medical records in research. Science, 1980, 207, 153-156.
- Hediger, M.L. Personal communication, March, 1980.
- Katz, S.H., Hediger, M.L., Schall, J.I., Bowers, E.J., Barker, W.F., Aurand, S., Evelith, P.B., Guskin, A.B. & Parks, J.S. Blood pressure, growth and maturation from childhood through adolescence: Results of a mixed longitudinal study from the Philadelphia blood pressure project. Hypertension, 1980, in press.
- Marecek, J. Economic, social and psychological consequences of adolescent childbearing: An analysis of data from the Philadelphia Collaborative Prenatal Project. Final report to the National Institute of Child Health and Human Development, June, 1979.
- McFalls, J.A. Social science and the collaborative perinatal project: An opportunity for research. Public Data Use, 1976, 4(5), 37-47.
- Myriantopoulos, N.C. & French, K.S. An application of the U.S. Bureau of the Census Socioeconomic Index to a large, diversified patient population. Social Science and Medicine, 1968, 2, 283-299.
- Niswander, K.R. & Gordon, M. The women and their pregnancies. Philadelphia.
- Rigg, L.I. & Scott, T.F.M., Childhood behavioral deviations in an inner city population: I. Prevalence at ages of 4 and 7 years. Unpublished manuscript. Philadelphia: Children's Hospital of Philadelphia, undated.
- The Collaborative Study on cerebral palsy, mental retardation and other neurological and sensory disorders of infancy and childhood, Parts I-IV. Perinatal Research Branch, National Institute of Neurological Diseases and Blindness, Bethesda, MD, 1966 & 1974.

The Psychosocial Development of Aggressive Behavior*

Leonard D. Eron

University of Illinois at Chicago Circle

L. Rowell Huesmann

University of Illinois at Chicago Circle

Leopold O. Walder

Behavioral Service Consultants

Monroe M. Lefkowitz

New York State Department of Mental Hygiene

*Various aspects of this research have been supported by the following grants and contracts from NIMH, Grant No. M 1726 (L.D. Eron); and 28280 and 31866 (L.R. Huesmann) and Contract No. HSM 42-70-60 (M.M. Lefkowitz); and from the Office of Child Development, Grant No. CB 364 (L.D. Eron). Requests for reprints should be sent to L.D. Eron, Univ. of Ill. Chicago Circle, Box 4348, Chicago, Ill.

The Psychosocial Development of Aggressive Behavior

Leonard D. Eron

L. Rowell Huesmann

Leopold O. Walder

Monroe M. Lefkowitz

Introduction

The longitudinal study to be described here was originally sponsored by the Rip Van Winkle Foundation (now called the Esselstyn Foundation), an organization interested in public health in rural areas. The initial charge to the research group was to do an epidemiological study of mental health in rural areas. A review of past efforts in this field, however, revealed two major shortcomings in epidemiological research in mental health: (a) lack of a valid and reliable criterion measure, and (b) the use of biased, unrepresentative samples (Eron, 1956). It was decided to circumvent the whole thorny criterion problem (i.e., What is mental health and when do you know you are looking at it so that you can count it accurately?) by isolating and concentrating on just one aspect or variable of mental health, which was more manageable for research than the global concept. We selected a variable which we felt was definable, observable and measurable and that variable was overt aggression. Aggression was defined for our purposes as an "act whose goal response is injury to another object," a slight variation on the definition used by Dollard and others (1939) and by Sears and others (1953), and which deliberately excludes both intropunitive aggression and nonhostile expression of aggression.

Method

Once we decided what we wanted to study, we then had to decide how to study it. There are a number of research strategies that can be used in studying aggression -- one is to study a small number of subjects intensively, either by clinical case method or by random assignment to groups which are then given different treatments. Another strategy is to survey a large number of subjects with less intensive methods. Because we were committed to an epidemiological point of view, we decided on the latter approach primarily. We were anxious to know about the incidence of aggressive behavior, how it varied with social class, ethnic group, religion, rurality, etc., and, very much aware of sampling biases in previous studies, we wanted to be sure we included among our subjects representatives of each of these variables in all of its major categories. Therefore, we wanted to approximate, as nearly as possible, a 100% sample of at least one segment of the population, cutting across all pertinent variables.

We selected third grade school children as our population and set out to study patterns of aggressive behavior in all third graders in Columbia County, New York. Every school in the county participated in the study, including the one parochial school. To accomplish this, it was necessary to get permission from over 20 individual school boards and to contact teachers in 40 different classrooms. The chief reason we were able to get such a high degree of cooperation was the high regard the community had for the organization sponsoring the research. The Rip Van Winkle Foundation was an integral part of the community and its employees participated in community activities at all levels.

We decided on school children as subjects primarily because they were available for study, a captive population; also, we wanted to get our subjects as early in life as possible since we were interested in longitudinal studies. The third grade was the earliest that we felt we could begin to use children as subjects in this type of research -- by this time youngsters can read and write and their cooperation in and understanding of the kinds of survey procedures we wanted to use could be elicited. Also, by the time they are in the third grade, i.e., 8 years old, children usually have formed stable patterns of relationships with their peers; the pecking order by then has been fairly well established. In general, we felt that meaningful results could be obtained from third graders on the basis of which it might be possible to predict future behavior of these children, who could then be followed for a minimum of nine years, at least until high school graduation, in order to test the validity of the predictions.

Columbia County was an ideal location for such a study. According to census data, the county was representative of approximately 38% of the rural counties in the nation at the time the study was implemented (U.S. Department of Commerce, 1953). A wide range of social class, ethnicity and rurality existed in a relatively stable population; people did not move around much in Columbia County. There had been no great industrial expansion or mass migration in or out of the county around that time. The population in 1900 was 43,000 and in 1960 it was 43,000. This stability permitted long-term follow up with less likelihood of the marked attrition of the original sample which plagues most longitudinal studies, especially those centered in large university communities with their very mobile populations. A county-wide survey, cutting across all social and cultural variables, helped insure that each variable would be represented in most of its categories in any analysis of the data. Adequate representation in most of the pertinent categories, while no problem in manipulative experiments, usually is difficult with small samples in correlational studies where categories are filled by selection rather than manipulation. It is for this reason that so large and unrestricted a sample was necessary. However, the sample was not really so large that it was unmanageable. The size of the county population and the integration of the research team in the community made it possible to follow the sample individuals in many contexts other than school, e.g., courts, newspapers, community agencies, etc.

The initial wave of this study was conducted in 1960. Five years prior to this were devoted to devising and validating large scale measures of aggression and its presumed antecedent conditions. During that time over 2,000 subjects were seen in various pretests and pilot studies. Our major thesis was that aggression is a learned behavior. Therefore, we were interested in relating aggression of children at school to the learning conditions which existed in the home. Our principal measure of aggression was a peer nomination measure in which every child in a classroom rated every other child on a series of aggressive behaviors. We had started out with approximately 1,000 different items of aggression -- ways in which 8 year old children can display aggressive behavior. By the time our measurement operations were completed just prior to the major survey, the definition of aggression had changed to "an act which injures or irritates another person" and the measure consisted of 10 specific items of aggressive behavior which conformed to this definition. These 10 items were embedded among 12 other peer nomination items tapping popularity, aggression anxiety, activity level and success in aggression (Walder, Abelson, Eron, Banta & Laulicht, 1961).

Information about the learning conditions for these aggressive behaviors was taken from a precoded, objective interview used with the parents in an individual, face-to-face situation. The final interview

contained 286 items, comprising 41 variables consisting of 1 to 27 items each, with a median number of 5 items per variable. The variables were classified largely into 4 types: reinforcers of aggression, instigators to aggression, identification and sociocultural variables. In addition, there were a number of variables to control for validity of parents' responses, including the respondent's candor, consistency, reliability and knowledge of the child. In assembling this interview a number of test-construction procedures not usually associated with interview-building were utilized. These included: (1) judgment of content validity of items by a panel of experts; (2) selection of items on the basis of their difficulty and variability in responses as well as relation to the criterion individually; and (3) establishment of homogeneous scales by inter-item and item-total correlations. Extensive pretests were done. Responses of mothers and fathers were compared with each other and in relation to the criterion, singly and in combination. Factor analyses of mothers' and father's interviews done separately strengthened our belief that it was necessary to interview both parents individually since there were some areas of parent-child interaction in which mothers were better informants, some where fathers were better informants and still others in which responses of both parents had to be combined. Because of the variety of controls instituted in these data-gathering procedures, we were reasonably satisfied that the information obtained and the relations demonstrated were valid representatives of the real-life behaviors of these subjects.

Results and Discussion

The initial wave included 875 children, the entire third grade population of Columbia County, New York in 1960, 85% of their mothers and 71% of their fathers. The basic demographic information on this sample seems to be representative of what one would expect to find for third graders across much of the United States. The modal age was 8; the predominant socioeconomic environment was middle class and the mean IQ at that time for the children in the sample was 104.4 ± 14 .

The analysis of the 1960 data yielded a large amount of information about the characteristics of an aggressive child, the family and the environment (Eron, Walder & Lefkowitz, 1971). It is fair to say that all the major findings of that analysis were consistent with the hypothesis that aggression may be learned by a child from interactions with the environment. For example, it was found that punishment of the child at home was correlated with the aggressiveness of the child at school. However, there were contrasting effects of punishment for two types of

children -- low-aggressive children who had close identification with their parents and high-aggressive children who identified only moderately with their parents. Punishment seemed to inhibit aggression in the former group and facilitate it in the latter group.

The difficulty with such a one-time field study is that one cannot tease apart causation from correlation. In the example above, was punishment correlated with aggression because an aggressive child is punished more or because many children imitate the punishments they receive, or both? In order to separate causation and correlation, one needs to obtain repeated measurements on a child during development. With such longitudinal data one can perhaps go beyond correlational theories and distinguish between the plausibility of rival causal theories. With this in mind, we undertook a reinvestigation of our 875 subjects in 1970, ten years after the original data collection. By this time most of the subjects had been out of high school for a year and thus we identified them as being in the 13th grade.

The residential mobility of the American population makes any longitudinal study extremely difficult. From school records, we discovered that approximately 52% of the original sample had graduated from high school in June 1969 in Columbia County. To locate as many of the subjects in the original sample as possible, we queried school officials, examined high school yearbooks, telephone directories, voter lists, tax lists and the county directory. In addition, each subject who did appear was asked about the whereabouts of any of the missing subjects. In this way we finally obtained addresses for 735 of the 875 original subjects (84%).

Obtaining addresses was only part of the problem, however. Getting subjects to come to an interview was an equally imposing task. Letters were sent to each of the 735 subjects inviting them to come to our offices at a central location in Columbia County for an interview. As lures, we appealed to their sense of responsibility and offered each of them \$20 for appearing. Of the 735 subjects to whom letters were sent, 460 indicated willingness to be interviewed. Of the 460 subjects who agreed to be interviewed, 436 actually appeared. Of the 436 who appeared, 427 completed the procedure and contributed usable data.

It is judicious to ask how representative of the population of the United States these 427 subjects are and how representative of the original 875 subjects they are. The 427 subjects had a modal age of 19 years. While 211 were boys, 216 were girls. On the average they had completed 12.6 years of school (in addition to kindergarten). Based on the 25% for which current IQ scores could be obtained (we have no reason to believe this is a biased sample), the mean 12th-grade IQ was 109.1 ± 11.6 . Using father's occupation as a measure of social status (Warner,

Meeker & Eells, 1960), the sample could still be described as predominantly middle class. It seems fair to say that our longitudinal sample of 427 children was not too different from what one would find in similar localities across the country.

On the other hand, the 427 subjects who were reinterviewed were not a completely representative sample of the original 875. In particular, the 13th-grade sample included more of the original low-aggression subjects and less of the original high-aggression subjects than one would have expected by chance. Of the boys in the lower quartile of aggression in the third grade, 57% were reinterviewed. However, of the boys who had been in the upper quartile in the third grade, only 27% were interviewed. The corresponding figures for girls were 63% and 33%, respectively. Why would almost double the number of low-aggressive subjects as high-aggressive subjects appear for the reinterview ten years later? The most compelling single explanation we can offer is based on a relation we discovered between a family's residential mobility and its children's aggressiveness. We found these factors to be significantly positively correlated within our reinterview sample ($r = .17$ for boys); so it seems possible that the families of high-aggressive children were most likely to have moved between the times of the two interviews than the families of low-aggressive subjects.

The follow up interview was approximately two hours in length. Like our earlier interview with the parents it was a completely objective, precoded interview conducted in a face-to-face situation. The questionnaire interview yielded information about the subjects' current status, behavior and attitudes. In addition we had the subjects again rate their peers on aggressive behaviors. A peer rating measure of aggression with items similar to the ones the subjects had originally completed was used. Two of the original ten items in the peer rating procedure were changed because they were deemed inappropriate for 19 year old persons. The question, "Who gives dirty looks and sticks out their tongue at other children?" was changed to "Who makes unfriendly gestures?" The item, "Who says: 'Give me that!'" (spoken with emphasis) was eliminated from the 13th grade questionnaire. The pool from which nominations could be made by each subject was also different. Originally there had been 38 different third grade classes. These 38 classes fed into five different high schools. In the follow up interview each subject was presented with the third grade lists of all the feeder classes for the high school he or she attended, in addition to his or her own third grade class and was asked to indicate all those students he or she knew. Thus each subject had a much larger group from which he or she could make nominations and similarly each subject was judged by a much larger group of nominators in the follow up than in the original phase.

One of the most obvious findings in the longitudinal study was the stability of aggressive behavior over a 10-year period for both boys and girls. The correlation for boys was .38 and for girls .47 ($p < .001$). These high correlations, not too different from correlations in IQ over a similar period of time, are due, however, to something more than reliability of rating by the same judges. Aside from the fact that there was a 10-year lag, making it highly unlikely that memory of earlier ratings was an important factor, each subject was rated by a somewhat different set of raters in the later period than had rated him in the former period, as noted above. Thus there were two different groups of nominators at the two points in time -- they were overlapping to be sure, but different.

Further, peer-rated aggression at age 19 correlated positively with the various self ratings of aggression made when the subjects were that age. During the reinterview in the 13th grade each subject was tested with the MMPI (Hathaway & McKinley, 1969). It has been reported that elevations on scale 4 (psychopathic deviate) and 9 (hypomania) of this test are indicative of delinquency (Hathaway & Monachesi, 1963). Hence, we added together each subject's T-scores on scales 4 and 9 of the MMPI to get a score reflecting potential antisocial behavior. This measure correlated positively with 13th grade peer-rated aggression (.28 $p < .01$ for girls; .39 $p < .01$ for boys). However, the correlation improved (.41 $p < .001$ for boys; .31 $p < .001$ for girls) when scale F (a validity scale) was added to these two clinical scales as in Huesmann, Lefkowitz and Eron (1978).

In order to secure self-reports of aggression from our subjects in the 13th grade, we included two sets of questions in the interview. One set was designed to have face validity as a measure of a subject's propensity for antisocial behavior and the other set was designed to measure the intensity of a subject's aggressive habits. In the former set there were 26 questions such as -- "In the last three years, how many times have you taken something from a store without paying for it?" "How many times in the last three years have you hit someone badly enough to need bandages or a doctor?"

In the latter set there were rating scales on which the subject checked one of the following: almost always true, often true, sometimes true, seldom true, never true, which best expressed his acceptance of items like: "I feel like swearing," "I feel like being a little rude to people," "I feel like picking a fight or arguing with people." The questions in both these sets had been derived through extensive pretesting. A Total Aggressive Habit score was derived by summing the scores on all the questions from both sets. The self-rating score correlated highly with 13th grade peer-rated aggression (.50 for males and .45 for girls, $p < .001$).

Each of the above measures of aggression possesses reasonable face validity. In addition, their demonstrated statistical interrelatedness supports their validity. With regard to reliability, however, we only had information on the MMPI and peer-rating measures which are both highly reliable (Eron, Walder & Lefkowitz, 1971; Hathaway & McKinley, 1969).

Do these aggression variables measure the type of aggression in which we are interested: "an act which injures or irritates another person?" To try to answer this question, we asked the New York State Division of Criminal Justice, which collects data on arrests within the state, to determine the number of arrests of male subjects who were low and high on aggression in the 3rd grade. The results indicated that three times as many subjects who were in the high quartile as were in the low were mentioned as having been arrested sometime before age 20. Also the final MMPI measure, scales F + 4 + 9, was subsequently shown to be an excellent discriminator between delinquent and general populations of males and females even when intelligence and social status were controlled (Huesmann, Lefkowitz & Eron, 1978). These data support the contention that our aggression measures are valid and the consistency in peer-rating over a 10-year period indicates that aggressive behavior is stable over time and predictive from grade 3 to a year beyond high school (i.e., age 8 to age 19).

It is interesting that we were able to get such high correlations between self and peer ratings in the 13th grade although there had been no such correlation in the 3rd grade. Perhaps at age 19 subjects can describe themselves better; or they feel it's less incriminating to admit these behaviors to an interviewer who is a stranger and whom they will probably never see again than when they were in a classroom in grade 3 and had no real assurance that the teacher wouldn't see their answers. Or perhaps it might even be the difference in the times -- in 1960, it wasn't the "in thing" for young people to engage in or admit to certain antisocial behaviors which now are acceptable at least to persons in this age group. "Ripping off" is the term currently used to legitimize stealing and make it socially acceptable. In 1960, there was no such term.

Having established the credibility of these as measures of aggression we may now turn our attention to what the relation is of the child rearing variables measured when the child was 8 to his or her aggressive behavior at age 19. As stated previously, we were predicting from four classes of antecedent variables: instigators, contingent responses, identification and social class. This information was obtained primarily from parent interviews but there were also classroom procedures with the children which measured variables considered to be antecedents.

An instigator to a behavior is a stimulus that usually elicits the behavior in question as a response. In the 3rd grade we measured four major characteristics of our subjects and their families which we considered potential instigators of aggression: parental rejection of the child, parental disharmony, lack of nurturance of the child by the parents and the child's IQ.

While instigators may trigger aggressive acts, the development of more enduring aggressive habits should be influenced by how subjects' environments respond to their aggressiveness. Among the most important reinforcing agents for an 8 year old or younger child are undoubtedly the parents; so in the 3rd grade we selected the parents' use of punishment as a measure of how a child's aggressive acts are reinforced. The higher the score, the more punishment a parent said was used in controlling the child.

As for identification, the importance of modeling and observational learning in the acquisition of behaviors has become increasingly clear in recent years. Children do what they see being done, especially if they see it being reinforced (Bandura & Walters, 1963; Bandura, Ross & Ross, 1963). While various characteristics of the model and the situation influence the likelihood of immediate imitation (Bandura, 1969), observational learning would seem to have great potential as a determiner of life-long patterns of behavior.

A second type of identification relevant to the development of personality is the psychoanalytic concept of internalization of parents' values, desires and standards (Freud, 1923). The child who successfully internalizes parental standards finds it easy to control his or her own behavior in line with the parents' proscriptions.

Both types of identification variables, modeling and internalization, were measured in the 3rd grade. The modeling category included the child's copying of parental behaviors, sex role, modeling and potential for modeling aggressive behaviors observed on television. To measure the child's identification with each parent, we calculated an Expressive Behavior Profile for both parents and for the child. In this procedure, a variation of the Semantic Differential Technique (Osgood, Suci & Tannenbaum, 1957), subjects rated several of their own expressive behaviors, e.g., walking, talking, on 18 five-point scales with bipolar adjectives as anchors, e.g., fast-slow. Since parents were also asked to rate themselves on the behaviors, a measure of discrepancy between the child and each of the parents could be obtained. The total discrepancy between a parent and child was the square root of the sum of the squared discrepancies on each bipolar adjective scale.

The child's sex role identification was measured by the Games and Activities Preference List (Lefkowitz, 1962). Each 3rd grade child was presented with a list of questions requiring him or her to choose between two activities. For examples, "Would you rather go shooting or go bowling?" or "Would you rather use lipstick and powder or use a razor and shaving cream?" The hypothesis which led to the development and use of this measure of identification was that preference for masculine activities would be positively related to aggression (Lefkowitz, 1962).

The child's exposure to aggressive models on television was measured in the 3rd grade by asking the mother to name the child's three favorite television programs. Ten years later each subject was asked for four current favorite television programs. Scores were assigned on the basis of the number of programs mentioned which had been independently rated by two judges as violent. The designation by these raters of violent and nonviolent programs agreed very well with the assignment of programs by Feshbach and Singer (1971) to aggressive and nonaggressive diets in their field experiment.

Identification with parents through internalization of values was assessed by the amount of confessing to parents that the child performed and the amount of guilt the child expressed to the parents. Both of these behaviors were reported by each parent when the subjects were 8 years old.

Sociocultural variables can be hypothesized to affect the development of personality in a variety of ways ranging from genetic predispositions, to nutrition, to learning. In choosing the variables to represent a subject's sociocultural environment, we entertained no particular hypotheses about the processes underlying their effect; rather we tried to be exhaustive in our coverage of potential sociocultural predictors of aggression. In both the 3rd and 13th grades we rated the status of the father's occupation on a seven-point scale (Warner, Meeker & Eells, 1960). Besides occupational status we evaluated the parents' general mobility orientation. A parent high on mobility orientation would be very willing to learn new skills, leave his friends, move, take more responsibility and give up spare time in order to get ahead. The ethnic background of the family was also assessed. A highly ethnic family would be one in which one parent or all the grandparents were born in another country. Finally, both the religious background of the family and the frequency of church attendance as reported by the parents were recorded.

Of the four classes of antecedent variables that had been found to predict aggression in the 3rd grade -- instigation, contingent response to aggression, identification and sociocultural variables -- only the latter two, identification and sociocultural variables, predicted to aggression consistently in the 13th grade.

TABLE OF CONTENTS

DESCRIPTIONS OF LONGITUDINAL RESEARCH

VOL. II A

1. The California Child Health and Development Studies by B.J. van der Berg	1 - 13
2. The St. Louis Baby Study: Theory, Practice and Findings by T.E. Jordan	14 - 28
3. A Neuropsychiatric Follow-up of Children in the Collaborative Perinatal Project Population by P.A. O'Connor, D. Shaffer, S. Shafer & C. Stokman	29 - 50
4. The Philadelphia Collaborative Perinatal Study by J. Marecek	51 - 66
5. The Psychosocial Development of Aggressive Behavior by L.D. Eron, L.R. Huessmann, L.O. Walder & M.M. Lefkowitz	67 - 81
6. ETS Study of Academic Prediction and Growth by T.L. Hilton	82 - 96
7. The Youth in Transition Project by J.G. Bachman & P.M. O'Malley	97 - 118

9. An Epidemiological Study of the Behavioral Problems of Young Children by H.Z. Reinherz, et al.	119 - 131
10. Self-Attitudes and Deviant Responses: Toward the Validation of a General Theory of Deviant Behavior by H.B. Kaplan	132 - 140
11. Family and Peer Processes in Adolescent Drug Use by D.B. Kandel	141 - 154
12. Adolescence to Young Adulthood : A Twelve-year Prospective Study of Problem Behavior and Psychosocial Development by R. Jessor & S.L. Jessor	155 - 168
13. Using Longitudinal Data to Study College Impact by A.W. Astin	169 - 184
14. Project TALENT: A Longitudinal Study of the Development and Utilization of Individuals' Capabilities by L. Steel, L.L. Wise & R.P. Abeles	185 - 208
15. Early Signs of Delinquency and Related Behavioral Difficulties: A Longitudinal Study by G. Spivack	209 - 228
16. The National Longitudinal Surveys of Labor Force Behavior by H. Parnes & K. Sproat	229 - 254
17. The Southern Youth Study: Longitudinal Survey Research in the Deep South by J.S. Picou, J.K. Thomas & A.G. Cosby	255 - 264
18. The Florida Longitudinal Project : A Review by J.M. Fletcher, P. Staz & R. Morris	265 - 284

19. Longitudinal Health Research in the U.S. Navy by E. Gunderson, R.E. Mitchell & R.J. Biersner	285 - 304
20. The Veterans Administration Normative Aging Study by R. Bosse, D.J. Ekerdt & J.E. Silbert	305 - 324
21. Longitudinal Studies of Impaired Competence in the Community: A Description of the Riverside Community Research Project (1963 - 1979) by E.W. Butler & S.S. Friedman	325 - 342
22. The Nebraska Annual Social Indicator Surveys by D.R. Johnson & L.L. White	343 - 353
23. The Woodlawn Mental Health Longitudinal Community Epidemiological Project by S.G. Kellam, M.E. Ensminger, J. Branch, C.H. Brown & J.P. Fleming	354 - 367
24. The Midtown Manhattan Longitudinal Study by L. Srole & A.K. Fischer	368 - 382
25. City Children: The Family Research Project by T.S. Langner	383 - 404
26. Life Events, Role Conditions and Adaptation: A Panel Study of Social Stress by E.G. Menaghan, L.I. Pearlin, M.A. Lieberman & J.T. Mullan	405 - 415
27. The New York Longitudinal Study (NYLS) by A. Thomas & S. Chess	416 - 429
28. Longitudinal Predictive Study of Success or Failure in Birth-Control Programs by S. Oskamp & B. Mindick	430 - 454

While parental punishment for aggression at home when the child was 8 years old related to aggression in school at that time, it was not a predictor of how aggressive the child would be at age 19. Similarly, while rejection by parents when the child was 8 was the best predictor of concurrent peer-rated aggression, it was not a significant predictor of his or her aggression at age 19 as rated by peers. However, various measures of identification which related to aggression at age 8 continued to maintain this relation to aggression at age 19. This included both measures of internalization, such as the extent to which the children manifested guilt and confessed when they committed transgressions, and measures of modeling, such as how much the youngsters' expressive behaviors resembled either or both of their parents' behaviors, and especially what models they were exposed to on television. Social class-related variables continued to exhibit relations over the years, such as father's occupational status (the higher the social class, the more aggressive the female subjects) and mobility orientation of the parents (the more ambitious the father, the more aggressive the male at age 19).

Two cross validations of these findings were done. The original multivariate analysis had been based on data from 128 boys and 120 girls. These included all subjects for whom there were complete data, i.e., a score on every independent and dependent variable. In the first cross validation, all 427 subjects were included whether or not they had a score on each variable. The second cross validation used a 50% random sample of those subjects in the first cross validation. Those variables that appeared on the original regression analysis and held up on both cross validations are described as having excellent validity; those that held up only on one cross validation are termed fair; and those that appeared on neither cross validation are termed as of poor validity. For boys, TV violence and nurturance are of excellent validity; identification with mother, preference for girls' games and mobility orientation of parents are of fair validity; while ethnicity is of poor validity. For girls, parents' religiosity and child's guilt are of excellent validity; TV violence, identification with father and father's occupation, education and mobility orientation are of fair validity; IQ is of poor validity.

Let us examine the implications of these findings, which have thus far been stated in a very general way. In the ensuing discussion we will become more specific. The effects of punishment and instigation are short lived. While instigators are perhaps necessary antecedents to aggression, their urgency is strong only at that very time and any specific instigation probably dissipates over an extended period. Thus, after 10 years the circumstances in the home atmosphere that may have led to aggression in the 3rd grade no longer have much effect and other variables are more important in predicting later aggression.

CONTINUED

1 OF 10

Why do the identification variables successfully predict aggressive behavior both at age 8 and then again 10 years later? Two of the identification measures, children's confessing and guilt as reported by parents, are often viewed as indications of the child's internalization of parental interdictions and, when viewed in a dynamic model of behavior, are indications of conscience which develops through identification with parents.

It is instructive to consider the children's behaviors that are associated with high amounts of confessing and of guilt. These behaviors are a form of communication from the children -- self-disclosure about some negative or undesirable action on their part. The fact that a child admits a transgression to a parent suggests that that parent has supported such communications and is a probable reinforcing agent for the child. If so, a child with high scores on our measures of identification might well have a parent who uses a child management system that includes positive aspects. Such systems are much more effective in building behavior controls which are mediated by the child alone and are not as situationally or time bound as predominantly negative control systems. This would explain why identification variables are effective across time and why externally imposed conditions such as punishment are not. Further, by the time a child is 19, physical punishment by parents is no longer a viable control.

Perhaps the most consistent relation between identification and aggression occurred as a function of a child's identification with the parents in expressive behavior. As indicated above, the measure of identification was the similarity in ratings between parent and child. Examination of the aggression scores for high, medium and low identification subjects revealed that low identification with both parents is the most potent predictor of aggression, irrespective of the subject's sex. Thus, the hypothesis that identification with parents in certain motor behaviors such as walking, eating and in perceptions of body image would be related inversely to aggressive behavior was substantiated. These measures, obtained from both parents independently and from children in the 3rd grade classroom, were correlated significantly to aggression synchronously in the 3rd grade and longitudinally to aggression in the 13th grade. These findings lend support to the idea that what is termed conscience or internalization of parental proscriptions not only is copying of moral precepts and guilt for transgressions but also copying of manifest motor behaviors of the socializing agents.

As noted above, identification in terms of sex role was investigated by examining specifically sex typed behaviors in which the children engaged. Indeed it was found that a boy's preference for girls' games and activities was a highly significant indicator of lower aggression both synchronously and in later years. Boys' preference for girls' games and activities was inversely correlated with peer nominations of aggression

both in the 3rd and 13th grades. Although not statistically significant, preference by boys for boys' games and activities was also in the hypothesized direction: the greater the preference, the more the aggression, both in the 3rd and 13th grades. For girls, no statistically significant relations occurred but again the correlations were all in the hypothesized direction: girls' preference for boys' games and activities in the 3rd grade was positively related and their preference for girls' games negatively related to aggression in the 13th grade.

These data lend support to the notion that aggressive behavior may be in part learned. Although for every measure of aggression on which we collected data, boys consistently scored higher than girls, there were some boys who scored well within the range of girls' scores and some girls who scored like boys. These tended to be the subjects who preferred activities inappropriate to the stereotype of their own sex role. When boys opt for feminine games and activities the choice in itself seems to act as a suppressor of early and later aggression. Preference for feminine activities may simply be incompatible with aggressive responses.

The most dramatic of our findings relevant to identification (i.e., modeling of behavior) had to do with the television habits of the youngsters. As has been reported elsewhere (Eron, Huesmann, Lefkowitz & Walder, 1972), one of the best predictors of how aggressive a boy will be at age 19 is the violence of television programs he prefers at age 8. This longitudinal relationship is even stronger than the synchronous one between television violence at age 8 and aggression at age 8. The use of such statistical techniques as cross lagged correlations, path analysis, multiple regression and partial correlation, added support to the interpretation that early viewing of violent television caused later aggression. Thus it seems that the possibility of the child's imitating aggressive behavior is not limited exclusively to behavior exhibited by parents. As Bandura (1969) points out, children will copy the behavior of any significant model, presumably including models they view on television. It should be emphasized that this direct positive relation between violence of preferred programs and later aggression was true only for boys.

For those investigators concerned with longitudinal research, probably the most interesting finding was the relation of early to later aggression. Not only did the peer ratings taken at both age periods relate significantly to each other but peer ratings at age 8 also related significantly to self ratings of aggression at age 19. The consistency of aggressive behavior over time and across situations strongly suggests that aggression is a stable behavior which conforms to the definition of a trait. This does not imply any necessary underlying biological correlates but only that a child learns very early in life a distinctive way of responding which is then perpetuated because of the success it brings.

References

- Bandura, A. Principles of behavior modification. New York: Holt, Rinehart and Winston, 1969.
- Bandura, A., Ross, D. & Ross, S.A. Imitation of film-mediated aggressive models. Journal of Abnormal and Social Psychology, 1963, 66, 3-11.
- Bandura, A. & Walters, R.H. Social learning and personality development. New York: Holt, Rinehart and Winston, 1963.
- Darlington, R.D. Multiple regression in psychological research and practice. Psychological Bulletin, 1969, 69, 161-182.
- Dollard, J., Doob, L.W., Miller, N.E., Mowrer, O.H. & Sears, R.R. Frustration and aggression. New Haven: Yale University Press, 1939.
- Eron, L.D. Social and cultural factors in mental illness. Proceedings of the Rip Van Winkle Clinic, 1956, 7, 16-38.
- Eron, L.D., Huesmann, L.R., Lefkowitz, M.D. & Walder, L.O. Does television violence cause aggression. American Psychologist, 1972, 27.
- Eron, L.D., Walder, L.O. & Lefkowitz, M.M. Learning of aggression in children. Boston: Little, Brown, 1971. Reprinted, New York: Pergamon, 1977.
- Feshbach, S. & Singer, R.D. Television and aggression. San Francisco: Jossey-Bass, 1971.
- Freud, S. The ego and the id. London: Hogarth, 1923.
- Hathaway, S.R. & McKinley, J.C. The Minnesota Multiphasic Personality Inventory. New York: The Psychological Corporation, 1969.
- Hathaway, S.R. & Monachesi, E.D. Adolescent personality and behavior: MMPI patterns of normal, delinquent, dropout and other outcomes. Minneapolis: University of Minnesota Press, 1963.
- Huesmann, L.R., Lefkowitz, M.M. & Eron, L.D. Sum of MMPI Scales F, 4 and 9 as a measure of aggression. Journal of Consulting and Clinical Psychology, 1978, 46, 1071-1078.
- Lefkowitz, M.M. Some relationships between sex role preference of children and other parent-child variables. Psychological Reports, 1962, 19, 911-914.
- Osgood, C.E., Suci, G.J. & Tannenbaum, P.H. The measurement of meaning. Urbana: University of Illinois Press, 1957.
- Sears, R.R., Whiting, J.W.M., Nowlis, V. & Sears, P.S. Some child rearing antecedents of aggression and dependency in young children. Genetic Psychology Monographs, 1953, 47, 135-234.
- U.S. Department of Commerce, Bureau of the Census. County and City data book, 1952. Washington, D.C.: U.S. Government Printing Office, 1953.
- Walder, L.O., Abelson, R., Eron, L.D., Banta, T.J. & Laulicht, J.H. Development of a peer rating measure of aggression. Psychological Reports, 1961, 9, 497-556.
- Warner, W.L., Meeker, M. & Eells, K. Social class in America. New York: Harcourt, 1960.

ETS Study of Academic Prediction and Growth*

Thomas L. Hilton

*This chapter first appeared in New Directions for Testing and Measurement, San Francisco: Jossey-Bass, 2, 1979.

ETS Study of Academic Prediction and Growth

Thomas L. Hilton

Background and Description

Goals

The Study of Academic Prediction and Growth, sponsored by Educational Testing Service and the College Entrance Examination Board, was designed in 1959 primarily to provide "information on the academic growth of boys and girls at different times and under different circumstances of schooling" (Anderson & Maier, 1963, p. 212).

More specifically the study was designed to answer questions such as the following:

1. Do students grow at a fairly constant rate from grades 5 through 12 in ability to solve mathematics problems, to read with comprehension, to understand social studies materials, to understand science materials, to listen effectively and to write intelligently?
2. If students are "bright," do they grow faster than if they are "dull"?

3. To what extent are growth patterns influenced by the curriculum in which students find themselves, and by characteristics of their school?
4. What special individual background and experience factors influence rate of academic growth--factors like home environment, recreation, hobbies, reading, and career goals?
5. How well can 12th-grade academic standing be predicted by samples of performance taken at the 5th grade, at 7th grade, at 9th grade, and at 11th grade? How early can academic talent be positively identified?
6. Which students drop out of school? What factors might have allowed prediction of the likelihood of their dropout?

Data Collection

The testing plan for the Growth Study incorporated both a cross-sectional and a longitudinal design: In 1961, different students in grades 5, 7, 9, and 11 were tested and after that the same students were tested every two years until they were graduated from high school. This design enabled the Growth Study research staff to proceed immediately with the descriptive summary of the cross-sectional data while waiting for the longitudinal data to come in. It also permitted the investigation of differences in growth estimates which are obtained from cross-sectional as compared to longitudinal data (see Hilton & Patrick, 1970).

Sample Design

The schools were selected to provide a range in geographical location, size of system, proportion of senior class graduates who subsequently attended college, and rural-urban orientation. Although all the high schools were comprehensive, the proportion of students enrolled in vocational education curricula varied markedly from one school to another. The average enrollment in such programs was approximately 40 percent.

Although the schools were not randomly selected from the population of high schools in the United States, comparisons that have been made with national probability samples indicate that the Growth Study sample fairly closely approximates the randomly selected samples in aptitude and achievement. For instance, in the Study of Equality of Educational

Opportunity (the Coleman Study), 13 percent of the fathers were classified as laborers, 19 percent as skilled, and 7 percent as professional. The comparable percentages for the Growth Study total sample are 12 percent, 14 percent, and 9 percent. Certain differences in percentages can be attributed to differences in the wording of relevant questionnaire categories. (The Growth Study staff could not resist the temptation to improve the wording of successive questionnaires.) In any case, the occupational profiles of the two samples are sufficiently similar to suggest that the Growth Study sample probably provides a sampling of U.S. students from grade 5 to grade 12 that is suitable for studies of academic growth and its correlates.

As far as school characteristics are concerned, comparison of the Growth Study sample with the demographic characteristics of the nation's schools indicate the Growth Study is deficient in its representation of small rural high schools. Again, however, this is not regarded as serious for most analytical purposes.

Sampling Within Schools. The Growth Study sample consisted of the students in these public schools and in all the junior high schools and elementary schools feeding them. Whenever a test administration was conducted, all the students enrolled in the school were tested except those not in attendance because of prolonged illness and those students classified as "mentally retarded." (In 1960, schools were less precise in their classification of handicapped children than they are now.) The initial sample consisted of approximately 9,000 fifth graders, 9,000 seventh graders, 9,000 ninth graders, and 5,000 eleventh graders, giving a total of approximately 32,000 subjects. (The precise number in any particular subsample of the total Growth Study sample depends on what restrictions individual researchers place on selection of the sample from the total file, that is, on whether complete data are required.)

In the spring of 1963, the original eleventh graders were tested as high school seniors.

By the fall of 1963, when all subjects--except for the original eleventh graders who had graduated--were tested for the second time, 80 percent of the original subjects remained but (since the Growth Study made a practice of testing all the students in the relevant grades of the school, whether or not they had participated in previous administrations) the total sample--with full or partial data--had grown to slightly less than 40,000. By the time of each subsequent testing, about 20 percent of each cohort had left the sample school systems or, for one reason or another, were not tested. But "new" students entered the sample, keeping the volume of testing for each cohort approximately level, until senior high school when dropouts took their toll. When students left "Growth Study"

schools, no effort was made to follow them to their new location, primarily for financial reasons. The number who were tested at each grade level is as follows. (note that for each cohort the number tested at each grade level is equal to the number tested at the previous grade level plus the number of "new" students minus the number who left the sample.)

Cohort	5	7	9	11	12
4	8,939	8,361	7,651	6,304	5,674
3		8,891	8,724	7,383	5,810
2			9,245	7,790	6,750
1				5,294	4,854

In all, 45,901 students took one of more tests as part of the in-school data collection. Of these, 15,124 had complete data on all variables included for their cohort and comprise the core sample. For the cohort of 8,939 tested in grade 5 in 1961, 3,476 had complete data for grades 5, 7, 9, 11, and 12. This shrinkage from 8,939 to 3,476 reflects the high mobility of American public school students from one high school sending district to another and the drop-out of students from senior high schools.

Most analyses based on the Growth Study data file have been restricted to Cohort 4 who were tested as fifth graders in September or October of 1961 and as twelfth graders in January or February of 1969. To the author's knowledge, this sample is the only national sample of public school students who have been tested on five occasions over an eight year span.

Variables and Tests. The original focus of the Growth Study was on academic growth, as measured by objective tests of ability and achievement, recognizing that this domain is but one aspect of the total development of the student. Supplementary biographical information--on interests, both academic and nonacademic, and educational and vocational plans--was obtained from questionnaires, but throughout the data collection the primary objective remained that of thoroughly describing and explaining the student's acquisition of knowledge, understanding, and intellectual skills.

The basic instruments of the study are the School and College Ability Test (SCAT) and the Sequential Tests of Educational Progress (STEP). The SCAT, which yields verbal and quantitative scores, measures general ability to do school work. The STEP series of achievement tests measure the students' ability to apply their skills to the solution of problems in six areas: reading, writing, listening, social studies, science, and mathematics. Lower level SCAT and STEP forms are scored on the same continuous scale as higher level forms (that is, the tests are vertically

equated) but the tests are not equated horizontally: A given scaled score on one test cannot be interpreted as the same score on another test.

A Test of General Information (TGI) was designed for the Growth Study in order to identify nonacademic factors that are related to students' growth during this period. It has been used to determine when and at what rate students learn the kinds of facts that are not systematically taught in school but which an alert and reasonably well-informed adult could be expected to know. There are three forms of this test at three different levels of difficulty with questions in each of eight areas: home arts, industrial arts, physical sciences, art and music, biological sciences, history and literature, entertainment and recreation, and public affairs.

The SCAT subtests are long enough (twenty-five minutes, with fifty to sixty questions, for Verbal and forty-five minutes, with fifty questions, for Quantitative) to provide reliable measurement of individual performance. Reliabilities for SCAT Verbal range from .92 to .94, for SCAT Quantitative from .88 to .93, and for SCAT Total from .95 to .96.

In January or February of grade 12, the sample subjects were given the College Board American History, the College Board English Composition, and the Preliminary Scholastic Aptitude Test--unless the subjects had already taken these tests, in which case their scores were retrieved from the files of ETS.

Questionnaires. A Background and Experience Questionnaire (BEQ) was also designed especially for the Growth Study. It provides information on the relationships between the students and their experiences and activities in and out of school. These questionnaires attempt to minimize generalization and vagueness by requiring the students to document many of their responses with facts, such as the titles of some books they have read, the plays they have seen or participated in, the instrument played in the school band, or a description of the job held during the summer. The questionnaires also yield information on the students' opinion of the courses they have taken and indicate how often and with whom they have discussed topics such as their future education and career plans, news events of the day, or personal values.

Questions on the students' background focus on the education and occupations of their parents and the amount of encouragement and support they give. Forms of the Background and Experience Questionnaire were developed for the seventh, ninth, eleventh, and twelfth grades.

The twelfth grade form of the BEQ was given only to the 1963

seniors. The 1965, 1967, and 1969 seniors, who earlier completed the eleventh grade versions, were given a shortened version asking only about their educational and vocational plans.

Other Data. For Cohort 4, which graduated in 1969, transcripts of academic grades were purchased from the participating schools. In addition, the senior class yearbooks for all Cohort 4 sample members were obtained. The yearbooks have never been used systematically except as a partial source of racial data.

Two types of variables have been defined for the purposes of various analyses of the Growth Study data file. The first, classification variables, were selected for the purpose of subdividing the total sample into subgroups of special interest. The second type were the dependent variables whose distributions were examined.

Classification Variables. In the most exhaustive investigation of the Growth Study data file, the Study of Stability and Instability in Academic Growth (Hilton, Beaton, and Bower, 1971), the classification variables were sex, ethnic group, high school curriculum, and father's education. These particular variables were selected, primarily on the basis of previous analyses of the Growth Study data file and the educational research literature, as most likely to be major sources of variance in the dependent variables and to be of interest to researchers as control variables.

The ethnic variable included only two categories: Black, White and other. Because at the time the Growth Study was begun the sociopolitical climate was such that it was not feasible to include an item for self-identification of ethnic group, the identification was accomplished several years later by engaging school guidance counselors who knew the students to make the identification from school rosters. After the students were graduated from high school in 1969, the Growth Study staff repeated the process, using high school yearbooks as the source of identifying information. In the few cases where there were disagreements between the staff and the counselors (less than 1 percent) the student in question was classified as White. Thus the "white" category--while mostly white students--also includes Mexican-Americans, Orientals, Latin-Americans and students whose identification was uncertain.

The academic-nonacademic dichotomy was made on the basis of a questionnaire item included in the Background and Experience Questionnaire in 1967 in which the students, then in Grade 11, identified their course of study:

125. From the list below, which course of study are you taking in high school?

- A. Academic or college preparatory
- B. Agricultural
- C. Business or commercial
- D. General
- E. Home economics
- F. Vocational
- G. Other*
- H. Undecided

*What? _____

Students responding "A" were classified as academic and all others were classified as nonacademic.

Father's education was selected as the single variable most likely to reflect the socioeconomic background of the student and also the educational opportunities the student may have had in the home setting. An exploratory study conducted by the authors indicated that--unlike father's occupation--father's education was related to achievement in a positive linear fashion for both white and black students. Father's education was obtained from the following BEQ item:

113. How much formal education does your father or male guardian have?

- A. Grade school
- B. Some high school
- C. Graduated from high school
- D. Some college, junior college, business or trade school (after completing high school)
- E. Graduated from college
- F. Some graduate or professional school (e.g., law, medicine)
- G. Obtained a graduate or professional degree
- H. Don't know

G. Categories A and B were combined, and also categories D, E, F, and

Dependent Variables. The test scores which served as dependent variables for the descriptive summary are shown in Table 1.

Included are eleven scales derived from items in the BEQ on the basis of a factor analysis of all the items in the BEQ (Rock and Freeberg,

1969). The items selected for each scale were those which had the highest loading on each successive factor and also were consistent with the other items on logical grounds. In computing scores for each subject, the problem of missing item responses arose. It was decided that the subject had to have responses for 75 percent of the items before a score on that scale could be computed for him or her. For all subjects, the final score was the average score for the items responded to; these mean scale scores were standardized across the entire sample, with a mean of 50 and a standard deviation of 10.

Operation

The data collection was conducted essentially as a small ETS testing program. Test materials and questionnaires were shipped to the participating schools in the early fall with instructions that the tests be kept under secure conditions. The tests typically were administered by classroom teachers who were provided with detailed instructions on exactly how to handle the materials and conduct the testing. A coordinator in each school, usually a member of the guidance department, supervised the test administrations. The very small fraction of unusable answer sheets received suggests that the test administrators were unusually successful in conducting orderly test sessions.

The scheduling of the test administrations was left to the schools, some of which set aside two or three full days for the data collection while others spread it over several weeks. Typically, the schools conducted one make-up session for students who were absent from a particular test administration. Some schools used auditoriums, gyms, or cafeterias for the test administration while others used classrooms. In many schools, the STEP Listening test was played over the school public address system.

After the answer sheets were checked and scored at ETS, the schools were sent both rosters with test scores and score reports for individual students. Typically, these were filed with the students' records and used for guidance purposes.

After the first data collection in 1961, those data and data collected in subsequent data collections were merged in a central data file. This proved to be a time consuming and expensive process as a result of the many errors students made in gridding answer sheets, especially in gridding their birthdates. Quality control measures were executed to minimize errors in the data tapes.

From the inception of the study, the Growth Study staff has kept in close touch with the sample schools through newsletters and periodic visits.

During the data collection period, annual meetings (held in Princeton) to which representatives from each Growth Study school system were invited were a particularly useful form of communication. Research results were presented and operational problems were discussed.

Data Analyses and Reporting

Most of the analyses of the file have been designed by Growth Study staff members with an eye to the questions the Study was designed to address, or by other ETS staff members and non-ETS researchers pursuing their individual research interests. Initially, most of the analyses conducted involved simple descriptive statistics (means, variances, and zero-order correlations); later, the analyses became more complex, employing conventional multivariate techniques (multiple regression and multivariate analyses of variance). More recently the analyses have made use of path analytic techniques and Jöreskog's procedures (1970).

Reporting of the research has taken three main forms: (1) final reports submitted to government agencies that supported the research (most of these reports are in the Education Research Information Clearinghouse (ERIC); (2) ETS Research Memoranda and Research Bulletins, copies of which are available from ETS; and (3) journal articles. An annotated bibliography describing seventy-eight reports and articles is available from the author of this chapter.

Nature and Availability of Data

Much of the information carried on the Growth Study master data tapes is of a confidential nature, and has been collected with the understanding that ETS would respect the privacy of both individual students and individual schools. As a matter of policy, therefore, release of data from the Growth Study files to persons or organizations outside of ETS is subject to at least the following conditions: (1) Student names will be deleted, and students will be identified only by their internal Growth Study I.D. number; (2) schools will be identified only by their internal Growth Study school code. The key to this code will not be released; and (3) in any cases where students or schools must be identified--as in the case of follow-up studies, for example--all processing and file matching will be performed by ETS, and the data will be returned in such a form as will preserve the anonymity of individual

students and schools.

Guided by these restrictions, ETS has frequently released copies of the Growth Study tape to qualified researchers. In fact, approximately twenty copies of the Growth Study data file are currently in use or accessible in various educational research laboratories and centers in this country. At least five doctoral dissertations are known to have been based on Growth Study data, as well as a large number of journal articles and parts of monographs and books.

The only requirements imposed by ETS are, first, that the researcher agree never to publish results in such a way that a particular student or school can be identified. Even though all identifying information is removed from the tape before it is sent, a person with enough ingenuity might be able to single out individual students or schools. Second, the staff requests that the researcher acknowledge the source of the data in any reports that result from use of the tapes. Third the staff requests that ETS receive a copy of any reports resulting from use of the data file.

A modest charge is made for the tape. The amount takes into consideration the staff's experience that usage of the tape usually entails a number of time-consuming telephone calls and correspondence with ETS staff members, even though a tape layout and description of the variables is provided.

The tape is BCD format and contains the following information:

- Subject's cohort number
- Date of birth
- Social security number (if any)
- Code for schools attended in 1961, in 1963, in 1965, in 1967 and for senior year
- Curriculum enrolled in each year
- Code for reason for not attending particular testing session
- Scores for each subtest of SCAT, STEP, and the TGI in four occasions
- Rank in senior class
- Responses to 7 items of Senior Questionnaire
- Senior year SAT-V or PSAT-V
- Senior year SAT-M or PSAT-M
- Senior year History Achievement
- Senior year English Achievement
- Responses to 169 items of BEQ
- Drop-out indicator
- Race code
- (Information on the last item has not been collected for the entire sample.)

Plans for Subsequent Work

Although the Growth Study data file is in use in a number of universities and research centers throughout the United States, no analyses are currently in progress at ETS. It is expected, however, that additional studies will be conducted as relevant substantive and methodological questions arise. In the opinion of the author, a number of studies conducted previously should be pursued further provided that funding for the studies can be obtained. The studies conducted to date have by no means exhausted the data file. It would be especially valuable to conduct a long-term follow-up of the students who graduated from high school in June, 1969, for the purposes of investigating relationships between patterns of academic growth, personal background, experiences during the elementary and high school years, and long-term career development.

Table 1. Dependent Variables

SCAT	Verbal Quantitative
STEP	Math Science Reading Social Studies Listening Writing
TGI Scale	A Industrial Arts B Home Arts C Physical Sciences (Math) D Biological Science E Musical Arts F History, Literature G Recreation, Entertainment H Government, Public Affairs
BEQ Scale	1 Female Activities 2 Sports and Masculine Activities 3 General Leisure Activities 4 Breadth of Reading 5 Talk with Others 6 Academic Course Interests 7 Musical Activities 8 Lower Level Social and Leisure Activities 9 Technical Hobbies and Interests 10 General TV Viewing 11 Academic Effort

References

- Alexander, K. L., Cook, M., & McDill, E. L. Curriculum tracking and educational stratification: Some further evidence. American Sociological Review, 1978, 48, 47-66.
- Anderson, J. E. The limitations of infant and pre-school tests in the measurements of intelligence. Journal of Psychology, 1939, 8, 351-379.
- Anderson, S. B., & Maier, M. H. 34,000 pupils and how they grew. Journal of Teacher Education, 1963, 14, 212-216.
- Atkin, R., et al. Cross-lagged panel analysis of sixteen cognitive measures at four grade levels. Child Development, 1977a, 48, 944-952.
- Atkin, R., et al. Ability factor differentiation, grades 5 through 11. Applied Psychological Measurement, 1977b, 1 (1), 65-76.
- Clearly, T. A. The prediction of performance in independent secondary schools (SSAT Report 9). Princeton, N.J.: Educational Testing Service, 1966.
- Coleman, J. S., et al. Equality of education opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.
- Connolly, J. A. The contribution of clusters of biographical data to prediction. Princeton, N.J.: Educational Testing Service, 1968.
- Evans, F. R. & Patrick, C. Antecedents and patterns of academic growth of school drop-outs (RB 69-81). Princeton, N.J.: Educational Testing Service, 1969.
- Hilton, T. L. A study of intellectual growth and vocational development. Interim report, Project No. 6-1830. Princeton, N.J.: Educational Testing Service, 1967.
- Hilton, T. L. A study of intellectual growth and vocational development. Final report to USOE, Project No. 6-1830, PR 71-1. Princeton, N.J.: Educational Testing Service, 1971. (ERIC Document Reproduction Service No. ED 056 063).
- Hilton, T. L. Predictability and intellectual growth--some comments on the degree and interpretation of growth correlations (RM 72-1). Princeton, N.J.: Educational Testing Service, 1972.

- Hilton, T. L., Beaton, A. E., & Bower, C. P. Stability and instability in academic growth--a compilation of longitudinal data. Final report to USOE, Research No. O-0140. Princeton, N.J.: Educational Testing Service, 1971. (ERIC Document Reproduction Service No. ED 065 514)
- Hilton, T. L., & Berglund, G. W. Sex differences in mathematics achievement--a longitudinal study. The Journal of Educational Research, 1974, 67 (5), 231-237.
- Hilton, T. L., & Gaines, L. K. Academic growth in Penn City (Chapter I). Volume II of a report from Educational Testing Service to the State Board of Education of Pennsylvania, Princeton, N.J.: Educational Testing Service, 1965.
- Hilton, T. L., & Myers, A. E. Personal background, experience, and school achievement: An investigation of the contribution of questionnaire data to academic prediction. Journal of Educational Measurement, 1967, 4 (2), 69-80.
- Hilton, T. L., & Patrick, C. Cross-sectional versus longitudinal data: Empirical comparison of Mean score changes. Journal of Educational Measurement, 1970, 7 (1), 15-24.
- Hilton, T. L., & Pingel, L. A. Community socio-economic status, aptitude, and school achievement. Paper presented at the meeting of the National Council on Measurement in Education, Chicago, 1965.
- Humphreys, L. G. Investigations of the simplex. Psychometrika, 1960, 25, 313-323.
- Humphreys, L. G., & Parsons, C. K. A simplex process model for describing differences between cross-lagged correlations. Psychological Bulletin, 1979, 86 (2), 325-334.
- Jöreskog, K. G. Factoring the multitest-multioccasion correlation matrix (RB 69-62). Princeton, N.J.: Educational Testing Service, 1969.
- Jöreskog, K. G. A general method for analysis of covariance structures. Biometrika, 1970, 57 (2), 239-251.
- Kenny, D. A. The measurement and explanation of populations effects: Sex differences in mathematics and science in a longitudinal study. Unpublished dissertation, Northwestern University, 1972.
- McGaw, B., & Hilton, T. L. Similarity of factor structure in longitudinal data for black and white students. Unpublished manuscript, Educational Testing Service, 1970.
- Patton, M. J. Educational achievement prior to and during high school (RB 69-29). Princeton, N.J.: Educational Testing Service, 1969.

- Patton, M. J. & Morse, L. K. Curricular groups, school environments, and patterns of academic achievement. In T. L. Hilton et al. (Eds.), A study of intellectual growth and vocational development (Interim report, Project No. 6-1830). Princeton, N.J.: Educational Testing Service, 1967.
- Rock, D. A., & Freedberg, N. E. Factorial invariance of biographical factors over three grade levels. Multivariate Behavioral Research, 1969, 195-210.
- Rosenfeld, M., & Hilton, T. L. Negro-white differences in adolescent educational growth. American Educational Research Journal, 1971, 8 (2), 267-283.
- Werts, C. E., & Hilton, T. L. Intellectual Status and intellectual growth, again. American Educational Research Journal, 1977, 14 (2), 137-146.

The Youth In Transition Project*

Jerald G. Bachman and Patrick M. O'Malley
Survey Research Center, Institute for Social Research
The University of Michigan
Ann Arbor, Michigan

*Major funding for the data collections and analyses which form the basis of this chapter was provided by the U.S. Office of Education, the National Institute of Education, and the National Institute on Drug Abuse. The authors gratefully acknowledge the major professional contributions to the Youth in Transition project by Terrence Davidson, Jerome Johnston, Lloyd Johnston, Robert Kahn, and Martha Mednick.

The Youth in Transition Project

Jerald G. Bachman and Patrick M. O'Malley
Survey Research Center, Institute for Social Research
The University of Michigan
Ann Arbor, Michigan

Introduction

This chapter summarizes many of the findings of the Youth in Transition project. The design and findings have been spelled out in a series of six monographs (Bachman, Kahn, Mednick, Davidson and Johnston, 1967; Bachman, O'Malley and Johnston, 1978; Bachman, 1970; Bachman, Green and Wirtanen, 1971; Davidson, 1972; Johnston and Bachman, 1972; plus a number of other reports and articles (Johnston, 1973(a), 1973(b); Johnston, O'Malley and Eveland, 1978; Bachman and O'Malley, 1977; Johnston and Bachman, 1973; Bachman and Jennings, 1975; Bachman and O'Malley, unpublished; Wilson and Portes, 1975). In particular, the present chapter is drawn largely from Volume VI in the Youth in Transition monograph series (Bachman, O'Malley and Johnston, 1978).

Purposes of the Youth in Transition Project

Stated most broadly, the Youth in Transition project was designed to be an exploration of the effects of social environments and experiences, with special emphasis on the impacts of school and work. More specifically, the primary objectives were:

1. To learn more about the degree to which measures obtained at the start of high school (i.e., "input characteristics") can predict later educational and occupational attainments.
2. To assess the role of educational attainment in occupational attainments, particularly as compared with the role of family background and intellectual ability.
3. To explore the effects of dropping out of high school, particularly by comparing the occupational attainments of dropouts with graduates.
4. To explore whether different high schools have differential impacts on the outcomes of their students.
5. To determine the impacts of various post-high school environments and experiences (educational, occupational, military, marital, and parental) on values, attitudes, and behaviors. Included are such dimensions as: occupational attitudes; job satisfaction; self-esteem; self-reported delinquent or criminal behavior and use of drugs; views on race relations, government activities, and the military.

The multiple and overlapping objectives outlined above require that variables serve several different "roles" as we move from one type of analysis to another. Thus, for example, self-esteem and delinquent behaviors are among the variables we explore as possible predictors of educational outcomes; but we also explore educational attainment as a possible predictor (and determinant) of changes in self-esteem and delinquent behaviors.

The fact that many of our variables serve in dual analysis capacities illustrates a point which may be obvious, but is nevertheless very important: there are a variety of data collection and analysis strategies that can make use of a variety of data collection and analysis strategies that can make use of longitudinal designs. The Youth in Transition project concentrates heavily on the analysis of change, thus requiring repeated measurement of the same dimensions (values, attitudes, behaviors) at multiple points in time. But the project also employs the longitudinal strategy in which one set of measures obtained at an early point in the study is used to predict a different set of outcome measures obtained later.

Some comments may be in order concerning the fact that the Youth in Transition project is limited to an all-male sample. The project began with a sharp focus on the causes and consequences of dropping out of high school, and a broader interest in the dynamics of educational and occupational attainment. For males during the high school years the two major role possibilities are student and/or worker (or neither, in the case of unemployed dropouts). The additional role possibilities of parent and homemaker are remote for practically all males during the high school years; but for females, especially those who drop out, these roles are often quite central. In other words, the range of causes and consequences of dropping out of high school

differ for males and females, and in the case of females the impacts of school and work environments may be less important than other forces. Given these additional complexities for females, it seemed clear to us that if we were to use a sample of both males and females we would have to carry out two rather separate streams of analysis and reporting (much as Parnes and his colleagues have done with the separate male and female samples in the National Longitudinal Surveys). We opted, instead, to begin our research with what we considered to be the least complicated approach--a study of males. We recognized, of course, that a parallel study of young women could be undertaken at a later point.

Now, some fifteen years after our initial decision to concentrate on an all-male sample, we strongly suspect that our most important findings about change and stability apply just about as clearly to young women as to young men; but we regret that we cannot demonstrate that directly using a Youth in Transition sample that included both genders. (On the other hand, it may be worth noting that we are currently involved in a new longitudinal study, which samples both young men and young women; and preliminary analyses from that study provide some support for our suspicion that many of the most important findings from the Youth in Transition project apply to both genders.)

Research Design and Methods

The research design centers on a panel of adolescent males chosen in 1966 to be representative of young men entering tenth grade in public high schools in the contiguous United States. The panel agreed to be surveyed at intervals of one year or more for an indefinite period. For those who did not drop out of high school, the first three data collections occurred while panel members were still in school; the first took place in the fall of 1966 when the subjects had just entered tenth grade, and follow-up data collections occurred in the spring of 1968 (the end of eleventh grade for most) and the spring of 1969 (just before most were graduated). The last two data collections correspond to one year and five years beyond high school: the spring of 1970 and the spring of 1974. (Additional data concerning school environments were obtained from principals, counselors, and samples of teachers in each of the schools which participated in the study.)

This sequence of five data collections from the panel is, incidentally, a considerable improvement over the three data collections projected at the beginning of the study in 1965. Details on the original study design can be found in Volume I of the Youth in Transition monograph series (Bachman et al., 1973).

Sample Design and Response Rates

The study began with 2,213 tenth-grade boys in 87 public high schools. The schools were located in communities across the United States, corresponding to the primary sampling units used by the University of Michigan's

Survey Research Center. Schools and boys were selected through a multi-stage sampling design in such a way that the probability of a school's selection was proportionate to its size (i.e., the estimated number of tenth-grade boys), and roughly equal numbers of boys (about 25) were selected from each school. The net effect of this design was to provide an essentially unbiased representation of tenth-grade boys in public high schools throughout the contiguous United States.

Response rates for the initial data collection were considered at two levels: schools and student. In the initial phase of sampling, 88 schools were invited to participate; 71 consented, and replacement schools in the same geographic areas were secured for all but one of the remaining schools. In the resulting 87 participating schools, 2,277 boys were invited to participate in the study; 2,213 (over 97 percent) of those students participated in the first data collection.

Table 1 presents a summary of data collections and the response rates throughout the study. As the table indicates, more than 73 percent of the young men who began the study in 1966 were still participating in 1974. For further details on sample characteristics and data collection procedures, see Youth in Transition, Volume VI (Bachman et al., 1978).

Measurement Content

A large number of dimensions were measured in the five data collections from the young men. The initial data collection included tests of ability and academic skills, measures of family background characteristics, and a large number of "criterion" dimensions: affective states, self-concepts, values and attitudes, plans, and behaviors. These measures are described in Youth in Transition, Volume II (Bachman, 1970), and most of them are summarized in Appendix A of Volume VI (Bachman et al., 1978). Many of the criterion dimensions were repeated in all five data collections (see Table 2), thus permitting a fairly detailed assessment of change during the eight-year span of the longitudinal design.

Sources of Bias in a Longitudinal Study

Two special types of bias are possible in longitudinal studies. First, repeated data collections may lead to changed answers; that is, individuals who answer some of the same questions two or three years in a row may show systematic changes in responses simply because of the repetition. Second, certain types of individuals may be especially likely to leave the panel of subjects, thereby biasing the sample that remains in later data collections.

Repeated Measurement

The first problem, repeated measurement effects, was explored by using a small control group sample of individuals who were selected at the time of the initial data collection (1966) but not contacted until the fourth

data collection (1970). One hundred and fifteen control subjects completed the fourth data collection, and their responses were compared with corresponding subjects from the regular sample. The results clearly indicated the absence of repeated survey effects; those who had been participating in the study for nearly four years gave answers similar to those of their classmates from the control group who were answering for the first time.

Panel Losses

As noted earlier, 74 percent of the original panel were still participating in the study some eight years after the initial data collection. This is a highly acceptable retention rate for a panel study and compares favorably with the response rate for many one-time, cross-sectional surveys. But what about those who left the study? Do they come disproportionately from particular subgroups, thus biasing the sample that remain? On the basis of the initial (time 1) data collected from nearly all subjects in the sample, it was possible to compare those who left the study with those who participated throughout the study. On the whole, the composition of the sample at Time 5 was remarkably similar to the original sample from Time 1. The most important exception involved high school dropouts, who were lost disproportionately from the sample. This bias should not affect the shape of any of the relationships noted, and should affect the population estimates of the strength of relationships only in the case of variables highly correlated with dropping out. For an extensive discussion of panel losses, see Youth in Transition, Volume VI (Bachman et al., 1978).

Predicting Educational and Occupational Outcomes

Why do some individuals extend their educations through college and into graduate school while others drop out of high school? Why do some get high status and/or high paying jobs while others do not? What are the factors of background, ability, earlier educational experience, plans, values, attitudes, and behaviors that predict--and perhaps help to determine--later levels of attainment? To what extent do such basic factors as family background and ability show a direct or continuing effect on educational attainment, and to what extent do they appear to have an "indirect" impact via their influence on other factors such as educational success in the pre-high school years?

The Youth in Transition project attempted to deal with such questions by examining a large number of possible predictors of educational and occupational attainment. More specifically, we focused on how factors measured in tenth grade relate to the attainments of young men in their early twenties. Two considerations governed our selection of a large array of potential predictors. First, we were heavily influenced by sociological studies of educational and occupational attainment (Wilson and Portes, 1975; Blau and Duncan, 1967; Alexander, Eckland and Griffin, 1975; Sewell and Hauser, 1975, 197). Second, we sought to understand the extent to which factors which appear as differences after various levels of education have been attained are actually present as differences by the start of

high school. Consequently we included a wide range of attitude and personality dimensions. (The inclusion of such variables, and the emphasis on repeated measures, as a means for distinguishing between the causes and consequences of educational attainment, are special foci of the Youth in Transition project. Both are developed more fully in the next section.)

Educational Attainment as an Outcome

After a considerable amount of exploration, we decided to use an eight-category measure of educational attainment for correlational analyses. The eight categories, and the percentages of Youth in Transition panel members classified into each as of 1975 (five years beyond high school for most), are as follows:

1. No high school diploma (dropouts)--7 percent
2. High school diploma only--25 percent
3. Some post-high school training other than two-year or four-year college--13 percent
4. Some two-year college--10 percent
5. Some four-year college--14 percent
6. Associate degree--9 percent
7. Bachelor degree--15 percent
8. Bachelor degree plus some graduate work--7 percent

This eight-category measure of educational attainment, which is keyed largely to academic degrees, shows a strong and linear correlation with a separate measure of the total number of years of schooling completed ($r = .89$). When both of these measures were correlated with a large variety of potential predictors, in the majority of cases the eight-category measure (treated as an equal-interval scale) showed very slightly higher correlations. For this and other reasons we chose to concentrate on the eight-category measure in our analyses; however, the results would be very similar had we used the years-of-schooling variable as our measure of educational attainment.

About fifty different variables measured in the initial data collection (done in tenth grade in 1966) were explored as possible correlates of educational attainment. We began by using a wide array of background and ability measures in a multiple regression analysis, and were able to account for 36.3 percent of the variance in educational attainment (adjusted for degrees of freedom). Three of the most important of these measures, a composite measure of socioeconomic level, a composite of three ability tests, and number of siblings, were able to explain nearly as much variance (34.5 percent, adjusted) as the larger set.

In sum, it appears that by knowing background, ability, and a few indicators of success in school and attitudes toward schooling, all measured at the start of tenth grade, we can do a rather good job ($r = .71$) of predicting later educational attainment. (As a matter of fact, we could

"account for" still more variance in educational attainment if our predictive equation included college plans measured at the end of twelfth grade. However, most young men who go to college have made and received firm commitments by the end of their senior year in high school. Thus, our twelfth-grade measure of "colleg Plans," which correlates .67 with later educational attainment, might better be termed a measure of "college arrangements.")

A key finding from our multivariate analysis is that much of the impact of background and ability overlaps the impact of schooling factors. We take this as evidence that a substantial proportion of the impact of background and ability is indirect. In other words, background and ability contribute to pre-high school educational experiences and successes, which in turn contribute to later educational attainment.

Occupational Attainments

One of the strongest selling points for education has been that higher levels of education lead to better jobs--better in the sense of status and income, freedom from insecurity (risk of unemployment), and satisfying and rewarding working conditions. In this section, and several which follow, we consider ways in which educational attainment and a number of other factors may affect occupational outcomes five years beyond high school.

Some of our findings can be summarized very briefly. For example, we found that amount of education showed very little effect on the hourly wage rates of young men five years out of high school. This lack of relationship is attributable to the fact that these young men were in the very early stages of their occupational careers. After a longer time interval it seems certain that wage rate differences linked to educational level will emerge quite clearly.

There were also few differences in employment rates (versus unemployment rates) associated with level of attained education, with the very important exception of high school dropouts (discussed below). Excluding dropouts, about 7.5 percent of the respondents were unemployed. Perhaps more differences among the educational levels would have been evident had the unemployment rate been still higher, since historically the less educated have been most affected by increased rates of unemployment.

Unlike wages and unemployment rates, job status was rather strongly correlated with educational attainment. In particular, college graduates held the higher status jobs, while there were only rather small status differences among the other educational levels.

In addition to exploring impacts of educational attainment, we examined variations in occupational outcomes as a function of a variety of other environments and experiences, including military service, marital/parental status, urbanicity, region, and county-level labor market conditions. In general, relationships were weak, particularly when a few control

variables were introduced. More unemployment was found among veterans, among single respondents, in urban areas, and in the non-South regions. Wages and status varied little, but were higher in the most urban areas.

The one critical difference between high school dropouts and graduates lay not in the kinds of jobs they held, but in whether they held any job at all. Being a dropout--lacking a high school diploma--seems to double the risk of being unemployed. After adjusting for family socioeconomic level, number of siblings, and intellectual ability, the dropouts were still almost twice as likely as graduates to be unemployed. The contrast here is somewhat sharper and stronger than that found in our earlier analysis (Bachman et al., 1971) it indicates a serious continuing social problem, which we discuss later.

Educational Attainment and Job Satisfaction

Another aspect of occupational attainment is the individual's subjective sense of job satisfaction. One of our research questions was to see whether education seems to increase job satisfaction. The answer to that question, based on our sample of young men five years beyond high school, is a qualified no. One of the reasons for qualifying the answer is that our analyses involve individuals only five years beyond high school; also those still pursuing their education (or just out of college) were not included in this particular set of analyses. Thus it is quite possible that those with the highest levels of educational attainment will, perhaps ten years after high school, show substantially above-average levels of job satisfaction. By that time, or perhaps still later, others with somewhat above-average levels of education may also show more than average job satisfaction. But for young men in their early twenties it seems fairly clear that the benefits of education do not include the assurance of above-average job satisfaction.

The implications of this finding are neither simple nor straightforward. For many members of the high school class of 1969 the promised effects of education had not--at least not yet--been realized. It seems increasingly clear that the message to future classes should be a more balanced one: the high status jobs are likely to go to the college graduates; but the total number of such jobs seems limited, and this means that some graduates are likely to be disappointed with the jobs they actually find. Higher education still carries a promise; but it also contributes to high aspirations, and thus carries a risk. In the short run, at least, the two effects seem more nearly equal than might have been expected.

Analyses of Change, Stability, and the Impacts of Education

It is widely believed that social environments and experiences can have lasting impacts on personality characteristics and behavior patterns; and, as we noted at the outset of this chapter, one of the most basic purposes of the Youth in Transition project has been to measure and analyze

some of these impacts. But it is also commonly assumed that variations in personality and behaviors lead different people to be exposed to different environments and experiences; and the preceding section confirms that differences in background, ability, and prior experiences greatly influence educational (and also occupational) attainment. Thus the causal linkage between environments and personality must be viewed as a two-way street: environments and events shape people, but people also play an important part in selecting and shaping their own experiences.

In this section we examine a wide range of personality and behavior dimensions which have been measured at several points in time, in many cases over the full eight-year span of the Youth in Transition project. The previous section summarized analyses in which such characteristics were viewed as predictors of educational and occupational attainment; now we consider these dimensions of values, attitudes, self-concepts, and behaviors as outcomes--as dependent variables or criterion dimensions. In each case we ask whether different levels of educational attainment (as well as other experiences during the late teens and early twenties) are linked to differences in the outcome dimensions--and whether such differences are stable or represent differential patterns of change across time.

Self-Esteem and Related Measures

Self-Esteem

Self-esteem, an individual's self-evaluation or judgment of his own worth, has long been the subject of theoretical speculation; and it has also recently been the topic of several major empirical studies (Rosenberg, 1965; Coopersmith, 1962; Rosenberg & Simmons, 1971).

A number of authors have presented or summarized evidence that self-esteem is linked to educational (and also occupational) attainment (Rosenberg, 1965; Coopersmith, 1967; Rosenberg & Simmons, 1971; Gergen, 1971; Purkey, 1970; Luck & Heiss, 1972; Weidman, Phelan & Sullivan, 1972). The problem, of course, is distinguishing cause from effect. We can conceptualize three logically distinct patterns of causation, all of which could play some part in a relationship between self-esteem and educational attainment: self-esteem could contribute to educational attainment; educational success could contribute to heightened self-esteem; and both could be determined (at least in part) by other prior factors such as background, ability, and earlier educational successes.

Our longitudinal analysis of self-esteem suggests that it is not a characteristic of personality that is firmly fixed by the time a young man enters high school. On the contrary, we found a gradual, but quite substantial rise. The pattern of findings does not fit any notion that coming to the end of high school produces a sudden and substantial rise in self-regard. A more likely explanation is that the rise in self-esteem reflects a gradually increasing maturity and the corresponding increase in status, opportunities, and privileges.

Although we found a substantial amount of change in self-esteem over time, we also found a good deal of stability in scores from one year to another. We estimate that, among young men in their late teens and early twenties, self-esteem (after adjustments for measurement unreliability) has a stability of nearly .9 for one-year intervals. Over longer periods the stability is proportionately lower, so that for the total eight-year span of the study we estimate the stability of self-esteem to be .4. These findings concerning stability, coupled with the fact that the overall rise in scores was quite gradual, provide some support for the view of self-esteem as a relatively enduring characteristic rather than something which shifts abruptly from one situation to another. Change certainly does occur during and following late adolescence, but the change seems to be gradual and developmental rather than revolutionary.

We found self-esteem linked with educational attainment; however, the linkage was complex and showed evidence of change over time. The young men in our sample who eventually completed college and entered graduate training showed the highest mean self-esteem, those who never attained the high school diploma had the lowest mean self-esteem, and the groups between these two extremes had mean self-esteem scores which neatly matched their level of educational attainment. The finding that self-esteem was positively related to educational attainment is not surprising. But the relationship was not strongest and clearest at the end of the study, when individuals had sorted themselves into the several levels of educational attainment. Instead, the differences were strongest among self-esteem scores obtained at the beginning of the study, apparently "anticipating" educational attainment; and that is surprising.

Our findings permit us to conclude that educational success became a less vital part of the self--at least for young men who were in high school during the late 1960s. However, this was a period of important social change, including public views about the value of education. Thus our longitudinal data left some uncertainty as to whether the changing link between educational success and self-esteem reflected a genuine developmental process in late adolescence, or a secular trend which affected the society as a whole. (Incidentally, this illustrates a very general problem in drawing inferences from a single-cohort longitudinal study: maturational changes are completely confounded with secular trends.) Fortunately, it has been possible to carry out a partial replication and extension of this self-esteem analysis; the results showed that the correlations between self-esteem and educationally relevant measures for Youth in Transition participants as seniors in 1969 were matched almost identically by data from both male and female seniors in the high school class of 1977 (O'Malley & Bachman, 1979). Thus we feel confident, in this case, that the pattern is a fairly general one and the change represents a maturational process.

While level of educational attainment seems to have had little direct impact on later self-esteem, the same cannot be said for some aspects of occupational attainment. The status of the respondent's current job was correlated with self-esteem at the end of the study, and that

relationship remained largely intact after controlling for background, ability, educational attainment, and earlier levels of self-esteem. Another aspect of occupational success, employment versus unemployment, also showed an impact on self-esteem at the end of the study--a difference of about one-third of a standard deviation. Further examinations of the effects of unemployment on self-esteem indicated that dropouts were more likely than others to show a disadvantage in self-esteem (a smaller overall increase) as a result of being unemployed. Given the small number of cases involved, we must view these findings as only suggestive. But what they suggest is that the impact of unemployment may be felt most heavily by those who fail to attain a high school diploma. We may speculate that unemployment is particularly likely to lead to self-blame among these young men, since their lack of a diploma is something which they are often told that they could, and should, take steps to correct.

Motives, Affective States, and Values

The first four data collections of the Youth in Transition project included measures of needs for self-development and self-utilization, happiness, negative affective states, somatic symptoms, social values, and internal control (personal efficacy). Because of restrictions in questionnaire length, these dimensions were not measured in 1974. As a result, we were unable to examine ways in which these dimensions might have been affected by post-high school experiences. We did, however, examine their stability over four years, as well as their relationships with educational attainment. Similar to self-esteem, these dimensions showed a good deal of stability during the high school years, thus adding further support to the view that changes in personality dimensions during late adolescence are primarily gradual rather than abrupt. A number of the dimensions showed declining correlations with eventual educational attainment--a pattern similar to that found for self-esteem. In particular, need for self-development, need for self-utilization, social values, and internal control showed rather sharp declines in correlations. Measures of happiness, negative affective states, and somatic symptoms did not mimic the self-esteem pattern so closely.

Delinquent Behavior and Drug Use

It is interesting to consider some ways in which the undesirable or "negative" behaviors of delinquency and drug use are both similar to, and different from, the desirable or "positive" behaviors of educational and occupational attainment. All of these behaviors are considered vitally important for youth and for the nation in general. They are also similar in that they are relatively concrete events that can be reported fairly objectively, and thus they stand in sharp contrast to some of our subjective measures such as self-esteem or job satisfaction. A critical difference between the two sets of dimensions is that educational and occupational attainment measures are really not very meaningful until

at least several years after high school, whereas the behaviors of delinquency and drug use can be measured during high school and even earlier. Ironically, it appears that these socially disapproved behaviors provide our best opportunity to study changes in actual behaviors, and to see whether they can be linked to particular environments and experiences during the late teens and early twenties.

Delinquent Behavior

We focused on three indexes of delinquency: interpersonal aggression, theft and vandalism, and seriousness of delinquency. All three showed an appreciable drop in scores from 1966 to 1974. Along all three dimensions there is a negative correlation with eventual educational attainment. Multivariate analysis indicated that the different levels of delinquent behaviors, especially interpersonal aggression, shown by those with different levels of education, reflect deeply ingrained patterns of behavior that were fully in evidence as early as junior high school. In particular, it should be stressed that the higher delinquency rates among dropouts were most clearly apparent during junior and senior high school. These data fail to support the notion that dropping out causes delinquency, it seems more likely that delinquency is one of the causes of dropping out.

The correlations between delinquency and job status are smaller than those for educational attainment, but they show a similar pattern of consistency over time; accordingly, we conclude that these correlations appear simply because job status is a weak proxy for educational attainment.

Unemployment is one aspect of environment or experience which appeared to have some direct impact on our measure of interpersonal aggression. The effect showed up most clearly for unemployed dropouts--a finding which matches the relatively low self-esteem scores we observed for this group.

On the whole, we conclude that the delinquency differences linked to educational and occupational circumstances in 1974 were largely a reflection of long-standing patterns and were not heavily influenced by the specific post-high school experiences we examined.

Drug Use

The 1970 and 1974 questionnaires included a number of items about the use of illegal drugs, and the use of alcohol and cigarettes. From 1969 (measured retrospectively) to 1974, respondents reported considerable increases in usage rates for alcohol, marijuana, and other illegal drugs. Cigarette use showed little change during that period.

The largest drug use differences between those who went to college and those who did not was not in the use of alcohol or marijuana or any other illicit drug. The really big difference was in the daily use of cigarettes--the higher the level of education the lower the likelihood of being a smoker. This was not primarily an effect of post-high school education, however; the differences were present by the time most respondents were seniors in high school. The fact that smoking was much less frequent among young men who got good grades in junior and senior high school, and then went on to college, does not provide us with a clear explanation as to why their smoking rates were lower; but it does suggest one strong possibility. During the last decade there has been increasing evidence, widely reported in the media, pointing to the health hazards involved in smoking. It seems likely that those young men who are more successful in school are also more likely to read and take seriously such reports.

Alcohol use showed a negative correlation with educational attainment; however, the strength of that relationship declined from 1969 to 1974. The illegal substances showed very little relationship with education. The fact of simply going to college or not showed little relationship to the use of illegal drugs. But the major program of study elected in college did show a correlation with the use of illegal drugs. In particular, humanities and arts majors were more likely to be users of illegal drugs. The differences among academic majors were considerably less for tobacco and alcohol.

Job status and employment (versus unemployment) showed little relationship with alcohol and cigarette use. But the data indicated that failure to find a job may contribute to illegal drug use.

The findings with respect to marriage suggested that being married may "tone down" the excesses of young men in their early twenties. Single men tended to increase substantially their use of alcohol and illegal drugs from 1969 to 1974, while married men showed much less of an increase.

Assessing the Impacts of Environments and Experiences on Delinquent Behavior and Drug Use.

We can summarize and interpret our findings on delinquent behavior and drug use as indicating again that educational attainment, and the related dimension of occupational status, are correlates more than causes of differences among young men. Educational attainment is linked to delinquency, cigarette use, and alcohol consumption. But in each case the findings suggest that the differences in behavior preceded the sorting out by level of education. Thus the more accurate interpretation is that young men with high rates of delinquency, cigarette use, and alcohol use are less likely to reach high levels of educational attainment.

We do have some evidence of other factors having a causal impact on these behaviors. Specifically, it appears that unemployment may contribute to aggression and drug use, while marriage may tend to reduce such behaviors. Young men unemployed in 1974 reported fully twice as much interpersonal aggression and illicit use of amphetamines, barbiturates, and hallucinogens as did those who were employed. The earlier measures showed much smaller differences between those who were, and were not, unemployed by 1974. During junior and senior high school, those who married and became fathers by age 23 were well above average in interpersonal aggression; but five years later they were very similar to the rest of the sample in mean level of aggression. During high school there were few differences in drug use between those who soon married and those who did not, but by the fifth year after high school the proportions of single men using marijuana and/or other illicit drugs were half again as large as the proportions of married men doing so.

Does unemployment really heighten aggression and drug use? Does marriage actually tend to "reform" some overly aggressive individuals, and does it decrease the likelihood of drug use? Our findings in this area are suggestive, but not definitive. In each case an alternative path of causation is possible. Increases in aggression and drug use in the years following high school may contribute to unemployment and decrease the likelihood of marriage. But the fact that we are dealing with different patterns of change in behaviors, rather than stable differences in behaviors, leads us to favor an explanation in terms of environmental/experiential impact.

The Limitations of Single-Cohort Longitudinal Studies

It seems appropriate at the conclusion of this chapter to call attention to an important limitation in a study which follows just one cohort longitudinally. Our study followed young men from 1966 (when they averaged fifteen years of age) to 1974 (average age of twenty-three), and during that period we documented a number of important shifts: self-esteem rose substantially, trust in government plummeted, support for U.S. policy in Vietnam eroded, and drug use increased sharply. Only in the case of the shift in self-esteem can we be reasonably comfortable that our research has uncovered a "maturational change"--a shift attributable to increasing age and the accompanying experiences. In the case of drug use we attribute some of the change to increasing age, but some of it also must be attributed to "secular trends"--shifts in the society in general quite apart from changes in age. Views about government and about Vietnam changed substantially for the population as a whole during the period from 1966 to 1974; thus it would be a mistake to interpret our study as demonstrating that young people become more cynical about government during their late teens and early twenties.

The problem, of course, is that any longitudinal study of a single cohort confounds age with historical period. In our case, age fifteen corresponds to 1966 and age twenty-three corresponds to 1974; and sometimes we are unable to decide whether an observed change in our longitudinal panel is best interpreted in terms of changing age (maturation) or short-term historical shifts (secular trends). Much has been written recently about efforts to distinguish between these two types of change (Schaie & Baltes, 1975; Palmore, 1975; Adams, 1978; Riley & Warning, 1978), and perhaps the one area of consensus is that there are no totally satisfactory solutions. Nevertheless, it does seem clear that one important step is to provide data from more than a single cohort.

We are now doing exactly this in a new research venture in which we are collaborating with Lloyd Johnston. The new project, called Monitoring the Future, has sampled a new cohort of high school seniors each year since 1975 and will follow-up a portion of each sample for at least six years beyond graduation (Bachman & Johnston, 1978). The research concentrates heavily, but not exclusively, on drug use and related attitudes (Johnston, Bachman and O'Malley, 1979(a); Johnston and Bachman, 1975; Bachman, Johnston and O'Malley, 1979(a); Johnston, Bachman and O'Malley, 1979(b); Bachman, Johnston and O'Malley, 1979(b)). Our hope is that this multi-cohort longitudinal design will enable us to make much clearer distinctions than would otherwise be possible between maturational changes, secular trends, and differences between cohorts.

TABLE 1

Data Collections from Young Men

	TIME 1	TIME 2	TIME 3	TIME 4	TIME 5
Date	Fall, 1966 (tenth grade)	Spring, 1968 (eleventh grade)	Spring, 1969 (twelfth grade)	Spring, 1970 (grade 12 + 1 yr)	Spring, 1974 (grade 12 + 5 yrs)
Procedure	Individual inter- views; group-ad- ministered tests and question- naires	Individual in- terviews and questionnaires; \$2 payment	Group adminis- tered question- naires; \$5 pay- ment	Individual inter- views and ques- tionnaires; \$10 payment	Mail question- naires; \$10 pay- ment
Location	Schools	"Neutral Site"	"Neutral Site"	"Neutral Site"	Respondent's home
Number of Respondents*	2213	1886	1799	1620	1628
% of Original Sample (N = 2277)	97.2%	82.8%	79.0%	71.1%	71.5%
% of Time 1 Panel (N = 2213)	100%	85.2%	81.3%	73.2%	73.5%

* Probability sample located in 87 schools.

TABLE 2

Measures Included in Each Data Collection (partial listing)

	1966 1	1968 2	1969 3	1970 4	1974 5
<u>Educational Attainment</u>					
Number Years Schooling				X	X
Education Attained				X	X
<u>Occupational Attainment</u>					
Job Status - Duncan				X	X
Employment Status				X	X
Hourly Pay Rate				X	X
Job Satisfaction				X	X
Job Characteristics (Actual)				X	X
<u>College</u>					
College Status Ranking					X
College Mean ACT Score					X
<u>Environments/Experiences</u>					
Military Service				X	X
Marital/Parental Status				X	X
Urbanicity	X	X	X		X
<u>Family Background</u>					
Socioeconomic Level	X				
Father's Occupation	X				
Father's Education	X				
Mother's Education	X				
Number of Possessions in Home	X				
Number of Books in Home	X				
Number of Rooms/Person	X				
Number of Siblings	X				
Intactness of Home	X				
Parental Punitiveness	X				
Race	X				
Region	X				
Urbanicity (Origin)	X				
<u>Ability</u>					
Composite Ability	X				
Quick Test	X				
GATB-J	X				
Gates Reading Test	X				
<u>Educational Behaviors</u>					
Repeated Grade	X				
Average Grades (High School)	X	X	X		
Number of Hours Homework	X	X	X		
Rebellious Behavior in School	X	X			
Curriculum	X	X	X		

	1966 1	1968 2	1969 3	1970 4	1974 5
<u>Educational Plans</u>					
College Plans	X	X	X		
<u>Educational Attitudes</u>					
Interest in Courses	X	X	X		
Positive School Attitudes	X	X	X		
Negative School Attitudes	X	X	X		
Academic Achievement Value	X	X	X	X	
<u>Self Concepts</u>					
School Ability	X	X			
Does Best Work in School	X	X			
Works Harder than Average	X	X			
Satisfaction with School Work	X	X			
Self-Esteem	X	X	X	X	X
<u>Motives</u>					
Need-Social Approval	X				
Test Anxiety	X	X			
Need-Self-Development	X	X	X	X	
Need-Self-Utilization	X	X	X	X	
<u>Affective States</u>					
Happiness	X	X	X	X	
Negative Affective	X	X	X	X	
Somatic Symptoms	X	X	X	X	
Impulse to Aggression	X	X	X	X	
<u>Social Values & Attitudes</u>					
Social Values Cluster	X	X	X	X	
Internal Control	X	X	X	X	
Trust in People	X	X	X	X	
Trust in Government	X	X	X	X	X
Interest in Government	X	X	X	X	X
Racial: Strong Government			X	X	X
Racial: Social Distance			X	X	X
Racial: Perceived Discrimination			X	X	X
Vietnam Dissent			X	X	X
Military Influence			X		X
Abortion				X	X
Population Concern				X	X
Desired Family Size				X	X
<u>Occupational Attitudes & Aspirations</u>					
Job that Pays Off	X	X	X	X	X
Job that Doesn't Bug	X	X	X	X	X
Ambitious Job Attitudes	X	X	X	X	X
Status of Aspired Occupation	X	X	X	X	X
Job Characteristics (Ideal)	X	X	X	X	X

	1966 1	1968 2	1969 3	1970 4	1974 5
<u>Delinquent Behaviors</u>					
Delinquent Behavior in School	X	X	X		
Seriousness of Delinquency	X	X	X	X	X
Interpersonal Aggression	X	X	X	X	X
Theft and Vandalism	X	X	X	X	X
<u>Drug Use</u>					
Cigarettes			X*	X	X
Alcohol			X*	X	X
Marijuana			X*	X	X
Amphetamines			X*	X	X
Barbiturates			X*	X	X
LSD			X*	X	X
<u>High School-Aggregate</u>					
Mean Socioeconomic Level	X				
Mean Quick Test	X				
Mean GATB-J	X				
Mean Gates Reading Test	X				

*Measured retrospectively in 1970.

Note. An "X" indicates the particular dimension was measured at the indicated time point.

References

- Adams, J., Sequential Strategies and the Separation of Age, Cohort, and Time-of-Measurement Contributions to Developmental Data. Psychological Bulletin, 1978, 85, 1309.
- Alexander, K. L., Eckland, B. K. and Griffin, L. J. The Wisconsin Model of Socioeconomic Achievement: A Replication. American Journal of Sociology, 1975, 8, 324.
- Bachman, J. G. Youth in Transition: Volume II, The Impact of Family Background and Intelligence on Tenth-Grade Boys, Ann Arbor: Institute for Social Research, 1970.
- Blau, P. M. and Duncan, O. D., The American Occupational Structure. New York: Wiley, 1967.
- Bachman, J. G., Green, S. and Wirtanen, I. D. Youth in Transition: Volume III, Dropping Out--Problem or Symptom? Ann Arbor: Institute for Social Research, 1971.
- Bachman, J. G. and Jennings, M. K. The Impact of Vietnam on Trust in the Government. Journal of Social Issues, 1975, 31, 141.
- Bachman, J. G. and Johnston, L. D. The Monitoring the Future Project Design and Procedures. (Monitoring the Future Occasional Paper Series #1). Ann Arbor: Institute for Social Research, 1978.
- Bachman, J. G. and Johnston, L. D. A Profile of Today's Colleg Freshmen. Psychology Today, 1979, 13.
- Bachman, J. G., Johnston, L. D. and O'Malley, P. M. Monitoring the Future: Questionnaire Responses from the Nation's High School Seniors, 1976. Ann Arbor: Institute for Social Research, 1979(a).
- Bachman, J. G., Johnston, L. D. and O'Malley, P. M. Monitoring the Future: Questionnaire Responses from the Nation's High School Seniors, 1978. Ann Arbor: Institute for Social Research, 1979(b).
- Bachman, J. G., Kahn, R. L., Mednick, M. T., Davidson, T. N., and Johnston, L. D. Youth in Transition: Volume I, Blueprint for a Longitudinal Study of Adolescent Boys. Ann Arbor: Institute for Social Research, 1967.
- Bachman, J. G. and O'Malley, P. M. Self-Esteem in Young Men: A Longitudinal Analysis of the Impact of Educational and Occupational Attainment. Journal of Personality and Social Psychology, 1977, 35, 365.
- Bachman, J. G. and O'Malley, P. M. The Search for "School Effects"--Some New Findings and New Perspectives. Unpublished manuscript.
- Bachman, J. G., O'Malley, P. M. and Johnston, J. Youth in Transition: Volume VI, Adolescence to Adulthood--Change and Stability in the Lives of Young Men. Ann Arbor: Institute for Social Research, 1978.
- Coopersmith, S. The Antecedents of Self-Esteem. San Francisco: W. H. Freeman and Company, 1967.
- Davidson, T. N. Youth in Transition: Volume IV, Evaluation of a Strategy For Longitudinal Analysis of Survey Panel Data. Ann Arbor: Institute for Social Research, 1972.
- Duncan, O. D. in Occupations and Social Status. A. J. Reiss, Jr., ed. New York: The Free Press, 1961.
- Gergen, K. J. The Concept of Self. New York: Holt, Rinehard, & Winston, 1971.
- Johnston, L. D. "The American High School, Its Social System and Effects" Ann Arbor: Institute for Social Research, 1973(a).
- Johnston, L. D. "Drugs and American Youth." Ann Arbor: Institute for Social Research, 1973(b).
- Johnston, J. and Bachman, J. G. Youth in Transition: Volume V, Young Men and Military Service. Ann Arbor: Institute for Social Research, 1972.
- Johnston, J. and Bachman, J. G. The Transition from High School to Work: The Work Attitudes and Early Occupational Experiences of Young Men. Ann Arbor: Institute for Social Research, 1973.
- Johnston, L. D. and Bachman, J. G. Monitoring the Future: Questionnaire Responses from the Nation's High School Seniors, 1975. Ann Arbor: Institute for Social Research, 1979.
- Johnston, L. D., Bachman, J. G. and O'Malley, P. M. Drugs and the Class of 1978: Behaviors, Attitudes and Recent National Trends. (National Institute on Drug Abuse). Washington, D.: U.S. Government Printing Office, 1979(a). A shorter version, Highlights, is also available.
- Johnston, L. D., Bachman, J. G. and O'Malley, P. M. Monitoring the Future: Questionnaire Responses from the Nation's High School Seniors, 1977. Ann Arbor: Institute for Social Research, 1979(b).
- Johnston, L. D., O'Malley, P. M. and Eveland, L. K. in Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues. D. G. Kandel, ed. Washington, D.C.: Hemisphere Press, 1978.
- Luck, P. W. and Heiss, J. Social Determinants of Self-Esteem in Adult Males. Sociology and Social Research, 1972, 57, 69.
- O'Malley, P. M. and Bachman, J. G. Self-Esteem and Education: Sex and Cohort Comparisons Among High School Seniors. Journal of Personality and Social Psychology, 1979, 37, 7.
- Palmore, E. When Can Age, Period, and Cohort Be Separated? Social Forces, 1975, 57, 384.

- Purkey, W. W. Self Concept and School Achievement. New Jersey: Prentice-Hall, 1970.
- Rosenberg, M. Society and the Adolescent Self-Image. Princeton: University Press, 1965.
- Rosenbert, M. and Simmons, R. G. Black and White Self-Esteem: The Urban School Child. (Rose Monograph Series). Washington, D.C.: American Sociological Association, 1971.
- Riley, M. W. and Waring, J. Age, Cohorts, and Drug Use. in Longitudinal Research on Drug Use: Empirical Findings and Methodological Issues. D. G. Kandel, ed. Washington, D.C.: Hemisphere Press, 1978.
- Schaie, K. W. and Baltes, P. On Sequential Strategies in Developmental Research. Human Development, 1975, 18, 384.
- Sewell, W. H. and Hauser, R. M. Education, Occupation, and Earnings: Achievement in the Early Career. New York: Academic Press, 1975.
- Sewell, W. H. and Hauser, R. M. in W. H. Sewell, R. M. Hauser, and D. L. Featherman, eds., Schooling and Achievement in American Society. New York: Academic Press, 1976.
- Wilson, K. L. and Portes, A. The Educational Attainment Process: Results from a National Sample. American Journal of Sociology, 1975, 81, 343.
- Weidman, J. C., Phelan, W. T. and Sullivan, M. A. The Influence of Educational Attainment of Self-Evaluations of Competence. Sociology of Education, 1972, 45, 303.

An Epidemiological Study of the
Behavior Problems of Young Children*

Helen Z. Reinherz
Patrick C. Fowler
Mitchell S. Farkas
Corazon M. Raymundo
Myra O'Regan
Robert M. Park
Simmons College

*This research was supported by NIMH Grant No. 2 RO1 MH27458, Helen Z. Reinherz, Sc.D., principal investigator, Simmons College School of Social Work, Boston, Massachusetts. The authors acknowledge with thanks the opportunity to carry out this research with the collaboration of the Quincy Public Schools and its superintendent, Dr. Lawrence Creedon. Special thanks are due to the members of the Pupil Personnel Department including Dr. Carol Lee Friggin, Dr. Luleen Anderson, and Mrs. Barbara Hunt. Also, the contribution of the kindergarten teachers and the children and parents in the study is gratefully acknowledged. The assistance of Ms. Faye Lavrakas, Ms. Christine Johnson and Ms. Patricia Jameson has facilitated the completion of this report.

Patrick Fowler is now at the University of Virginia Medical Center, Charlottesville, Virginia; Robert M. Park is a graduate student at Harvard School of Public Health, Boston, Massachusetts.

An Epidemiological Study of the
Behavior Problems of Young Children

Helen Z. Reinherz

Patrick C. Fowler

Mitchell S. Farkas

Corazon M. Raymundo

Myra O'Regan

Robert M. Park

Simmons College

This report describes an ongoing, longitudinal study initiated when a cohort of children registered for kindergarten and participated in a comprehensive screening procedure. The goal of this research is (1) to describe the behavior and emotional status of a large group of children at several points in their development and (2) to examine the relationship between multiple assessments of children before entrance to kindergarten and their subsequent behavioral and emotional functioning.

The study began in the Spring of 1977 when extensive screening information was collected on all 774 preschool children who were scheduled to enter kindergarten in the Fall in Quincy, Massachusetts. One year later, the functioning of children in the same cohort was assessed with instruments administered to their parents and teachers. A third major period of data collection is scheduled for the Spring of 1981 when most of the children will be completing the third grade. At that time, self-reports of the children and descriptions of their environment will supplement evaluations of their functioning by

teachers, parents and other sources. The present report includes a description of the instruments used in the study, the plan for data collection and some methodological considerations. Finally, some selected analyses and preliminary implications of the findings from the preschool screening and kindergarten year are discussed.

The project's interest has been in the identification of children whose behavior, health and other areas of functioning deviated from expectations for their age group. This approach is based on the belief that children who deviate from developmental norms in any area are likely to be at greater risk of developing mental health problems later in their development than those who are not. Also, it may be important to focus on children who are at risk in a variety of domains in early identification and intervention efforts (Reinherz, 1978). This study represents an effort to determine if it is feasible to identify such children in order to implement programs which rely on these concepts of risk.

In the present research, the opportunity to collect salient information on children's development in a variety of areas arose from the movement toward early identification of children with "special needs." A number of states have mandated screening programs in the past several years.

When the Massachusetts screening program was initiated there were no screening procedures considered satisfactory by most of the school districts (Nuttall & Gomes, 1975). Thus, the need and the opportunity existed to design, develop and administer a comprehensive screening battery to assess the major areas of a child's functioning. An established long-term collaborative relationship with a large school system interested in improving its assessment procedures allowed the investigators to accomplish their goal of collecting data on a sizeable "normal" population of children who could be followed longitudinally. At the same time, the selection of data collection tools and methodology was subjected to the constraints of on-going resources of the school system (space, time, personnel) as well as other institutional considerations. For example, the resulting battery for pre-kindergarten data collection had to be brief (no more than 20-30 minutes per child). Also, it had to require minimal training and time involvement of school personnel. A positive concomitant of these constraints is that the resulting procedures are relevant to a variety of natural settings, especially schools.

The methodology employed in the study reflected ethical concerns as well as practical constraints. In order to insure confidentiality, the behavioral information obtained for each child was not placed in

school records. In addition, all data collection waves subsequent to screening were conducted directly by project staff. This procedure was employed to protect against the possible stigmatization of subjects and the emergence of self-fulfilling prophecies.

Description of the Site and Population

The site of the study is Quincy, a largely industrial city in southeastern Massachusetts. The city has a population of approximately 90,000 and a median family income that is slightly below that of the greater Boston area. Figures quoted are from the 1970 census (U.S. Bureau of Census, 1972). Most of the population are Caucasian, largely of Nova Scotian, Irish and Italian backgrounds. The main occupational grouping is skilled manual labor (35% of the population) and approximately 65% of the population can be categorized as Class III (low middle) or IV (working class) on the Hollingshead Scale of Social Class Position (1957). The working class character of the children's families is reflected not only in the occupational status of the fathers, but also by the fact that almost 78% of the mothers listed themselves as "homemakers" at the time of the screening. However, some of the mothers may also have engaged in part-time or seasonal occupations which are not noted by them.

The modal family size of five (both adults and children) included 69% of the study group. Fourteen percent of the 774 children in the study were raised by single parents. In most of the single parent families, the father was absent. The incidence of single parent households in the study group was slightly less than that for the city as a whole and considerably lower than the greater Boston area, where 25% of the families were headed by a single parent. The lower incidence of single parents in the study site is in keeping with the over-all impression of a stable predominantly working class community.

The study cohort consisted of all children (average age of 4 years and 11 months) registering for kindergarten and participating in the state mandated screening process. At the time of screening there was an equal proportion of boys and girls. According to school data, 13 children in the age cohort had already been assessed as having severely handicapping conditions and were removed from their age-mates by placement at day or residential schools. Seven of these children had severe emotional disorders; two were physically handicapped; and one each had diagnoses of deafness, cerebral palsy, brain damage and profound retardation. Thus, the cohort consisted of children who had not been previously identified as severely handicapped, and therefore, were candidates for public school entry.

Data Collection Plan

The first major wave of data collection consisted of a preschool screening battery which was administered in the Spring of 1977 (see Figure 1 for an outline of the overall data collection plan). The instruments included: a norm-referenced screening test of cognitive functioning (Hainsworth & Hainsworth, 1974), tests of visual and auditory acuity given by trained school health personnel who observed coordination and motor activity of the children, and a self-administered questionnaire completed by parents. The parental questionnaire was developed after pilot testing with over 800 subjects in the study community in the previous year. It solicited information on the child's past and present health and development, as well as a description of current behavior which was assessed through a newly developed Behavior Checklist (Reinherz, Kelfer, Griffin & Holloway, 1977).

Insert Figure 1 about here

In the Spring of 1978, kindergarten measures were obtained for the original 1977 cohort. During this wave of data collection the child's behavior and performance at home and school were the major areas of inquiry. The design allowed for the gathering of data from three sources: parents, through a self-administered questionnaire containing a short form of the Behavior Checklist; teachers, through a self-administered questionnaire; and school records, recording referrals and consultation concerning behavioral and other problems. The teachers' questionnaire also contained a previously developed age-appropriate behavior scale reported to have adequate reliability and validity (Behar & Stringfield, 1974).

The annual collection of data from school records will be continued as the cohort moves through the elementary school. A third major period of data collection is planned for the Spring of 1981 when most of the cohort will be completing the third grade. At that time, self-reports from the children themselves will supplement evaluations of their social, behavioral and cognitive performance by teachers, parents and other sources. The combination of parental reports and professional assessments at each major wave of data collection reflects the need to obtain complementary views of children in cross-situational contexts (Chamberlin, 1976; Mitchell & Sheperd, 1966).

Screening Information

Parent Questionnaire

Developmental Data

Parents provided information on the achievement of milestones by their children from birth to age 5. The children's functioning was evaluated in three major areas: motor, speech and language and self-help skills. Much of the data was retrospective. However, a number of questions about self-help skills were designed to reflect current functioning of the children. The overall distributions of results were similar to norms reported in the work of Knobloch and Pasamanick (1974). As reflected in reviews of other studies (Maccoby & Jacklin, 1974), boys had a generally slower rate of early development than girls.

Health Data

Several indices were constructed to summarize health data for key periods of development. The Prenatal Index summarized the mother's condition during pregnancy and at delivery. The index included: occurrences of bleeding during pregnancy, toxemia, difficult delivery, complications at delivery and other labor or delivery problems. The Neonatal Index summarized the status of the child on serious disorders occurring during the first month of life. These problems were: jaundice, anoxia, infection, breathing difficulties and congenital defects. The Health Index was a composite of the number of occurrences of five illnesses during the first five years of life. These included: ear infection, seizure, surgery, fever and poisoning. Overall, the data indicated that the health status of the Quincy group was somewhat superior to national health norms for similar age groups (U.S. Dept. of HEW, 1974). For the majority of items in the Neonatal and the Health Indices, fewer problems were reported for girls than boys. However, for most of the individual items the sex differences did not reach statistical significance.

Behavior Checklist

Mothers rated their children's behavior at screening on the 38-item Behavior Checklist which was developed by the project staff. The checklist items were designed to assess behaviors in problem areas germane for preschool children including: aggression/hostility, shyness/anxiety, attention-span/distractability, developmental delay and speech/language. Each item described a specific behavioral characteristic of four and five year old children, reportable by their mothers. Ratings were on a five-point scale consisting of the options "always," "often," "sometimes," "rarely" and "never."

One definition of the prevalence of significant problem behaviors for individual items of the Behavior Checklist was determined by parental ratings of either "always" or "often." Prevalence using these criteria ranged from a low of less than 1% for the reporting of bowel accidents to 14.6% for the reporting of impatient and demanding behavior. The five items with the highest prevalence ranged from 9.4% to 14.6%. They represented either hostility/aggression or attention-span/distractability behaviors designed as "externalizing" by Achenbach and Edelbrock (in press). However, in order to compare the prevalence findings of this study with those of some other investigators using broader ranges of frequencies to determine existence of problem behavior, prevalence was also estimated by adding responses in the "sometimes" category to "always" and "often."

Sensory Functioning and Observation by the School Nurse

Sensory functioning was evaluated through tests given by the school nurses. These included standardized measurements of visual and auditory acuity. More than 85% of the sample successfully completed both tests of acuity. In addition to the sensory tests, the nurses rated the children for coordination and motor activity. Twelve percent of the children were rated in the range from slightly awkward to very clumsy, and 11% were rated in the range from somewhat to very "fidgety."

Cognitive Functioning

Information regarding the children's cognitive development and school readiness was obtained from results of the Preschool Screening System (Hainsworth & Hainsworth, 1974). This screening instrument was administered in the Fall of 1977, prior to the children's entry into kindergarten. Both the Information Processing Subscale (47 items) and Verbal Subscale (10 items) of the instrument were administered. The information processing tasks require the child to imitate physical movements, follow simple directions, copy and label shapes, count and repeat phrases. The verbal component consists of a series of sentence completion items. In relation to the information published for the norm group, both girls and boys in the present cohort showed a greater concentration in the upper ranges for both subscales of this instrument. Although reasons for this relatively high level of performance are speculative, it could reflect the fact that over 60% of these children had some preschool experience, with an average duration of 6.7 months.

The Children at the End of the Kindergarten Year

Of the 774 children screened in 1977, 702 or 92% were enrolled in regular kindergarten classes at the end of the school year and, thus, rated by their teachers. In addition, information was available on

another 11 children who had been placed in special needs or pre-kindergarten classes. (Fifteen children were removed from the cohort for special programs during the kindergarten year.) Information on 79% (553) of the children in kindergarten was also received from parents who filled out a self-administered questionnaire at the end of the school year at school conferences.

Teacher Kindergarten Questionnaire

For each child the teachers filled out questionnaires that included general questions on the child's behavioral adjustment, academic performance and learning skills. Teachers also noted actions which were relevant to the child's academic and social functioning including consultation with guidance personnel and referral for diagnostic procedures. At the end of kindergarten, approximately 5.5% of the children were given comprehensive, in-depth evaluations, designed to assess special needs. These evaluations, given by the school, were mandated by state law.

A second component of the teacher questionnaire included a previously developed rating instrument designed to assess the existence of problem behaviors in young children (Behar & Stringfield, 1974). This scale contained many items representing "externalizing" (aggressive and distractible behaviors) and "internalizing" (shy/anxious behaviors) problems. On a three-point scale ("doesn't apply," "sometimes applies," "certainly applies") the percentage of children for whom the description that a behavior problem "certainly applies" ranged from less than 1% for "doesn't share toys" and "wet and soiled self" to about 7% for "restless," "squirmy" and "poor concentration."

Parent Kindergarten Questionnaire

The parent instrument contained several components in common with the teacher kindergarten questionnaire. Parents were asked a series of questions about: 1) areas in which their children were experiencing problems (e.g., home, school); 2) significant events in their family during the past year (e.g., illness, births, death); and 3) attitudes of their children towards school. Of interest was the finding that over 20% of the parents indicated their children had "some" behavior problems at home, but less than 10% indicated that such problems existed at school.

Parents' evaluations of their children's behavior were elicited with the Parent Followup Checklist, a shortened version of the Behavior Checklist used at screening. The Parent Followup Checklist consisted of 13 items selected from the Behavior Checklist to be representative of hostility, aggression, hyperactivity (externalizing behavior) and

shyness/anxiety (internalizing behavior). The percentage of responses which indicated the prevalence of a major problem behavior (defined as a parent rating of "often" or "always") ranged from less than 1% ("throws and breaks things" and "has behavior problems") to over 16% ("cannot sit still").

Relationships Between Parent and Teacher Ratings

A number of the questions asked of both parents and teachers in kindergarten ratings were highly comparable so that it was possible to examine the consistency of evaluation from the two important perspectives of these "natural ratings" of children (Kellam, Branch, Agrawal & Ensminger, 1975). For over 70% of the children both parents and teachers agreed in their positive evaluations of behavior and readiness for first grade entry. In almost every case in which inconsistencies did occur, parents were more positive in their evaluations than teachers. In the following sections some of the important differences, particularly in assessment of the behavioral functioning of the two sexes, will be discussed.

Explorations of Sex Differences: Some Initial Results

The earlier sections of this report have generally reported analyses of the total group. However, some interesting preliminary findings emerged when sex differences were considered. Several of these in the areas of health and development have been described briefly. These results are, as reported, comparable to those of other studies existing in the literature for children of this age. In behavioral functioning, although some findings are comparable to other studies, predominance of males with problem externalizing behaviors is not of the magnitude reported in other epidemiological studies (Beiser, 1972; Graham, 1979; Rutter, et al., 1970).

Contrast in Parental Ratings of Behavior at Home and School

In a section of the Parent Kindergarten Questionnaire, parents reported children's behavior at home and also indicated their view of their child's behavior in school. Parents indicated that almost equal proportion of the sexes (approximately 20%) had moderate to significant behavior problems at home. However, when reporting on incidence of behavior problems at school they stated that 13% of the boys and 6.3% of the girls had moderate to significant behavior problems at school. These assessments do agree with teacher ratings at school where sizeable sex differences and predominance of male behavior problems were revealed. Thus, although there was some disagreement on general ratings of the child's behavior, when parents and teachers rated in the context of behavior in the school environment, there was substantial agreement.

The indication of more than double the percentage of problematic behavior for boys as compared to girls at school highlights the greater difficulty boys have traditionally shown in the school environment (Beiser, 1972; Chamberlin, 1976; Rubin & Balow, 1978). Parents and teachers were essentially in agreement in rating the behavior of boys as more problematic than girls, within the school context.

Preliminary Findings on the Relationship Between Status at Screening and Kindergarten

In keeping with the objectives of the study, relationships between the children's status at screening and follow up were explored. Bivariate predictors representing each of the domains in the low to moderate range using occurred in the relationship between behaviors at screening and those at kindergarten as rated by parents and in the relationship between the cognitive screening measure and school readiness as rated by teachers.

Additional Substudies

Several substudies which are in various process of final analysis and reporting were initiated by the research staff. In the summer after the children were screened a subsample of 85 were interviewed by clinical psychologists and psychiatrists as a means of comparison of clinical judgment with screening assessments. The results of this study offer perspectives on the specific components of clinical judgment in assessing a child as at risk for behavior problems. In addition, the project utilized the Behavior Checklist in a setting contrasting children identified as having problems and special needs and normal children in the same classrooms. The Behavior Checklist results clearly differentiated the two groups (Kelfer, 1978). A third area of investigation was the study of family environments of children identified as behaviorally problematic and non-problematic (Fowler, 1979). The latter study served to pilot instruments considered for use in 1981 data collection.

The present study has only begun to follow this cohort as they leave home and enter the world of school with its expectations of performance and behavior. Assessments by parents and teachers of the children's functioning at the third grade (8 and 9 year old) level supplemented by the children's self reports and peer ratings will allow for a broad comprehensive description of these school aged children. In addition, analysis of the relationships between screening and various outcomes will be continued in order to further assess the fit between prediction and outcome and to identify those aspects of behavior and performance that are transient as contrasted with those that appear to be characteristic of behavioral disorder requiring intervention.

References

- Achenbach, T.M. & Edelbrock, C.S. Behavioral problems and competencies reported by parents of normal and disturbed children aged 4 through 16. In press.
- Anthony, E.J. & Koupernik, C. (Eds.). The child in his family: Children at psychiatric risk (Vol. 3). New York: John Wiley & Sons, 1974.
- Behar, L. & Stringfield, S. A behavior rating scale for the preschool child. Developmental Psychology, 1974, 10(5), 601-610.
- Beiser, M. The etiology of mental disorder: Sociocultural aspects. In B.B. Wolman (Ed.), Manual of Child Psychopathology. New York: McGraw-Hill, 1972.
- Bishop, Y., Feinberg, S. & Holland, P. Discrete multivariate analysis. Cambridge: MIT Press, 1975.
- Bureau for Education of the Handicapped. Better education for the handicapped: Annual report. Washington, D.C.: U.S. Government Printing Office, 1970.
- Chamberlin, R.W. The use of teacher checklists to identify children at risk for later behavioral and emotional problems. American Journal of Diseases of Children, 1976, 130, 141-145.
- Chess, S. The plasticity of human development: Alternative pathways. Journal of the American Academy of Child Psychiatry, 1978, 17, 81-91.
- Commonwealth of Massachusetts. Regulations for the implementation of Chapter 766 of the acts of 1972: The comprehensive special education law. Boston: Department of Education, May 1974.
- Escalona, S.K. Intervention programs for children at psychiatric risk: The contribution of child psychiatry and developmental theory. In E.J. Anthony & C. Koupernik (Eds.), The child in his family: Children at psychiatric risk (Vol. 3). New York: John Wiley & Sons, 1974.
- Feinberg, S.E. The analysis of cross-classified categorical data. Cambridge: MIT Press, 1977.
- Fowler, P.C. (Ed.) Relationships between early behavior, family environmental structure, and current developmental status. Unpublished report, Simmons College School of Social Work, 1979.
- Graham, P. Epidemiological studies. In H.C. Quay & J.S. Werry (Eds.), Psychopathological disorders of childhood (2nd ed.). New York: John Wiley & Sons, 1979.
- Hainsworth, P.K. & Hainsworth, M.L. Preschool screening system: Start of a longitudinal preventive approach (Field trial edition). Pawtucket, R.I.: Early Recognition Intervention Systems, 1974.
- Hobbs, N. The futures of children. San Francisco: Jossey-Bass, 1975.
- Hollingshead, A.B. Two-factor index of social position. New Haven: Yale University Department of Sociology, 1957.
- Johnson, S.C. Hierarchical clustering schemes. Psychometrika, 1967, 32(3), 241-254.
- Joint Commission on the Mental Health of Children. Crisis in mental health: Challenge for the 1970's. New York: Harper & Row, 1970.
- Kelfer, D. (Ed.) Development of the preschool screening instrument: Validation of the behavior checklist, design and assessment of the teacher rating scale. Unpublished report, Simmons College School of Social Work, 1978.
- Kellam, S.G., Branch, J.D., Agrawal, K.C. & Ensminger, M.D. Mental health and going to school: The Woodlawn program of assessment, early intervention, and evaluation. Chicago: University of Chicago Press, 1975.
- Kohn, M.L. Class and conformity: A study in values. Homewood, Ill.: Dorsey Press, 1969.
- Kohn, M. & Rosman, B. A two-factor model of emotional disturbance in the young child: Validity and screening efficiency. Journal of Child Psychology and Psychiatry. 1973, 14, 31-56.
- Knobloch, H. & Pasamanick, B. (Eds.). Gesell and Amartrada's developmental diagnosis (3rd Ed.). New York: Harper & Row, 1974.
- Maccoby, E.E. & Jacklin, C.N. The psychology of sex differences. Stanford, Ca.: Stanford University Press, 1974.
- Mitchell, S. & Sheperd, M. A comparative study of children's behavior at home and at school. British Journal of Educational Psychology, 1966, 36, 248-254.

Nuttall, E.V. & Gomes, L.A. Early identification of children with special needs: A survey of current screening practices in Massachusetts Public Schools. Chestnut Hill, Ma.: Boston College, 1975.

President's Commission on Mental Health. Report to the president (Vol. 1). Washington, D.C.: U.S. Government Printing Office, 1978.

Reinherz, H. Primary prevention of emotional disorders of children: Mirage or reality? Paper presented at the annual meeting of the American Association of Psychiatric Services for Children. Atlanta, Ga., November 1978.

Reinherz, H., Kelfer, D., Griffin, C.L. & Holloway, S. Developing a tool for assessing the social-emotional functioning of preschool children. Paper presented at the annual meeting of the American Association of Psychiatric Services for Children. Washington, D.C., November 1977.

Rubin, R.A. & Balow, B. Prevalence of teacher identified behavior problems: A longitudinal study. Exceptional Children, 1978, 45(2), 102-111.

Rutter, M., Tizard, J. & Whitmore, K. (Eds.). Education, health and behavior: Psychological and medical study of childhood development. New York: Wiley, 1970.

Sameroff, A.J. & Chandler, M.D. Reproductive risk and the continuum of caretaking casualty. In F.D. Horowitz, M. Hetherington, S. Scarr-Salapatek & G. Siegel (Eds.), Review of child development research (Vol. 4). Chicago: University of Chicago Press, 1974.

Sapir, S.G. & Wilson, B. A developmental scale to assist in the prevention of learning disability. Education and Psychological Measurement, 1967, 27(4), 1061-1068.

Schwab, J.J. & Schwab, M.E. Sociocultural roots of mental illness: An epidemiological survey. New York: Plenum Medical Book Co., 1978.

Stringer, L. School-entry interviews: How much priority should they have? Paper presented at the Midwest School Social Work Conference. St. Louis Mo., September 1973.

U.S. Bureau of the Census. 1970 census of population and housing census tracts. (Document No. PHC(1)-29). May, 1972.

U.S. Department of Health, Education and Welfare. Vital and health statistics. Series 11, No. 125, 1974.

Figure 1
Data Collection Plan

Time	Sources and Instruments	Type of Data
Pre-Kindergarten May, 1977	Parent Questionnaire	Health Demographic Developmental Behavior
	Hainsworth Preschool Screening System a/	Cognitive
	School Nurse's Testing	Sensory Functioning
Kindergarten May, 1978	Teacher Questionnaire	Behavior School Adjustment
	Parent Questionnaire	Behavior Family Events School and Personal Adjustment
	School Records	School Status
Grade 1 May, 1979	School Records	School Status
Grade 2 May, 1980	School Records	School Status
Grade 3 May, 1981	Teacher Questionnaire	Behavior School and Personal Adjustment
	Parent Questionnaire	Behavior School and Personal Adjustment
	Pupil Questionnaire	Peer Ratings Self-concept Locus of Control
	School Records	School Status
	Observations and Questionnaire	Family and Class- room Environment

a/ The Hainsworth Preschool Screening Instrument was administered in the Fall of 1977.

SELF-ATTITUDES AND DEVIANT RESPONSES:
TOWARD THE VALIDATION OF A GENERAL THEORY OF DEVIANT BEHAVIOR

Howard B. Kaplan
Professor of Sociology
Department of Psychiatry
Baylor College of Medicine
Texas Medical Center
Houston, Texas 77030

SELF-ATTITUDES AND DEVIANT RESPONSES:
TOWARD THE VALIDATION OF A GENERAL THEORY OF DEVIANT BEHAVIOR

Howard B. Kaplan
Professor of Sociology
Department of Psychiatry
Baylor College of Medicine
Texas Medical Center
Houston, Texas 77030

The study to be described below was designed to test components of a general theory of deviant behavior (Kaplan, 1975b).

Briefly, the theoretical model (Kaplan, 1975b) is based upon the postulate of the self-esteem motive. According to this theory, a person is said to behave so as to maximize the experience of positive, and to minimize the experience of negative, self-attitudes. Self-attitudes refer to the person's positive and negative emotional experiences upon perceiving and evaluating his or her own attributes and behavior.

Intense self-rejecting attitudes are said to be the end result of a history of membership group experiences in which the subject was unable to defend against, adapt to, or cope with circumstances having self-devaluing implications. Self-devaluing implications include disvalued attributes and behaviors, and negative evaluations of the subject by valued others. By virtue of the association between past membership group experiences and the development of intensely distressful negative self-attitudes the person loses motivation to conform to, and becomes motivated to deviate from, membership group patterns. These patterns are those specifically associated with the genesis of negative self-attitudes and, by a process of generalization, other aspects of the membership groups' normative structures. Simultaneously, the unfulfilled self-esteem motive prompts the subject to seek alternative (that is, deviant) response patterns which offer hope of reducing the experience of negative self-attitudes and increasing the experience of positive self-attitudes. Thus, the person is motivated to seek and adopt deviant response patterns because of a loss of motivation to conform to the normative structure which has an earlier association with the genesis of negative self-attitudes. Moreover, the deviant patterns represent the

only motivationally acceptable alternatives that might effectively serve self-enhancing functions. Which of several deviant patterns is adopted, then, would be a function of the person's history of experiences influencing the visibility and subjective evaluation of the self-enhancing/self-devaluing potential of the pattern(s) in question.

Adoption of the deviant response has self-enhancing consequences insofar as the deviant patterns facilitate intrapsychic or interpersonal avoidance of self-devaluing experiences associated with the predeviance membership group. It also serves to attack, symbolically or otherwise, the perceived basis of the person's self-rejecting attitudes (that is, representations of the normative group structure). Finally, it offers substitute patterns with self-enhancing potential for behavior patterns associated with the genesis of self-rejecting attitudes.

METHOD

The hypotheses comprising the general theory of deviant behavior under consideration were tested by data collected for the purpose in the course of a longitudinal survey research project.

Sample

The target sample included all of the seventh grade students in 18 of the 36 junior high schools in the Houston Independent School District as of March, 1971. The registered seventh grade students in the selected schools comprised 49.77 percent of the seventh grade students in all 36 schools. The selected schools appeared to be representative of the 36 schools along a number of parameters. The selected schools contained ten of the twenty schools with above 1500 total enrollment and eight of the 17 schools in which at least 25 percent of the total student enrollment were black.

The seventh grade was selected as an optimum year since it would minimize instances of prior involvement in deviant activities. This is an important consideration since certain of the hypotheses assume an initial deviance free state. Choosing seventh graders as the target population also permitted the observation of an adequate number of instances of adoption of deviant responses over the next two years for purposes of hypothesis testing. In addition, the period between ages 11 and 13, roughly the junior high school years, was said to be the developmental period that is most predictive of adult status--as contrasted with both younger and older age groups.

Of the eligible seventh grade students in the selected schools, 7,618 (81.6 percent) returned usable questionnaires at the time of the first administration. The demographic characteristics of these subjects are summarized in Table 1.

Table 1 about here

Those students who provided usable questionnaires at the first testing were not uniformly willing or able to continue participation in the study over the next two years. Of the students who returned usable questionnaires at the first testing 30 percent did not continue in the study. The remaining 70 percent responded to either or both of the Time 2 and Time 3 questionnaires. Of those subjects providing usable first administration questionnaires, 61.6 percent also provided data at Time 2 and 41.3 percent of those subjects providing usable first administration questionnaires provided questionnaire data at the second and third points in time as well.

Sample attrition was associated with ethnicity (Mexican-American), mother's education (did not graduate from elementary school), subject's age (older), religious affiliation (unaffiliated), mobility patterns (mobile), deviant behavior of friendship groups, attitudinal predisposition to leave school, and self-reports of early deviant behavior. There was no significant relationship observed between sample attrition on the one hand and gender, self-derogation level, psychophysiological indicators of subjective distress, birth order, broken family, reports of deviant activity on the part of "many of the kids at school," or stability of self-attitudes on the other hand.

Data Collection

The seventh grade students who were to take part in the study were convened at common locations in each school during the morning of a school day. At that time they responded to a 209-item structured self-administered questionnaire.

The test was administered three times, at annual intervals, during March or April of 1971 (T_1), 1972 (T_2), and 1973 (T_3). The questionnaires were identical each year with the exception of the time reference in items dealing with self-reports of deviant behavior.

Operational Definitions

Of the several variables operationalized in the analyses to be reported below the two most meaningful variables are self-attitudes and deviant behavior.

Self-attitudes. Self-attitudes were measured by scores on a self-derogation scale. This scale was composed of seven items, derived in an earlier study (Kaplan & Pokorny, 1969) from a factor analysis of response to 10 items (Rosenberg, 1965). Sample items are: At times, I think I'm no good at all; and, I certainly feel useless at times.

For students present at all three test administrations at annual intervals the correlations of self-derogation scores between the first and second tests was $r=.46$; between the second and third tests $r=.55$; and between the first and third tests (a two-year interval) $r=.40$.

virtue of the observation of theoretical relationships based on the assumption that the self-report items are indeed deviant as defined above. The second issue, regarding the relationship between the self-reports of behavior and the actual behavior, was addressed by considering the relationship between self- and other reports (the reports of vice-principal or school counselors) about the students' behavior.

On the basis of the observed relationships it was concluded that student self-reports generally could be used as a rough indicator of deviant behavior in the sense that the probability is far greater that a student who reports the act will have in fact committed the act (by the criteria of school personnel reports) than that a student who denies the act will have done so; or, conversely, that the probability is far greater that externally validated cases of deviant performers will so identify themselves than will performers whose deviant behavior is unknown. However, the self-report data should not be used to estimate the absolute number of deviant actors in a population.

Other scales. In addition to the strategic measures of self-attitudes and deviant behaviors, a number of other variables were considered as antecedents and/or consequences of these factors, or as moderators of these relationships.

With some exceptions, these measures were derived by principal components factor analyses with the factors rotated to the normalized varimax criterion of orthogonal simple structure. These scales include (with sample items indicated parenthetically): Perceived devaluation by peers (the kids at school are usually not interested in what I say or do); perceived devaluation by family (my parents do not like me very much); perceived devaluation by school (my teachers are usually not interested in what I say or do); perceived devaluation for ascribed characteristics (people often put me down because of my color); defenselessness/vulnerability (are you bothered by nervousness; when the kids at school dislike something I do, it bothers me very much); avoidance of personal responsibility for self-devaluing circumstances through the adoption of ineffective/deviant response patterns (it's mostly luck if one succeeds or fails; I don't care much about other people's feelings; I would like to travel with a circus or carnival); guilt deflection (if someone insulted me I would probably figure it was his own problems that made him do it; when things aren't going too well for me I try to think that things will be better in the future); self-enhancing potential of the deviant environment (the kids who mess up with the law seem to be better off than those who play it straight); self-enhancing potential of the normative environment (by the time I am 25 I will probably be happily married); awareness of deviant response patterns (do many of the kids at school take narcotic drugs).

Other single item measures reflect self-descriptions of traits and behaviors and subject valuation of the traits and behaviors.

Results

The results of analyses to date for the most part are reported in a series of publications (Kaplan, 1975a, c, d; 1976a, b, c; 1977a, b; 1978a, b, c, d; 1979; 1980a, b; Kaplan & Pokorny, 1976a, b, c; 1977; 1978).

Considering the findings for the study as a whole, the results of the analyses are highly compatible with central features of the theoretical structure. The consistency of the results with the theoretical structure is in many ways quite remarkable. However, the relatively large portion of the variance left unaccounted for does indicate the need for further specification of the theoretical conditions under which the mutual influences of self-attitudes and deviant behavior will be observed and for research processes that are appropriate to the testing of these specifications. Nevertheless, the theoretical structure under consideration, centering as it does around the concept of self-attitudes, continues to show promise as a useful framework for integrating existing, and guiding future, research on the antecedents and consequences of deviant behavior. Further opportunity to evaluate the usefulness of the theoretical structure will be provided by the follow-up study now underway of the original cohort approximately ten years after the initial testing.

References

- Cronbach, L. J. & Furby, L. How we should measure "change"--or should we? Psychological Bulletin, 1970, 74, 68-80.
- Kaplan, H. B. Increase in self-rejection as an antecedent of deviant responses. Journal of Youth and Adolescence, 1975a, 4, 281-292.
- Kaplan, H. B. Self-attitudes and Deviant Behavior. Pacific Palisades, Goodyear Publishing Company, Inc., 1975b.
- Kaplan, H. B. Sequelae of self-derogation: Predicting from a general theory of deviant behavior. Youth and Society, 1975c, 7, 171-197.
- Kaplan, H. B. The self-esteem motive and change in self-attitudes. Journal of Nervous and Mental Disease, 1975c, 161, 265-275.
- Kaplan H. B. Antecedents of negative self-attitudes: Membership group devaluation and defenselessness. Social Psychiatry, 1976a, 11, 15-25.
- Kaplan H. B. Self-attitude change and deviant behavior. Social Psychiatry, 1976b, 11(2), 59-67.
- Kaplan H. B. Self-attitudes and deviant response. Social Forces, 1976c, 54, 788-801.
- Kaplan, H. B. Antecedents of deviant responses: Predicting from a general theory of deviant behavior. Journal of Youth and Adolescence, 1977a, 6 (1), 89-101.
- Kaplan, H. B. Increase in self-rejection and continuing/discontinued deviant response. Journal of Youth and Adolescence, 1977b, 6 (1), 77-87.
- Kaplan, H. B. Deviant behavior and self-enhancement in adolescence. Journal of Youth and Adolescence, 1978a, 7, 253-277.
- Kaplan, H. B. Self-attitudes and multiple modes of deviance. In D. J. Lettieri (Ed.), Drugs and Suicide: When Other Coping Strategies Fail. Beverly Hills, California: Sage Publications, 1978b.
- Kaplan, H. B. Self-attitudes and schizophrenia. In W. E. Fann, I. Karacan, A. D. Pokorny, & R. L. Williams (Eds.), Phenomenology and Treatment of Schizophrenia. New York: Spectrum Publishing Co., 1978c.
- Kaplan, H. B. Social class, self-derogation and deviant response. Social Psychiatry, 1978d, 13, 19-28.
- Kaplan, H. B. Anxiety, self-derogation, and deviant behavior. In W. E. Fann, I. Karacan, D. D. Pokorny, & R. L. Williams (Eds.), Treatment of Anxiety. New York: Spectrum Publishing Co., 1979.
- Kaplan, H. B. Deviant Behavior in Defense of Self. New York: Academic Press, 1980a, in press.

- Kaplan, H. B. Self-enhancing functions of alcohol abuse among male adolescents. In W. E. Fann, I. Karacan, A. D. Pokorny, & R. L. Williams (Eds.), Phenomenology and Treatment of Alcoholism. New York: spectrum Publishing Co., 1980b, in press.
- Kaplan, H. B., & Pokorny, A. D. Self-derogation and psychosocial adjustment. Journal of Nervous and Mental Disease, 1969, 149 (5), 421-433.
- Kaplan, H. B., & Pokorny, A. D. Self-attitudes and suicidal behavior. Suicide and Life-Threatening Behavior, 1976a, 6 (1), 23-35.
- Kaplan, H. B., & Pokorny, A. D. Self-derogation and suicide--I: Self-derogation as an antecedent of suicidal responses. Social Science Medicine, 1976b, 10 (2), 113-118.
- Kaplan, H. B., & Pokorny, A. D. Self-derogation and suicide--II: Suicidal responses, self-derogation and accidents. Social Science Medicine, 1976c, 10 (2), 119-121.
- Kaplan, H. B., & Pokorny, A. D. Self-attitudes and alcohol use among adolescents. In F. A. Seixas (Ed.), Currents in Alcoholism, (Volume II). New York: Grune and Stratton, Inc., 1977.
- Kaplan, H. B., & Pokorny, A. D. Alcohol use and self-enhancement among adolescents: A conditional relationship. In F. A. Seixas (Ed.), Currents in Alcoholism, (Volume IV). New York: Grune and Stratton, Inc., 1978.
- Rosenberg, M. Society and the Adolescent Self-image. Princeton: Princeton University Press, 1965.

Table 1

Percentage Distributions of Subjects Completing First Questionnaire by Selected Demographic Characteristics

Gender	
Male	49
Female	51
Race/Ethnicity	
White-Anglo	61
Black	28
Mexican-American	11
Mother's Education	
Did not complete elementary school	4
Completed elementary school	13
Graduated high school	54
Graduated college	30
Age	
11 or younger	4
12	33
13	48
14 or older	16
Religion	
Protestant	59
Catholic	25
Jewish	3
Other Non-Christian	4
Unaffiliated	9

Family and Peer Processes in Adolescent Drug Use*

Denise B. Kandel

School of Public Health and Department of Psychiatry

Columbia University

and

New York State Psychiatric Institute

722 West 168th Street

New York, New York 10032

*This research was supported by Research Grant DA 00064 from the National Institute on Drug Abuse, and by the Center for Socio-Cultural Research on Drug Use, Columbia University.

Family and Peer Processes in Adolescent Drug Use

Denise B. Kandel

School of Public Health and Department of Psychiatry

Columbia University

and

New York State Psychiatric Institute

722 West 168th Street

New York, New York 10032

Introduction

The original objectives of this study were to examine the relative influence of parents and peers on the use of drugs by adolescents, to identify the nature of family contexts and family behavior within which drug use develops, and to analyze the family processes which promote or deter drug use among different types of adolescents, both independent of and relative to the influence of peers.

Studies Conducted: Design and Samples

Four related studies were carried out.

Longitudinal Study of New York State Adolescents

A two-wave panel survey with an interval of 5 to 6 months was conducted on a multistage random sample of adolescents representative of public secondary school students in New York State in fall 1971 (N = 8,206) and spring 1972 (N = 7,250).

Dates were also collected from parents of the students in every school, and from a best school friend in a subsample of schools. The two-stage sampling procedure entailed selection of (1) a stratified sample of 18 high schools, and (2) subsequent cluster sampling of homerooms within schools. The universe for the study included all public high schools in New York State, with grades 7-12, 9-12, or 10-12. The criteria for the ordering of the schools into four strata were: (1) geography (New York City; other cities over 200,000 or suburbs; rest of state); (2) minority (black or Puerto-Rican) student enrollment; (3) proportion of students going to college; and (4) size of school (whether above or below median for its geographic area). Six of the sample schools were located in New York City, six in suburban areas, and six in small towns or rural areas in upper New York State. The high schools ranged in size from 360 to 5,000 students. The total number of students eligible for inclusion in the study and the number of those who participated are displayed in Table 1.

Table 1 about here

Structured self-administered questionnaires were given to adolescents in selected homerooms in 13 schools and to the entire student body in five schools, making it possible in the latter to collect data from the student's best school-friend.

At both times, two to three weeks after each of the 18 schools was surveyed, questionnaires were mailed to one of the student's parents (alternately mother or father). A maximum of three follow-up contacts per family were scheduled. The first mailing included a questionnaire with a postcard to be signed and returned separately. This initial mailing was followed for nonrespondents by a reminder follow-up letter and by a second questionnaire. At Time 1, usable questionnaires were returned by 5,574 parents or 62 percent of the initial group contacted. The return rate was higher for mothers (66 percent) than for fathers (53 percent) and varied in different parts of the state, being lowest in New York City. Return rates ranged from 44 percent in one New York City school to 78 percent in a suburban school. The return rate of parental questionnaires at Time 2 (47 percent) was lower than at Time 1.

At Time 1, both student and parent samples were weighted to reflect the variable probabilities of selection of schools and homerooms, and to correct for different absentee and nonparticipation rates among students in each school and nonresponse rates among parents in each community.

In order to protect the rights and privacy of the adolescents in the survey and to obtain informed consent for their participation, all parents in the sample homerooms were notified of the student survey prior to its administration in each school, via certified mail return receipt requested, and were given an opportunity to refuse to have their child

participate if they so desired. The refusal rate was generally low, averaging 3.6 percent for the state as a whole.

In addition, no respondent signed any of the questionnaires. Legal and ethical considerations in drug research led to a record-linking scheme based on self-generated identification code numbers, based on middle letters of the adolescent's first and last names, date of birth, and last two digits of the phone number. Adolescents were asked to generate a number for the best school friend similar to the number the friend constructed for himself/herself. Parents generated a number for themselves identical to that created by the adolescents. Using these codes, we were able to match 49 percent of all sampled students to their parents and 38 percent of the students in the five schools to their best school friends. In the five schools, 1,110 students, or 23 percent of those surveyed, could be matched to parents as well as to best school friends and thus formed triads. Sixty-six (N = 5,468) of all students at Time 1 could be matched to themselves at Time 2. The resulting number of matched cases is lower than those obtained when matching is done on the basis of names (Kandel and Lesser, 1972). Furthermore, the elimination of certain respondents from the relational and panel samples constitutes an important source of bias. Students who cannot be matched, either to themselves over time or with a parent or friend at one point in time, exhibit consistent differences from those matched. The most important bias involved the order-representation of drug users in the matched samples. For example, 24 percent of the students who could be matched to themselves at Time 2 reported marijuana use at Time 1, compared with 41 percent of those not matched; 7 percent and 16 percent, respectively, reported use 40 or more times ever. The higher the frequency of marijuana use at Time 1, the greater the likelihood that the adolescent would start using other illicit drugs by Time 2. Furthermore, as compared to matched students, those unmatched tend to be black or Puerto Rican, from schools in New York City, from families with low income and low parental education, absent from school often, have poorer grades in school and lower educational aspirations, are less religious and less close to their parents. They are also more oriented toward their peers. Thus, 28 percent versus 47 percent get together with friends every day; 11 percent versus 25 percent have been absent from school 16 or more days since the beginning of the term; 63 percent versus 47 percent are students with grades of A and B. Similar differences appear between students matched and unmatched to a friend and/or parent at Time 1. To correct somewhat for these biases, the T1-T2 student panel sample was weighted to reproduce the frequency distribution of marijuana use observed at Time 1 in the total adolescent sample. The loss of cases in the panel samples must be recognized as a problem, even though the matched samples were not used to estimate incidence or prevalence rates of drug use in the New York State population, but to analyze interrelationships among variables. Table 2 summarizes the size of the various cross-sectional and panel samples available for analysis.

Table 2 about here

Absentee Study

To estimate the biases inherent in self-reported rates of drug use based on school surveys, a small study of students who were absent from school on the day of the survey was also carried out in spring 1972. Two different procedures were used: (1) household interviews, and (2) group administration of the questionnaires in an office setting, when it became apparent that students interviewed in the households reported very little use of drugs. In spring 1972, 238 students from three New York City schools who were absent from school on the day of the spring survey were contacted and invited to participate in a new survey. Forty-four students of those contacted (18.5 percent) agreed to participate in the second absentee study.

Follow-Up of 1972 Senior Class

In winter-spring 1972-73, during the year following high school graduation and five to nine months after the second survey, a third survey was carried out by mail and by telephone with members of the graduating senior class from each of the 18 schools. Responses were obtained from 69 percent (N = 1,635) of the target population; 51 percent responded to the mail questionnaire and 18 percent participated in the telephone interview of nonrespondents to the questionnaire.

Formation and Dissolution of Adolescent Friendships

The design of the longitudinal high school study made possible the creation of an unusual set of longitudinal data on adolescent friendship pairs. As noted above, data were collected on the entire student body at two points in time in five schools. Each adolescent was asked to identify his or her best friend in school. Information was available at two points in time on the characteristics of the adolescents, of the friend named at Time 1, and of the friend named at Time 2. These longitudinal data are unique not only because they focus on characteristics of adolescents during the friend-selection process, but also because they provide data on the characteristics of the adolescents and their friends before and after a given friendship. A special panel sociometric file was constructed to include information at both points in time on each adolescent, the friend named at Time 1 and the friend named at Time 2. Such data are available for 959 dyads, including 669 dyads with the same friend and 290 dyads with different friends at each time period. In 266 dyads where adolescents changed friends between Times 1 and 2, data are available on what the new friend was like at Time 1, before he or she was chosen as a friend. In 233 dyads, data are available on what the former friend was like at Time 2, after he or she had been dropped as a friend. Thus, complementary data are available on (a) the adolescent and the current friend; (b) the adolescent and the friend-to-be; and (c) the adolescent and the former friend.

Data Collection

Data were collected through self-administered structured questionnaires, designed so that identical items appeared in the student and parent versions. Three versions of the student questionnaire were developed (for Waves 1, 2 and 3), and two versions of the parent questionnaire (Waves 1 and 2). The variables consisted of single-item questionnaire items as well as factor-based scales that included multiple items. (See Table 3.)

Findings

Cross-sectional and longitudinal analyses were carried out to investigate the epidemiology of drug use among high school students in New York State, and the social, interpersonal and psychological correlates, determinants and consequences of various patterns of drug use.

Methodological Analyses

The validity and reliability of drug use responses (Kandel, 1975c; Single, Kandel and Johnson, 1975).

Developing valid and reliable measures is potentially more problematic for drug use than for other behaviors or attitudes because drug use is illegal and generally disapproved of by adults in the society. There may be pressures for adolescent respondents not to be candid in their answer for fear of the consequences. There is not only the potential risk of under-reporting but also that of over-reporting drug use. Anticipating these problems, exceptional steps were taken in the administration of the questionnaires to insure the respondent's confidentiality and anonymity.

The analysis of patterns of responses to drug use items indicates that these efforts were well rewarded by a high rate of responses to the drug questions and by valid and reliable answers. The number of students not answering the illegal drug use questions was very small, ranging from 1.6 percent (LSD) to 5 percent (marijuana). These rates are similar to the rates of nonresponse to the questions immediately preceding and following the illegal drug use items.

A number of results establish the validity of adolescent responses to the drug use items: (a) There are striking differences between self-reported adolescent drug users and nonusers on a variety of characteristics, differences which would not appear if respondents had given unreliable and invalid answers. Indeed, the group differences would be obscured by the inclusion of users in the non-using group. (b) There is a very high correlation between drug use patterns independently reported by each of the adolescents in friendship pairs. (c) Only a tiny proportion (less than one percent) claimed to have used a fictitious drug.

(d) The responses to different drug questions at different points in the questionnaire are consistent with each other. (e) Responses to drug use questions at Time 1 and also consistent with answers given several months later at Time 2. The proportion of adolescents who claimed to be nonusers in the first wave but indicated in the second questionnaire that they had been using at the time of the first wave, ranged from 0.3 percent (heroin) to 2.4 percent (marijuana). These rates are not higher than those observed with respect to other variables (3.3 percent of the panel sample reported a different race in the two waves). (f) The rates of self-reported drug use are similar to those found in other studies of high school populations. However, self-reported illicit drug use is less consistent over time than at one point in time. The failure to report certain types of drug use is associated with low lifetime drug experience and appears to be the result of poor recall rather than willful concealment of use.

Strategies for the Study of Developmental Stages (Kandel, Margulies and Dvaies, 1978).

The identification of stages in drug use led to an interest in the most appropriate method for identifying the predictors of transitions into developmental stages. Two alternate strategies were developed and illustrated on our data, using transitions into the three stages of legal and illegal drug use as a test case. The strategies include the use of a decomposed sample that is partitioned into groups at a particular stage who are at risk for transition into the next stage. The second strategy involves the use of a pooled sample, with step-up as the dependent variable, where a series of increasingly more complex models are run to test the notion that certain predictors are stage-specific.

The principles illustrated by these analyses are relevant to the study of developmental transitions in domains other than drug use. In particular, the strategies have implications for general research in socialization where a person must choose between successive behavioral alternatives.

Stages in Drug Use (Single, Kandel and Faust, 1974; Kandel, 1975a; Kandel and Faust, 1975).

A key notion developed from our longitudinal findings is that adolescent involvement in drug use follows well defined steps and sequences in which the legal drugs play a crucial role.

The results of the cross-sectional scalogram analyses, based on the total New York State adolescent sample, suggested that there exist stages in drug use. Direct evidence for the existence of stages required longitudinal data. Such evidence was provided by Guttman scale analyses of drug use responses over time. Analyses were replicated on two different cohorts: (a) the total New York State high school sample, followed

over one school year, and (b) the sample of graduated seniors who were contacted 5-9 months following graduation from high school. At least four distinct developmental stages in adolescent involvement in legal and illegal drugs were identified: (1) beer or wine; (2) cigarettes or hard liquor; (3) marijuana; and (4) other illicit drugs. The legal drugs are necessary intermediates between nonuse and marijuana. Whereas 27 percent of high school students who smoked and drank progressed to marijuana within a five-month follow-up period, only 2 percent of those who had not used any legal substance did so. Marijuana, in turn, was a crucial step on the way to other illicit drugs. While 26 percent of marijuana users progressed to LSD, amphetamines or heroin, only 1 percent of nonusers of any drug and 4 percent of legal users did so. This sequence was found in each of the four years in high school and in the year following graduation. The same steps were followed in regression as in progression in drug use (Kandel, 1975a; Kandel and Faust, 1975).

The identification of cumulative stages in drug behavior has important conceptual and methodological implications for studying the factors that relate to drug use. Users of a particular drug must be compared not to all nonusers, but only to the nonusers among the restricted group of respondents who have already used the drug(s) at the preceding stage(s). Otherwise, the attributes identified as apparent characteristics of a particular class of drug users may actually reflect characteristics important for involvement in drugs at the preceding stage(s). The definition of stages allows one to define a population at risk and to isolate systematically, within that population, those individuals who succumb to this risk within a specific time interval.

These findings are also important in that they indicate that concern over adolescent use of drugs should not be focused exclusively upon the use of illicit drugs, but should encompass the legal ones as well.

Adolescent Illegal Drug Use and Parental Drug Behavior: Cross-Sectional Analyses (Kandel, 1973; 1974a, b; Kandel et al., 1976).

Adolescent illegal drug use was examined as a function of two indicators of parental drug behavior: (a) the children's perceptions of their parents' use of tobacco, alcohol and psychotropic drugs (tranquilizers, barbiturates or sedatives, and stimulants); (b) the parents' self-reported use of each of these substances.

Our findings confirm previously reported associations between adolescent's drug use and the perceived parental use of a variety of psychoactive drugs. However, these relationships decrease when parents' actual reports of their psychotropic drug use is examined.

Self-reported parental marijuana and alcohol use have greater impact on adolescent illegal drug use than parental use of psychotherapeutic drugs (Kandel, 1974b).

Peer Influences on Adolescent Drug Use: Cross-Sectional Analyses

(Kandel, 1973, 1974a, b; 1975b, c; Kandel et al, 1976).

One of the most striking findings to emerge from our study is the documentation of the crucial role which peers play in the use of drugs by other adolescents. While previous surveys had suggested the importance of peers in youth's drug use, these findings were based upon the youths' perceptions of their friends' behavior. The independent data obtained from the student's best friend in school allowed us to examine and establish the role of peers on the basis of the friends' self-reported drug behavior. Involvement with other drug-using adolescents was found to be the most important correlate of adolescent drug use. Adolescents' marijuana use was strongly related not only to friends' perceived marijuana use but to the friends' self-reported use. Only 7 percent of adolescents who perceived none of their friends to use marijuana used marijuana themselves in contrast to 92 percent of those who perceived all their friends to be users. The most dramatic demonstration of friends' influence appeared when adolescent marijuana use was correlated with the self-reported marijuana use patterns of their best schoolfriend in the sample of friendship dyads. The proportion of users ranged from 15 percent when the best friend reported never to have used marijuana to 79 percent when the friend had used it 60 times or over. Not only use per se, but extent of use was highly related to friend's drug behavior. The proportion of adolescents who had used marijuana 60 or more times increased from 2 percent when their friends reported never to have used marijuana to 48 percent when their friends had themselves used the drug 60 times or more.

Having a friend who uses marijuana was related, as a minimum, to a fourfold increase in the number of adolescents who were themselves users of that drug. The increase with respect to drugs other than marijuana was at least as large, and sometimes even larger (Kandel, 1974b).

The greater importance of friends than parents on adolescent marijuana use was illustrated in the sample of matched triads when adolescent's marijuana use, best schoolfriend's marijuana use and parent's overall use of psychoactive drugs were examined simultaneously. While parents and best friends both had an independent effect on adolescent's marijuana use, the effects of peers was far larger than the effect of parents. This was best seen in those triads in which the adolescent was exposed to conflicting role models because parent's and friend's behaviors diverged, one using drugs and the other not. When faced with conflicting role models of drug use, adolescents are much more responsive to peers than to parents. Parental influence can, however, synergize with and potentiate peer influence. The highest rates of adolescent marijuana use occurred when both parent and peer reinforced each other's influence on the adolescent (Kandel, 1973).

Psychological Correlates of Illegal Drug Use (Paton and Kandel, 1978)

Depressive mood, normlessness, the feeling of being all alone in the world, and self-esteem are related to the use of drugs, and particularly the use of drugs other than marijuana.

Ethnic Differences in Correlates and Predictors of Drug Use

(Marel, 1977; Paton and Kandel, 1978).

As noted earlier, there exist strong differences in rates of use among different ethnic groups. Differences also appear in the role of parental and peer influences as predictors of initiation into drugs. At each stage of drug use, parental influence, especially in the form of the closeness of the relationship between parent and child, is a much more important predictor of black involvement than of white. This factor is the single most important predictor for blacks (Marel, 1977). On the other hand, the psychological factors of depressive mood and normlessness are more strongly related to the use of illicit drugs among whites than among blacks (Paton and Kandel, 1978).

Consequences of Drug Use: Depressive Mood and the Use of Illicit Drugs

(Paton, Kessler and Kandel, 1977).

Analyses of the relationship of depressive mood to illicit drug use over time reveal differing interactive relationships depending upon the drug involved. Depressive mood is related to the onset of marijuana use among nonusers and to the termination of marijuana use among marijuana users. Depressive mood also predicts the use of other illicit drugs by marijuana users. While the beginning use of illicit drugs other than marijuana is positively associated with increased depressive mood, continued multiple drug use is related to reduced depressive mood, suggesting that certain illicit drugs may be used by certain adolescents as self-medication.

Selection Versus Socialization in Peer Influence (Kandel, 1973; 1974a; 1978b,d).

Using the relational data obtained from the adolescent's best friend, major advances were made in the understanding of processes of peer influence and of friendship formation and dissolution.

In addition to the findings cited earlier on the correlates and predictors of drug initiation and involvement, other data document the extent to which experience with marijuana and with other drugs is closely related to the extent of involvement with peers. Use of marijuana and of other illegal drugs is what friends have most in common. With the exception of certain demographic characteristics (such as age, sex, and race), on no other activity or attitude (such as school attitudes and performance,

deviant behaviors of various kinds, political attitudes, drug-related attitudes, and attitudes toward parents) is similarity between friends as great as on illegal drug use. Similarity is lowest on psychological factors, selected attitudes, and reported quality of relationship with parents (Kandel, 1973, 1975a, 1978b).

Stressing that youths who use drugs associate together does not answer a basic question about the influence of peers: Which came first: drug use or drug using friends? Do adolescents seek out drug users after they themselves have become involved with drugs or do they start using drugs because they come to associate with other drug using friends? The longitudinal analyses clearly indicate that prior association with drug using friends predicts subsequent initiation into drugs. There definitely is a process of socialization in which adolescents influence each other over time and in which association leads to similarity.

CONCLUSION and IMPLICATIONS

The research has made contributions not only to understanding drug behavior in adolescence but also to fundamental issues in adolescent socialization (Kandel, 1978a, 1978e, 1980). Most of these insights were made possible because of the availability of longitudinal and relational data in the research.

The concept of developmental stages in drug behavior, a merger of constructs derived from sociology and from cognitive psychology, has allowed a more fruitful specification of the role and structure of different causal factors at different stages of a behavioral sequence. Findings on involvement with various forms of drug use illustrate that different types of theories have differential explanatory power at different stages of a phenomenon: social and cultural factors explain initiation into the behavior, while psychological factors are more powerful in explaining further involvement. The same approach could be fruitfully applied to other aspects of psychosocial development.

In addition, important aspects of interpersonal influence in dyadic relationships pertaining both to friends and to parent were elucidated. The similarity on attributes observed among friends at one point in time was shown to result from two independent processes: selection and socialization. On the one hand, individuals who share prior common interests and attributes are more likely to select each other as friends than those without such common interests. On the other hand, interacting individuals become more similar over time as the result of continued interaction. The research has also shown that parents exert influence on their children by means of modeling and social reinforcement. Parents influence their children both as role models whose behavior is emulated and as setters of specific standards of behaviors.

EPILOGUE

A further follow-up in young adulthood at age 23-24 of a subsample of adolescents formally enrolled in high school grades 10 and 11 is planned for spring 1980. This follow-up will be conducted through personal interview 8 years after the first initial contact and will include approximately 1,000 former regular students and 250 former school absentees. The overall objectives of the new follow-up will be (1) to carry out developmental studies that build upon the earlier research in adolescence so as to examine patterns of involvement and regression in drug use at a later point in the life cycle, (2) to identify the psychosocial precursors in adolescence of drug use in young adulthood, and (3) to specify the psychosocial consequences of such use.

Publications: Published Articles and Doctoral Dissertations

- Andrews, K. H. and Kandel, D. B. Attitudes and behavior: A specification of the contingent consistency hypothesis. American Sociological Association, 1979, 44, 298-310.
- Jessop, D. The two worlds of the family: Agreement in parent/adolescent responses in a drug survey. Doctoral Dissertation, Department of Sociology, New York University, 1979.
- Kandel, D. B. Adolescent marijuana use: Role of parents and peers. Science, 1973, 181, 1067-1070.
- Kandel, D. B. Inter- and intragenerational influences on adolescent marijuana use. In V. Bengtson and R. Lauffer (Eds.), Special Issue on "Generations and Social Change," Journal of Social Issues, 1974, 30, 107-135. (a)
- Kandel, D. B. Interpersonal and influences on adolescent illegal drug use. In E. Josephson and E. Carroll (Eds.), Drug use: Epidemiological and sociological approaches. Washington, D.C.: Hemisphere, 1974 (b).
- Kandel, D. B. and Bust, R. Sequence and stages in patterns of adolescent drug use. Archives of General Psychiatry, 1975, 32, 923-932.
- Kandel, D. B. Stages in adolescent involvement in drug use. Science, 1975, 190, 912-914. (a)
- Kandel, D. B. Reaching the hard-to-reach: Illicit drug use among high school absentees. Addictive Diseases, 1975, 1, 465-480. (b)
- Kandel, D. B. The measurement of "ever use" and "frequency quantity" in drug use surveys. In J. Elinson and D. Nurco (Eds.), Operational definitions in socio-behavioral drug use research, Monograph Series 2. Rockville, Maryland: National Institute on Drug Abuse, 1975 (c).
- Kandel, D. B. Some comments on the relationship of selected criteria variables to adolescent illicit drug use. In D. Lettieri (Ed.), Predicting adolescent drug abuse: A review of issues, methods and correlates. Washington, D.C.: Government Printing Office, 1975 (d).
- Kandel, D. B. "Letter to the editor." Journal of Studies on Alcohol, 1977, 38, 1843-1844. (a)
- Kandel, D. B. Convergences in prospective longitudinal studies of drug use in normal populations. In D. B. Kandel (Ed.), Longitudinal research on drug use: Empirical findings and methodological issues. Washington, D.C.: Hemisphere-John Wiley, 1978; and in S. Sells, W. Pollin, J. Strauss and M. Roff (Eds.), Life history research in psychopathology, Volume 6 1978 (a).
- Kandel, D. B. Similarity in real life adolescent friendship pairs. Journal of Personality and Social Psychology, 1978, 36, 306-312. (b)

- Kandel, D. B. On variations in adolescent subcultures. Youth and Society, 1978, 9, 373-384. (c)
- Kandel, D. B. Homophily, selection and socialization in adolescent friendships. American Journal of Sociology, 1978, 84, 427-436. (d)
- Kandel, D. B. (Ed.), Longitudinal research on drug use: Empirical findings and methodological issues. Washington, D.C.: Hemisphere-John Wiley, 1978 (e).
- Kandel, D. B. Drug and drinking behavior among youth. In J. Coleman, A. Inkeles and N. Smelser (Eds.), Annual review of sociology, Volume 6, 1980 (in press).
- Kandel, D. B., Kessler, R. C. and Margulies, R. Z. Antecedents of initiation into stages of drug use: A developmental analysis. In D. B. Kandel (Ed.), Longitudinal research on drug use: Empirical findings and methodological issues. Washington, D.C.: Hemisphere-John Wiley, 1978 (a).
- Kandel, D. B., Kessler, R. C. and Margulies, R. Z. Antecedents of initiation into stages of drug use: A developmental analysis. Journal of Youth and Adolescence, 1978, 7, 13-40. (b)
- Kandel, D. B. and Lesser, G. Youth in two worlds. San Francisco: Jossey-Bass, 1972.
- Kandel, D. B., Margulies, R. Z. and Davies, M. Analytical strategies for studying transitions into developmental stages. Sociology of Education, 1978, 51, 162-176.
- Kandel, D. B., Single, E. W., and Kessler, R. C. The epidemiology of drug use among New York state high school students: Distribution, trends and change in rates of use. American Journal of Public Health, 1976, 66, 43-53.
- Kandel, D. B., Treiman, D., Faust, R. and Single, E. W. Adolescent legal and illegal involvement in drug use: A multiple classification analysis. Social Forces, 1976, 55, 438-458.
- Kessler, R. C. Rethinking the 16-fold table problem. Social Science Research, 1977, 6, 84-107. (a)
- Kessler, R. C. The use of change scores as criteria in longitudinal survey research. Quality and Quantity, 1977, 11, 43-66. (b)
- Kessler, R. C., Paton, S. and Kandel, D. B. Reconciling Unidimensional and multidimensional models of patterns of drug use. Journal of Studies on Alcohol, 1976, 37, 632-647.
- Marel, R. Drug behavior of Black and White adolescents: Patterns, correlates, and predictors of use. Unpublished doctoral dissertation, Teachers College, Columbia University, 1977.

Margulies, R. Z., Kessler, R. C. and Kandel, D. B. "A longitudinal study of onset of drinking among high school students. Journal of Studies on Alcohol, 1977, 38, 897-912.

Parry, H., Cisin, I., Balter, M., Mellinger, G. and Manheimer, D. Increased alcohol intake as a coping mechanism for psychic distress. In R. Cooperstock (Ed.), Social aspects of the medical use of psychotropic drugs. Ontario: Addiction Research Foundation, 1974.

Paton, S. and Kandel, D. B. Psychological factors and adolescent illicit drug use: Ethnicity and sex differences. Adolescence, 1978, 13, 137-200.

Paton, S., Kessler, R. and Kandel, D. B. Depressive mood and adolescent illicit drug use: A longitudinal analysis. Journal of Genetic Psychology, 1977, 131, 267-289.

Single, E. W., Kandel, D. B. and Faust, R. Patterns of multiple drug use in high school. Journal of Health and Social Behavior, 1974, 15, 344-357.

Single, E. W., Kandel, D. B. and Johnson, B. The reliability and validity of drug use responses in a large scale longitudinal survey. Journal of Drug Issues, 1975, 5, 426-443.

Book

Denise B. Kandel (Ed.), Longitudinal research on drug use: Empirical findings and methodological issues. Washington, D.C.: Hemisphere-John Wiley, 1978 (e).

Articles by Users Using Data From Study

Davis, James A. Studying categorical data over time. Paper prepared for Conference on Strategies of Longitudinal Research on Drug Use, San Juan, Puerto Rico, April 7-9, 1976. Published in Social Science Research 1978, 7, 151-179.

Lazarsfeld, Paul F. Some episodes in the history of panel analysis. In Denise Kandel (Ed.), Longitudinal research in drug use: Empirical findings and methodological issues, Washington, D.C.: Hemisphere-John Wiley, 1978.

Coleman, James S. Purposive actors and mutual effects. In James S. Coleman, Robert K. Merton and Peter Rossi (Eds.), Qualitative and Quantitative Social Research, papers in honor of Paul Lazarsfeld. New York: Free Press, 1979.

Adolescence to Young Adulthood:
A Twelve-Year Prospective Study of
Problem Behavior and Psychosocial Development*

Richard Jessor and Shirley L. Jessor
Institute of Behavioral Science
University of Colorado

*Support for the long term research effort described in this paper has been provided by Grant No. AA-00232 and Grant No. AA-03745 from the National Institute on Alcohol Abuse and Alcoholism. This paper is Publication No. 182 of the Institute of Behavioral Science, University of Colorado, Boulder.

Adolescence to Young Adulthood:
A Twelve-Year Prospective Study of
Problem Behavior and Psychosocial Development

Richard Jessor and Shirley L. Jessor
Institute of Behavioral Science
University of Colorado

Introduction

This paper is a report from a long term, prospective study of problem behavior and psychosocial development. The study has had several objectives: to describe the course of psychosocial development over time; to predict the initial occurrence or onset of several important developmental transition behaviors; to explore the degree to which there is continuity between earlier and later developmental stages; and to test the explanatory usefulness -- both cross sectionally and longitudinally -- of a particular social psychological framework, namely, problem behavior theory. When completed, the study will have spanned an historical period from the late 60s to the early 80s, an elapsed time period of a dozen years between first and last data collection, and a developmental period from early adolescence well into young adulthood.

The research has been carried out in two separate phases, each of them longitudinal. The initial phase began in 1969 as a project called "The Socialization of Problem Behavior in Youth," and findings from that phase have been published in numerous papers and ultimately a book (Jessor & Jessor, 1977). More recently, in 1979, and with our colleague Dr. John E. Donovan, we began the current phase of the research with a project called "The Young Adult Follow-Up Study." Two separate samples of participants have been followed through time across both phases: one sample is made up of cohorts originally drawn from the 7th, 8th and 9th

grades of junior high school; the other sample is a cohort that was drawn from the freshman class in the College of Arts and Sciences of a major university. Both of the samples and all their cohorts include both males and females.

Since the later presentation of problem behavior theory and of the empirical findings that bear upon its usefulness will be limited to what was done in the initial phase only, it will be helpful to begin with a description of the overall study and the larger design that encompasses both phases of the work.

Part I: The Overall Design

The initial research phase was concerned with the early and late years of adolescence and with the life tasks and transitions that mark its course -- strengthening peer relations and loosening parental ties, initiating sexual activity, beginning the use of alcohol and other drugs, living away from home, completing one's education, etc. The central focus was on behaviors such as drinking and heavy alcohol use, marijuana involvement, activist protest, precocious sexual experience and proto delinquent actions such as aggression and stealing -- youthful behaviors defined as "problems" by the larger society and constituting at that time one of its salient preoccupations. The aims of that phase of the research were to explain variation in problem behavior by embedding it in a theoretical network of personality and environmental attributes and to illuminate the role of problem behavior in the normal process of psychosocial growth and development. Toward those ends, four annual waves of data were collected on the junior high school and the college youth.

After the initial phase was completed, there was a seven year hiatus in which all contact with the participants was suspended. During this period of time, all of the participants developed into young adults, the high school sample youth reaching their middle 20s and the college sample youth their late 20s. In 1979, when the current phase of the longitudinal study began, contact was renewed and a fifth wave of data was collected. The sixth and final data wave is scheduled for the Summer of 1981. Many of the concerns of the current phase are the same as, and actually continue from, the concerns of the earlier phase, especially the interest in problem behavior and its personality and social antecedents and correlates. But there are also many new areas of research interest reflecting the life tasks of young adulthood -- marriage or entering into committed relationships, child rearing, work and career, establishing an enduring sexual identity, stabilizing leisure interests, coming to terms with an image of self, etc.

Although each phase of the work is longitudinal and each can be seen as a separate, self contained study of a stage of the life trajectory, taken together they make possible the exploration of a whole new set of psychosocial developmental issues, issues that have to do with the continuities between life stages, with the predictability of young adult outcomes from antecedent adolescent information and -- of major societal importance -- with the possibility of identifying early stage risk factors that can signal the likelihood of later stage problems in adjustment and behavior. The success of such a two phase, longitudinal followup of an earlier longitudinal study depends, of course, on maintaining the intactness of the samples of participants over an extended period of time. It is that aspect of the research to which we turn next.

At the conclusion of the initial phase, in 1972 for the high school sample and 1973 for the college sample, there were 432 high school participants and 205 college participants who had filled out four successive annual questionnaires. Between that time and late 1978, when efforts were begun to locate and solicit the renewed participation of the members of the various cohorts, three had died. All of the remaining 634 former participants were ultimately located, contacted about the new phase of the study, reminded of their earlier signed statement of willingness to be followed up "... even ten years from now..." and asked to resume participation. They were also told that a Confidentiality Certificate from H.E.W. would permanently safeguard the privacy of their responses from all inquiries, including those from judicial, legislative or law enforcement agencies. Relying on several successive followup letters, persistent local and long distance telephoning and even personal visits to the homes of local participants who had not responded to the letters or phone calls, we ultimately were able to achieve a rather impressive level of renewed participation. Of the 634 participants available from the initial phase, fully 596 resumed a role in the followup research, an overall retention rate of 94 percent. That percentage figure is all the more remarkable in view of the fact that there had been no contact whatsoever in the intervening years and considering that renewed participation meant filling out a 65 page questionnaire that required an average time of two and a half hours and, for many, took as long as five hours to complete.

Insert Table 1 About Here

The success achieved in maintaining the intactness of the samples and the integrity of the cohorts can be seen in Table 1. For the college sample, renewed participation reaches 95 percent for both the male and

female cohorts. In the high school sample, there is a slightly higher percentage of retention for the females (95%) than for the males (92%) and for both sexes the lowest percentage is in the youngest cohort, the one originally drawn from the 7th grade. Nevertheless, of the eight separate age by sex cohorts (two in the college sample and six in the high school sample) only one fails to reach the 90 percent mark and that one -- the 7th grade males -- with 89 percent retention is close. The data in Table 1 are important not only because they insure the viability of longitudinal inference across the two phases of the study, nor in addition because they preclude any concern about significant or differential attrition, but also because they reinforce the feasibility of time extended research designs, even when relatively long intervals are involved.

It is now possible to present the overall structure of the study showing how it encompasses both the initial phase and the current phase of the research. The cohort sequential longitudinal design for the high school sample is presented in Figure 1 and the simple longitudinal design for the college sample is presented in Figure 2.

Insert Figure 1 About Here

It can be seen in Figure 1 that the cohorts were in 7th, 8th and 9th grades when the study began in the Spring of 1969; by the end of the initial phase, in the Spring of 1972, they had all made the transition from junior to senior high school and were in grades 10, 11 and 12 respectively. Adolescents aged 13, 14 and 15 at the outset, it can be seen that they will be young adults of 25, 26 and 27 when the final data are collected in 1981.

Insert Figure 2 About Here

The college sample, all freshmen in the Spring of 1970, were seniors, transfers, dropouts or -- for a few -- early graduates by the end of the initial phase in the Spring of 1973. When the current phase of the research began in 1979, they averaged 28 years of age and, when the sixth and final data wave is completed in 1981, they will be -- at age 30 -- on the threshold of the fourth decade of life.

This two phase, two sample, longitudinal design with samples chosen to cover adjacent portions of the life trajectory permits a span of developmental study from age 13 to age 30. This is not only a substantial age range but it is a range in which psychosocial growth is rapid and life change is pervasive. In short, it is a critical interval for the shaping of lives and for establishing the contours of adulthood. When the study is completed, we hope to be in a position to identify some of the shaping forces in psychosocial development and to map some of the contours that have emerged in these cohorts of young adults.

The Longitudinal Design of the Initial Phase

In The Spring of 1969, a random sample of 1,126 students, stratified by sex and grade level, was designated in grades 7, 8 and 9 of three junior high schools in a small city in the Rocky Mountain region. Students were contacted by letter and asked to participate in a four year study of personality, social and behavioral development. Parents were also contacted by letter and asked for signed permission for their child's participation. Permission was received for 668 students and of these 589 (52% of the random sample) took part in the Year I testing in the Spring of 1969. By the end of the Year IV testing in 1972, 483 students were still in the study representing 82 percent retention of the Year I participants. Of these, there were 432 students (188 boys and 244 girls) for whom there was no missing year of data. It is this latter group that constitutes our high school core sample for longitudinal or developmental analyses. The sample is composed of six separate, sex by grade cohorts as was shown in Table 1.

The core sample represents good retention (73%) of the initial year participants over four annual testings; it provides a wide range of variation on all measures; and it is large enough to permit the kinds of breakdowns needed for the analyses reported later. Although generalization to the parent population is precluded by the fact that the core sample constitutes only 38 percent of the original random sample, the core sample is nonetheless satisfactory for the testing of hypotheses about variation in behavior and development. Demographically, the sample is relatively homogeneous: almost entirely Anglo-American in ethnic background and middle class in socioeconomic status.

Data were collected in April and May of each year by an elaborate questionnaire, approximately 50 pages in length, requiring about an hour and a half to complete. The questionnaire consisted largely of psychometrically developed scales or indexes assessing the variety of personality, social, behavioral and demographic variables shown in the conceptual framework. Although many of the measures derive from and were validated in previous work (for example, Jessor et al., 1968/1975),

prior to its present use the entire questionnaire was pretested and scales were revised to increase their appropriateness for the student samples. The majority of scales were kept constant over the four testing years, but modifications were made in some and new ones were added at various times. Administration of the questionnaires took place in small group sessions outside of class and strict confidentiality was guaranteed because questionnaires had to be signed to permit followup.

For obtaining the college participants, a random sample of 497 freshman students was designated in the College of Arts and Sciences of a university in the same Rocky Mountain city. When contacted by letter in the Spring of 1970 and asked to participate in the research over the next four years, 462 students were still in school. Of those contacted, 276 (60%) participated in the Spring 1970 initial testing. By the end of the Year IV testing in 1973, 226 students were still in the study and 205 of these had no missing year of data. The latter group (92 men and 133 women) constitutes the core college sample: the members represent 41 percent of the original random sample and 74 percent of the participants who had been tested in the freshman year.

Dropping out of school or moving away from the community was negligible in the high school sample. In the college core sample, by 1973, 64 percent were still at the same university, 20 percent were at another university and 16 percent had dropped out of school at some point and not returned, even though they remained in the study.

In the college sample, data were also collected by questionnaires administered in small group sessions with confidentiality guaranteed. The questionnaire was very similar to that used in the high school and many of the scales were in fact identical. For the most part, especially where scales have more than a few items, measurement properties were quite satisfactory.

Interest in the research was high among both the high school and the college students and the quality of the questionnaire data was generally excellent. Participants seemed especially to appreciate the comprehensiveness of the questionnaire and its coverage of a wide range of content. Analyses of the attrition subsequent to the initial year of testing indicate that those who left the study were very similar on their initial year data in both studies to those who stayed. Thus, selective dropout from the studies does not seem to be a source of additional bias beyond the original erosion from the designated random samples.

Several features of the research design are worth emphasizing in relation to the methodological orientation of the study as a whole. The first and most apparent one is the provision made for pervasive replication

of observations and findings. For example, in both the high school sample and the college sample, there is opportunity to carry out four separate annual cross sectional tests of the explanatory usefulness of the social-psychological framework. In addition, within any year, findings can be replicated across sexes, across age or grade groups and across the two different school contexts. Considering the six sex by grade cohorts in the high school study and the two sex cohorts in the college study, there are actually eight independent subsamples in which any theoretical relationships may be separately examined. The possibility for such replication over time and across samples lessens the likelihood that findings would reflect the vicissitudes of a particular testing year or that the idiosyncrasies of a particular sample would be given more credence than deserved. Second, the descriptive interest in psychosocial development can obviously be pursued by following the cohorts through time with repeated measures. Third, the design makes possible the testing of the predictive implications of the theory by permitting the accumulation of data temporally antecedent to the event being predicted. The fourth feature worth mentioning is the role played by the theory in the content of the measures employed. Most of the major measures were theoretically derived to capture the logical properties of the concepts in the framework; as such, they make the data they yield germane to the testing of the theory in a way that ad hoc measures usually do not.

Conclusion

Despite the convergence of findings and their numerous replications across years of data and cohorts of participants, it needs immediate acknowledgement that the causal texture of the relationships we have been dealing with remains very much a matter of presumption. None of our strategies, not even the prediction of onset where a time lag was involved, can do more than document an association and the temporal order of the events or processes involved. That the subsequent events were "produced" by those that were antecedent still eludes direct demonstration, and even if demonstrated, the possibility of the reverse direction in other samples at other times cannot be ruled out. For social-psychological concerns, such as those dealt with here, this latter point is of special importance. Given the nature of the processes involved, it would be strange indeed if causal influence could not in fact operate in different directions in different instances, for example, becoming more tolerant of deviance influencing the exploration of marijuana in one case and the exploration of marijuana influencing a more tolerant attitude toward deviance in another. It may be that the preoccupation with univocal directionality of cause is an unwarranted legacy from experimental method in the physical sciences.

In behavioral science, it may be preferable to adopt a network model of causal influence, with the possibility of traversing from one point to another by a variety of pathways and in alternative directions. In such a perspective, the critical question becomes the relevance of the network.

The emphasis on inference, whether to causality or to relevance, ought not to divert our attention from the importance of the sheerly descriptive information yielded by the time extended observations. The data suggest important developmental regularities through the adolescent period in personality, the perceived environment and behavior -- regularities that reflect a developmental move away from conventionality. These regularities may, of course, be restricted to these samples or to this period of history; no claim is being made for them as developmental invariants. On the other hand, the trends observed are not at all dissonant from descriptions of adolescence that transcend the most recent period of time. The general point we wish to stress is the value of longitudinal study for purposes of describing the natural course of psychosocial growth and development per se.

In the content of our findings, there is quite impressive coherence. If a single summarizing dimension underlying the differences in personality were sought, it might be termed conventionality-unconventionality. The adolescent least likely to engage in problem behavior is one who values and expects to attain academic achievement, who is not much concerned with independence, who treats society as nonproblematic rather than as an object for criticism, who maintains a religious involvement and a more uncompromising attitude toward normative transgression and who sees little attraction in problem behavior relative to its negative consequences. The adolescent most likely to be involved in problem behavior shows an opposite pattern: a concern with personal autonomy, a lack of interest in the goals of conventional institutions like church and school, a jaundiced view of the larger society and a more tolerant view of transgression.

In the environment, the youth likely to be involved in problem behavior perceives less parental support, less compatibility between parents' and friends' expectations, greater influence of friends relative to parents and greater friends' support of and models for such behavior. With respect to behavior, the adolescent likely to be involved in any one problem behavior is likely to be more involved in other problem behaviors as well and to be less involved in conventional behavior than others.

Our findings have been generally similar for both males and females, a fact worthy of emphasis. The similarity between high school and college youth, however, is attenuated, particularly in the personality

system and in the distal structure of the perceived environment system, suggesting that development is not homogeneous throughout the early to late stages of adolescence and youth.

Overall then, problem behavior theory has emerged as a useful explanatory framework for problem behavior. The various research strategies reported have yielded convergent support for the social-psychological formulations. While such convergence does strengthen our conviction about the relevance of the theoretical network, it is still not enough to sustain a claim for directly demonstrated causal influence. This conclusion will, of course, continue to apply as our efforts are extended to linking up these data from the initial phase of the overall study with those from the current phase. The fact that the life stage of young adulthood follows adolescence in time is not the same thing as saying it is caused by that earlier life stage. Nevertheless, establishing connections between the two stages cannot help but advance understanding of each and at the same time of the process of psychosocial development that underlies those connections. That, in the last analysis, is the special reward of longitudinal research.

CONTINUED

2 OF 10

References

- Donovan, J.E. & Jessor, R. Adolescent problem drinking: Psychosocial correlates in a national sample study. Journal of Studies on Alcohol, 1978, 39, 1506-1524.
- Elder, G.H., Jr. Age differentiation and the life course. In A. Inkeles, J. Coleman & N. Smelser (Eds.), Annual Review of Sociology (Vol. 1). Palo Alto, Calif.: Annual Reviews, 1975. Pp. 91-123.
- Jessor, R. The perceived environment in psychological theory and research. In D. Magnusson (Ed.), Toward a psychology of situations: An interactional perspective. Hillsdale, N.J.: Erlbaum, 1981, in press.
- Jessor, R., Chase, J.A. & Donovan, J.E. Psychosocial correlates of marijuana use and problem drinking in a national sample of adolescents. American Journal of Public Health, 1980, 70, 604-613.
- Jessor, R., Graves, T.D., Hanson, R.C. & Jessor, S.L. Society, personality, and deviant behavior: A study of a tri-ethnic community. New York: Holt, Rinehart and Winston, 1968. (Reprinted by Krieger Publishing Co., 645 New York Ave., Huntington, N.Y., 1975.)
- Jessor, R. & Jessor, S.L. The perceived environment in behavioral science: Some conceptual issues and some illustrative data. American Behavioral Scientist, 1973, 16, 801-828.
- Jessor, R. & Jessor, S.L. Problem behavior and psychosocial development: A longitudinal study of youth. New York: Academic Press, 1977.
- Jessor, R. & Jessor, S.L. Theory-testing in longitudinal research on marijuana use. In D.B. Kandel (Ed.), Longitudinal research on drug use: Empirical findings and methodological issues. Washington D.C.: Hemisphere (Halsted-Wiley), 1978. Pp. 41-71.
- Jessor, R., Young, H.B., Young, E.B. & Tesi, G. Perceived opportunity, alienation, and drinking behavior among Italian and American youth. Journal of Personality and Social Psychology, 1970, 15, 215-222.
- Neugarten, B.L. & Datan, N. Sociological perspectives on the life cycle. In P.B. Baltes & K.W. Schaie (Eds.), Life-span developmental psychology: Personality and socialization. New York: Academic Press, 1973. Pp. 53-59.
- Riley, M.W., Johnson, M.E., Foner, A. (Eds.), Aging and society: A sociology of age stratification (Vol. 3). New York: Russell Sage, 1972.
- Thomas, W.I. The child in America. New York: Alfred A. Knopf, 1928.

Table 1
1979 Follow-Up Outcome
by Sample, by Sex, and by Cohort

	Total Available	Number Obtained	% Obtained
I. TOTAL SAMPLE	634	596	94
Males	279	259	93
Females	355	337	95
I. COLLEGE SAMPLE	204	193	95
Males	92	87	95
Females	112	106	95
III. HIGH SCHOOL SAMPLE	430	403	94
Males	187	172	92
7th Grade	74	66	89
8th Grade	60	54	90
9th Grade	53	52	98
Females	243	231	95
7th Grade	95	87	92
8th Grade	82	80	98
9th Grade	66	64	97

List of Figures Titles

Figure 1: Cohort-sequential longitudinal research design from Year I (1969) to Year XIII (1981) for the High School Sample.

Figure 2: Simple longitudinal research design from Year I (1970) to Year XII (1981) for the College Sample.

Figure 1: HIGH SCHOOL SAMPLE: Cohort-Sequential Design

	<u>JUNIOR HIGH SCHOOL GRADES</u>			<u>SENIOR HIGH SCHOOL GRADES</u>			<u>YOUNG ADULT AGE</u>				
	<u>7</u>	<u>8</u>	<u>9</u>	<u>10</u>	<u>11</u>	<u>12</u>	<u>23</u>	<u>24</u>	<u>25</u>	<u>26</u>	<u>27</u>
YEAR I - 1969	X	X	X								
Year II - 1970		X	X	X							
Year III - 1971			X	X	X						
Year IV - 1972				X	X	X					
Year XI - 1979							X	X	X		
Year XIII - 1981 (to be obtained)								X	X	X	

Figure 2: COLLEGE SAMPLE: Simple Longitudinal Design

	<u>YEAR IN COLLEGE</u>				<u>YOUNG ADULT AGE</u>	
	<u>FRESHMAN</u>	<u>SOPHOMORE</u>	<u>JUNIOR</u>	<u>SENIOR</u>	<u>28</u>	<u>30</u>
YEAR I - 1970	\bar{X} (205)					
YEAR II - 1971		\bar{X} (205)				
YEAR III - 1972			\bar{X} (205)			
YEAR IV - 1973				\bar{X} (205)		
YEAR X - 1979						\bar{X} (193)
YEAR XII - 1981 (to be obtained)						\bar{X}

Using Longitudinal Data to Study College Impact*

Alexander W. Astin

*Adapted from A.W. Astin, Four Critical Years, San Francisco: Jossey Bass, 1977.

Using Longitudinal Data to Study College Impact

Alexander W. Astin

The Cooperative Institutional Research Program (CIRP) is a national longitudinal study of the American higher educational system. CIRP was initiated at the American Council on Education in 1966 and is now conducted jointly by the Council and the University of California, Los Angeles. It is the largest ongoing study of the American higher educational system. Its longitudinal data now cover some 300,000 students and a national sample of more than 1000 postsecondary institutions of all types. Its follow up surveys of several different entering classes during the 1960s and 1970s provide an opportunity to replicate findings from one sample to another. Its data cover a wide range of cognitive and affective student outcomes and have resulted in publications on how the college experience affects more than 80 different measures of attitudes, values, behavior, achievement, career development and satisfaction. And its size and scope make it possible to employ highly sophisticated multivariate controls over a large number of potentially biasing variables -- in particular several means to differentiate purely maturational changes from those directly attributable to the college experience.

The task of assessing how students are affected by their college is composed of three major undertakings: (1) understanding the meaning of student change; (2) developing student outcome measures; and (3) designing the analyses of college impact. The following sections describe how CIRP deals with each of these three requirements.

Developing Outcome Measures of College Experiences

Recent controversy over the "value" of higher education (Freeman & Hollomon, 1975; O'Toole, 1975; Solmon, 1975) implies that the most important outcome of college attendance is economic: having a college

degree is supposed to help students develop cognitive skills, get better jobs and make more money. But colleges potentially have much more pervasive influence than this. An 18 year old who is leaving home for the first time to attend college is subject to wide-ranging influences from faculty, staff and fellow students. The possible influence of parents is reduced proportionately, simply because they are no longer present. Many freshmen experience their first intensive encounter with peers who have markedly different beliefs, backgrounds and attitudes. For some, enrolling in college may also provide the students' first direct experience with drugs, sex, alcohol and political activism. For others, college presents the first real challenge to their academic motivation and skills. The fact that many students spend four or more years attending college under these circumstances highlights the great potential of the college experience for producing both short- and long-range changes in values, attitudes, aspirations, beliefs and behavior.

In short, a thorough examination of the impact of college must take into account a wide range of possible outcomes. There is no easy way to capture the impact of college adequately in one or two simple measures such as credits and degrees or job placement. The need for a variety of outcome measures thus was anticipated in the design of CIRP. Rather than simply generating a list of miscellaneous measures, we developed a conceptual scheme to guide the selection of various measures. This "taxonomy of student outcomes" involves three major dimension: (1) type of outcome; (2) type of data; and (3) time.

Type of Outcome

Behavioral scientists have traditionally classified human performance into two broad domains: cognitive (sometimes called intellectual) and noncognitive (sometimes called affective). Since cognitive outcomes involve the use of higher-order mental processes such as reasoning and logic, they are clearly relevant to the educational objectives of most students, faculty, administrators, trustees, parents and others concerned with higher education. Noncognitive or affective outcomes concern the student's attitudes, values, self-concept, aspirations and everyday behavior and are important to many educators. Information on affective outcomes is relatively easy to obtain through self-administered questionnaires, whereas measurements of cognitive outcomes often require more controlled conditions of administration and larger amounts of the student's time. But both deserve attention in studying the impact of college.

Type of Data

The second dimension of the taxonomy, type of data, refers to the types of information gathered to assess the cognitive or affective outcomes. Again, two broad classes can be identified: psychological data, relating

to the internal states or "traits" of the individuals; and behavioral data, relating to directly observable activities. The measurement of the psychological phenomena is usually indirect, in the sense that the investigator, from responses to questions, infers some underlying state within the individual. Behavioral measures, which might also be called sociological, reflect transactions between the person and the environment and are usually of intrinsic interest.

Any student outcome measure can be classified simultaneously by the type of outcome involved and the type of data used (see Table 1). Each cell provides examples of different types of outcome measures obtained using different types of data.

Insert Table 1 about here

The cell on the upper left, for example, includes psychological measures of noncognitive or affective states: the student's ambition, motivation and self-concept, as well as subjective feelings of satisfaction and well-being. The cell on the upper right includes cognitive measures such as the student's grade point average or performance on multiple-choice tests of ability and achievement. The lower left cell includes sociological or behavioral features of the individual's development that reflect primarily affective states. Under personal habits, for example, one might include such behaviors as reading, eating, use of drugs, tobacco and alcohol. Citizenship would include such outcomes as voting behavior, participating in community activities and earning special awards for community service or on the negative side welfare or arrest records. The lower right cell gives examples of behavioral or sociological measures of cognitive outcomes. Basically, this category contains outcomes that reflect the behavior of the student (or former student) in society and that usually require cognitive skills. Presumably, real-life achievements represent the behavioral manifestations of the cognitive traits listed in the cell immediately above it.

The two dimensions that make up Table 1 are really more continua than true dichotomies. For example, a person's earned income probably depends in part on cognitive abilities, but it may also be affected by noncognitive or personality traits.

Time Dimensions

Since attending college can have both short-and long-term effects, the four cells in Table 1 could be extended into a third dimension representing temporal differences in student outcomes. Table 2 shows examples of related measures taken at two points in time.

Insert Table 2 about here

Although timing is seldom considered in discussions of educational outcomes, it is of fundamental importance. Most colleges hope to produce long-term rather than short-term changes. The goals stated in college catalogues, for example, imply that the institution is primarily concerned with making an impact that will last throughout a lifetime. The college, it would seem, tries to provide experiences that will help the student make the fullest possible use of his or her talents and become an effective, responsible member of society. Presumably such effects will, in turn, result in a more satisfying and rewarding life.

For many prospective college students, however, such long-term effects may be too remote and too difficult to comprehend. These students are primarily interested in more immediate goals -- their actual experiences during the undergraduate college years -- rather than in how these experiences will affect their later development. Educators frequently overlook the fact that the two, four or eight years of college represent a sizable portion of the student's total life span. For students, then, college experiences are important in themselves, not merely for what they will mean later.

Data for Assessing College Impact

The need for multiple indicators of the entering student's propensity for change and also the need for multiple-outcome measures were a major consideration in designing the Cooperative Institutional Research Program. Each fall, the entering freshmen at institutions participating in CIRP complete an extensive questionnaire composed of two types of items: (1) pretests on possible outcome measures, and (2) personal characteristics (age, race, educational background and so forth) that might affect the propensity to change or to attain certain outcomes. Because of time constraints during freshman orientation and registration, it is not possible to include pretests on all relevant outcomes in any one questionnaire. Thus, different freshman surveys have pretested somewhat

different outcomes. In the Fall 1966 freshman survey, for example, the student's self-concept at college entry was pretested with a lengthy list of self-ratings; these same ratings were posttested four years later in 1970. In the 1967 survey of entering freshmen, however, these self-ratings were replaced by a list of "competencies" that were posttested one year later in 1968.

Although it is possible to assess change in some outcomes with freshman pretests and posttests, certain other outcomes do not lend themselves to a pre-and posttest design. Whether a student drops out, for example, is an event for which no pretest is available. (One might argue that dropping out of high school is an analogous pretest, but its occurrence would simply preclude any measurement of college performance.) Similar outcomes for which pretests are either unavailable or inappropriate are getting married, winning special honors or awards and being admitted to graduate school. To deal with this problem, CIRP freshman questionnaires incorporate two types of items: (1) personal background data that might relate to the student's chances of attaining such outcomes, and (2) the student's own predictions about the likelihood of each outcome. Thus, the entering freshman questionnaires ask the student to estimate his or her chances of getting married, dropping out, participating in extracurricular activities and so on.

Not all CIRP follow ups include posttests of items from the freshman questionnaires, primarily because of constraints imposed by the agencies that fund the follow ups. In the 1974 follow up of 1969 entering freshmen, for example, the data are of limited value because posttesting of most freshman questions was not possible. Major longitudinal surveys used for assessing college impact are as follows:

<u>Entering Freshman Year</u>	<u>Follow Up Years</u>	<u>Number of Students</u>
1961	1965-1971	16,674
1966	1969-1970	19,422
1966	1970	25,399
1967	1968	5,351
1967	1969-1970	20,958
1967	1971	34,346
1968	1969-1970	21,458
1968	1972	41,356
1969	1969-1970	17,771
1969	1974	24,395
1970	1977	9,037
1971	1979	25,000 (est.)
1975	1977	16,191

With the exception of the 1969 follow ups, all follow up surveys have been conducted in the late summer or early fall of the year indicated. The 1969 "ACE-Carnegie" follow up of four freshman classes, which was conducted jointly by the American Council on Education (ACE) and the Carnegie Commission on Higher Education, was carried out between late December 1969 and February 1970.

Follow ups of each entering freshman class are conducted by sampling approximately 300 students from each participating institution. (Students from institutions enrolling more than 300 freshmen are selected at random.) Follow up questionnaires are mailed to students' homes. With postcard reminders and second questionnaires sent to nonrespondents, response rates varied from 40 to 65%, depending on the complexity of the questionnaire. Because extensive freshman input information is available on both respondents and nonrespondents, it is possible to develop a sophisticated weighting procedure to adjust for nonrespondent bias. In essence, the procedure gives greatest weight to those respondents who most resemble nonrespondents. (For details, see Astin & Molm, 1972; Carter & Brown, 1976.) The results, in other words, are generalizable to the national population of entering freshmen.

Statistical Methodology

The analyses involve two stages. First, information on each entering freshman is combined statistically through multiple-regression techniques to obtain a predicted or expected score (prognosis) on each outcome measure under investigation. Those freshman characteristics most closely associated with the particular outcome receive the largest weights while freshman characteristics unrelated to the outcome are given no weight. Freshman pretests, when available, usually receive the largest weight in predicting the corresponding posttest outcome measure.

The second stage of the analysis determines the effects of college experiences by comparing predicted outcomes based on entering freshman characteristics with actual outcome measures separately for students in different college environments. If an outcome is positively affected by attending a particular type of college, the actual outcome scores should be consistently higher than the expected scores among students enrolling at that type of college and consistently lower than expected for students enrolling at other types of colleges. An important characteristic of multiple-regression analysis is that, for the total group of students used to derive weights for entering freshman characteristics (in this instance, students across all colleges), the discrepancies between expected and actual outcome measures will sum to zero. In other words, students whose actual outcomes exceed their expected outcomes are balanced by an equal number whose expected outcomes exceed their actual outcomes. If a particular type of college has no effect on

a given outcome, students entering colleges of that type should perform like the total group. The positive and negative discrepancies between expected and actual outcomes should balance out (that is, sum to zero). However, if a particular environment positively affects an outcome, the actual outcome measures for students in that environment will tend to exceed their expected measures. In other words, the mean actual outcome for those students will be higher than the mean expected outcome based on entering characteristics. The reverse will be true, of course, when an environment has a negative effect on an outcome: the mean expected outcome will exceed the mean actual outcome for students in that environment. The strength of the effect of a particular environment will be reflected in the magnitude of the difference between mean expected and mean actual outcomes for students in that environment. This difference can be expressed in terms of a partial correlation between the environment and the outcome (that is, the correlation, if any, that remains after the effects of entering characteristics have been "partialled out").

The logic of this type of research design can be illustrated with an example. In a follow up questionnaire completed 3-1/2 years after entering college, students were asked whether they smoked cigarettes frequently (score 3), occasionally (score 2) or not at all (score 1). Responses to this follow up question could be predicted with moderate accuracy from the students' entering freshman data, as indicated by a multiple-correlation coefficient of .57. It is not surprising that freshman pretest responses to the same smoking question carried by far the largest weight; but a number of other freshman items, such as being black and being female, also carried significant positive weights. In other words, blacks and women were more likely than nonblacks and men to have increased their smoking during the 3-1/2 year interval.

Control Groups

Many scholars in higher education believe that a definitive study of college impact should include a "control group" of persons who never attended college. By comparing such a group with college attenders, the investigator would presumably be able to differentiate purely maturational changes from changes specifically attributable to the effects of college.

There is, however, a somewhat different approach that can be taken to the task of separating college effects from maturational effects. Rather than comparing college attenders with nonattenders, students have been sorted into groups according to their degree of exposure to the college experience. The underlying rationale for this approach is quite simple: if certain outcomes are facilitated by the experience of attending college, the likelihood of such outcomes should be greatest for those students who have the greatest exposure to the college

environment. However, if an outcome is a normal part of maturation and is not affected by going to college, its occurrence should not depend on how much exposure to college the student has had.

Because of the great diversity of institutional types and the substantial variations in college attendance, students differ widely in their exposure to the college experience. The current study approaches the issue of exposure from two perspectives: time of exposure and intensity of exposure.

Time of Exposure

Time of exposure can be handled quite simply: how long was the student enrolled in college? Some students register for college and never show up for their first classes. Others drop out before completing the first term. Still others persist for the entire four undergraduate years. The relative importance of college effects and maturational effects can be assessed by determining answers to two questions: (1) are changes in people who stay in college for a short time comparable to changes in people who stay longer? and (2) are the effects of particular college characteristics stronger for people who stay longer? (See Astin, 1977.)

The lowest point on the time continuum lacks the extreme case of a group that had no exposure to college. Nevertheless, this continuum-of-exposure approach has certain advantages over the college-noncollege design. In laboratory experiments, the control group is supposed to be comparable to the experimental group in all respects except the "treatment" variable -- in this case, degree of exposure to the college environment. It may be that a group of students with little exposure to college is a less biased control group than one that has never attended college. The latter group would differ from college attenders in certain critical respects: most would not have gone through the process of applying to college and presumably many were never motivated to attend college in the first place. Although it is true that college dropouts differ from nondropouts in such characteristics as ability and achievement (Astin, 1971, 1975), similar differences between those who do and do not attend college may be even greater (Sewell & Shah, 1968; Trent & Medsker, 1967; Cooley & Flanagan, 1966). Furthermore, since much is known about factors that predispose students to drop out of college (Astin, 1975), this knowledge can be utilized to control for preexisting differences between early dropouts and college completers.

One problem with using time of exposure to the college environment is that student dropout rates differ widely among different types of institutions. Students entering private universities, for example, are

much more likely to complete four years of undergraduate work than are students initially entering other types of institutions. Ignoring such institutional differences would confound time of exposure with type of institution attended. To deal with this problem, analyses involving time of exposure to the college environment have been limited in two ways. First, all students who attended more than one institution are excluded. Next, within any given institution, equal numbers of students are selected from each time-of-attendance category (for example, less than one year, one and a half years and so forth). One effect of this latter procedure is to limit the students selected from each time category at any institution to the number in the category containing the fewest students -- usually, but not always, the shortest time interval. Although these selection criteria necessitate excluding some institutions (notably the two-year colleges, because few have any students attending more than two years), they remove any correlation between time of attendance and type of college attended.

Intensity of Exposure

The quality or intensity of the student's college experience can be measured in terms of student involvement in the college environment. The construct of involvement, first proposed in a recent study of dropouts (Astin, 1975) is the time and effort expended by the student in activities that relate directly to the institution and its program. Students at the low end of the involvement continuum are those who live off campus, who come only to attend classes, who devote minimum effort to their academic activities and whose lives are concerned primarily with persons and events outside the institution. At the high end of the continuum are students who spend most of their time on campus, are committed to their studies, are actively involved in campus organizations and interact frequently with faculty and other students.

These stereotypic descriptions of persons at two extremes on the involvement continuum suggest that the construct of involvement may be multidimensional. That is, a person can manifest involvement in different aspects of the college experience: academic, social, political and so forth. For this reason, a number of analyses have been performed to determine if there are identifiable patterns of student involvement in the undergraduate college experience. The results of these analyses are described briefly here.

The search for different patterns of involvement first means identifying items from the follow up questionnaires that reflect student involvement in the institutional environment. These items include the student's residence, extracurricular activities, membership in academic honor societies and patterns of interaction with fellow students and faculty.

Involvement items were identified in three follow up surveys: the 1970 follow up of 1966 freshmen, the 1972 follow up of 1968 freshmen and the Winter of 1969-1970 American Council on Education-Carnegie follow up of 1966 freshmen. Involvement items from each follow up were intercorrelated and factor analyzed to identify possible patterns of involvement.

These analyses did not reveal a single general involvement factor. Instead, the involvement items formed a number of relatively independent clusters or "factors" in statistical parlance. Several patterns appeared in all three analyses. Other patterns were unique to only one analysis, primarily because comparable items were not included in the other two. Results are summarized in the following sections for eight major areas: place of residence, involvement with faculty, familiarity with professors in major field, verbal aggressiveness, academic involvement, involvement in research, student government and athletics.

Place of Residence

Where the student lives in a potentially important index of student involvement. First, on purely intuitive grounds, one may assume that students who live on campus in a dormitory have many more opportunities and incentives to become involved in campus life than students who live at home. Second, a growing body of literature (Astin, 1973, 1975; Chickering, 1974) indicates that students who live in residence halls have more contact with faculty, interact more with student peers, do better academically and are more satisfied with their undergraduate experience than are commuters. Although residential factors appeared in all three longitudinal samples, the most detailed data were available from the 1968-1972 sample. The principal analysis of these data involved only those students who were in continuous attendance at the same institution for all four undergraduate years (N = 9,682). Results showed considerable consistency in the students' residential patterns during the first two years. Over half (58%) lived in a dormitory both years. About 8 percent moved from dormitories to fraternity or sorority houses between the first and second years, and about 5 percent moved from dormitories to private rooms or apartments. Only 2 percent moved out of dormitories to live with parents. Thus, of the nearly three in four students (72% of all freshmen) who lived in dormitories as freshmen, most (eight in ten) continued to live in dormitories as sophomores.

A comparable pattern was observed for students who spent their freshman year living with parents. One student in five (20% of the total group) lived with parents during the first year and moved into a dormitory during the second year, and only 1 percent left parents to live in private rooms or apartments during the second year. Again, of the nearly one freshman in four (24%) who lived with parents, the large majority (eight in ten) continued to live with parents during the sophomore year.

Only 2.5 percent of the students who stayed in the same college for four years initially lived in a private room or apartment as freshmen, and most of these (1.9% of the total sample) lived in this fashion during the first two years. Virtually no freshmen who lived either with parents or in private rooms or apartments moved into fraternity or sorority houses during their sophomore years. In summary, among students who attend the same college for four years, most who live in private rooms or apartments during the sophomore year and nearly all who live in fraternity or sorority houses as sophomores initially begin college living in dormitories.

Analyses involving the 1966-1970 sample yielded several additional patterns of residence involving the third and fourth undergraduate years. Basically, students who lived in dormitories during the first two years split into two groups of about equal size: those who continued to live in dormitories and those who moved into private rooms or apartments. The latter group was about equally divided between rooms and apartments.

These analyses suggest three distinct patterns of student residence: (1) living continuously at home; (2) living continuously on campus in dormitories or fraternity or sorority houses; and (3) mixed patterns. One would expect the students living on campus to be most involved in the institutional environment, students living at home to be least involved and students with mixed patterns or living in private, off-campus housing, to fall somewhere in between. Thus, if a particular student outcome is not merely maturational but is affected by the college experience, that outcome should be most likely for students living on campus and least likely for those living at home.

Involvement with Faculty

The factor analysis of the 1966-1969-1970 ACE-Carnegie data revealed a clear-cut pattern of student involvement with faculty. To obtain an involvement score for each student, responses to 13 items were summed (with the scoring reversed on all items with negative loadings). A similar procedure was used in scoring the other involvement factors, as follows.

Familiarity with Professors in Major Field

This factor was identical to one reported in an earlier study of the college environment (Astin, 1968). Students were asked to rate one of their major instructors as follows: "Think about the course you took during your most recent college term that was most closely related to your primary field of interest. Please mark 'yes' for all the following statements that apply to this course. (If the course had a lab portion, mark 'yes' only for those items that apply to the lecture portion.)" Four items made up the factor.

Still another way to estimate the effects of maturation is in terms of the age of the entering student. If a particular change is in part the result of maturation, the older the student the less likely that student would be to exhibit the change. That is, if a particular change occurs in most young people during the interval from, say, 18 to 20, regardless of whether the person attends college, then it is reasonable to expect that students who are already 20 when they first enter college would be less likely to exhibit these changes than students who are 17 or 18. In short, a negative relationship between any given change and age at college entry would constitute evidence that the change is in part maturational.

Summary

This essay has considered a number of conceptual and methodological problems associated with studying college impact, presenting in some detail the general design of the impact analysis used in studies from the Cooperative Institutional Research Program.

First, the fact that students change in certain demonstrable ways while attending college is not sufficient evidence to conclude that college attendance per se had produced the change. One also needs some basis for judging whether the same change would have occurred if the student had attended a different college, or no college at all.

Second, because the college experience has the potential to affect many aspects of students' lives, the impact of college cannot be adequately assessed in terms of one or two simple outcomes, such as attaining a degree or earning a certain income. An adequate assessment requires a variety of cognitive and affective outcome measures.

Third, the factory or production model of higher education in which credits, degrees and graduates are "produced" by the institution has been rejected as a conceptually inadequate representation of the process of higher education. To assess college impacts, a medical or treatment model is a more appropriate analogy, primarily because both medical and educational institutions provide services designed to enhance the development of the individual. The effectiveness of these services cannot be assessed solely in terms of the status of the individual at some end point; rather, the person's final status must be evaluated in terms of initial status at the point of entry into the institution. Initial status (the person's educational potential) is analogous to a medical prognosis.

Fourth, much early research on college impact has produced inconclusive findings, primarily because of limitations in the data and methods of analysis. The CIRP which provides the principal empirical basis for the longitudinal studies summarized in Four Critical Years (Astin, 1977) has a number of advantages in studying college impact. Its data are collected from students attending a wide variety of collegiate institutions representing all major types. Several independent student samples are included, making replication possible. A variety of outcome measures, covering both cognitive and affective as well as psychological and behavioral outcomes, are included. And CIRP data are longitudinal, with both freshman pretests and posttests available for most outcome measures. Freshman data also cover a wide range of personal and background information, which makes possible control of numerous potentially biasing variables.

Fifth and finally, many investigators have assumed that the "ideal" study of college impact requires comparing college attenders with a "control" group of nonattenders, but that design oversimplifies the college impact problem. Given the great variety of institutions and programs, it is necessary to assess the impact of college characteristics and collegiate experiences rather than "college" as such. In addition, the matter of attendance versus nonattendance is more a continuum than a dichotomy. Thus, CIRP has been designed so students can be compared in terms of their degree of exposure (minimal time in attendance versus full time attendance for four years) and intensity of exposure (degree of actual involvement in the collegiate experience). Through such comparisons the analyses can separate purely maturational changes from those changes attributable to college experiences.

References

- Astin, A.W. The college environment. Washington, D.C.: American Council on Education, 1968.
- Astin, A.W. The methodology of research on college impact (I). Sociology of Education, 1970a, 43, 223-254.
- Astin, A.W. The methodology of research in college impact (II). Sociology of Education, 1970b, 43, 437-450.
- Astin, A.W. Predicting academic performance in college. New York: Free Press, 1971.
- Astin, A.W. Impact of dormitory living on students. Educational Record, 1973, 54, 204-210.
- Astin, A.W. Preventing students from dropping out. San Francisco: Jossey-Bass, 1975.
- Astin, A.W. Four critical years. San Francisco: Jossey-Bass, 1977.
- Astin, A.W. & Molm, L.D. Correcting for nonresponse bias in follow-up surveys. Unpublished manuscript, Graduate School of Education, University of California, Los Angeles, 90024, 1972.
- Astin, A.W., Panos, R.J. & Creager, J.A. National norms for entering college freshmen -- Fall 1966. Washington, D.C.: American Council on Education, 1967.
- Carter, A.M. & Brown, M.D. Manpower survey of 1967-1969 college entrants. Los Angeles: Graduate School of Education, University of California, 1976.
- Chickering, A.W. Commuting versus resident students: Overcoming the educational inequities of living off campus. San Francisco: Jossey-Bass, 1974.
- Cooley, W.W. & Flanagan, J.D. (Eds.). Project talent: One year follow up studies. U.S. Department of Health, Education and Welfare Cooperative Research Project No. 2333. Pittsburgh: School of Education, University of Pittsburgh, 1966.
- Feldman, K.A. & Newcomb, T.M. The impact of college on students. Vol. 1: An analysis of four decades of research. San Francisco: Jossey-Bass, 1975, 7, 24-31, 62.
- O'Toole, J. The reserve army of the underemployed: II-The role of Education. Change, 1975, 7(5), 26-33, 60-63.
- Sewell, W.H. & Shah, V.P. Parents' education and children's educational aspirations and achievements. American Sociological Review, 1968, 33, 191-209.
- Solomon, L.C. The definition of college quality and its impact on earnings. Explorations in Economic Research, 1975, 2, 537-587.
- Trent, J.W. & Medsker, L.L. Beyond high school: A study of 10,000 high school graduates. Berkeley, Ca.: Center for Research and Development in Higher Education, University of California, 1967. (Published subsequently by Jossey-Bass, San Francisco, 1968.)

Table 1

Classification of Student Outcomes
By Type of Outcome and Type of Data

Data	Outcome	
	Affective	Cognitive
Psychological	Self-concept Values Attitudes Beliefs Drive for Achievement Satisfaction with College	Knowledge Critical Thinking Ability Basic Skills Special Aptitudes Academic Achievement
Behavioral	Personal Habits Avocations Mental Health Citizenship Interpersonal Relations	Career Development Level of Educational Attainment Vocational Achievements: Level of Responsibility Income Awards or Special Recognition

Source: Astin, Panos and Creager (1967, p. 16).

Table 2

Examples of Measures Representing
Different Types of Data and Outcomes

Types of Outcome	Type of Data	Time 1 (During College)	Time 2 (After College)
Affective	Psychological	Satisfaction with college	Job Satisfaction
Affective	Behavioral	Participation in student government	Participation in local or national politics
Cognitive	Psychological	Law School Aptitude Test Test (LAST) score	Score on law boards
Cognitive	Behavioral	Persistence in college (staying in versus dropping out)	Income

Project Talent:

A Longitudinal Study of the Development and
Utilization of Individuals' Capabilities*

Lauri Steel

Laurens L. Wise

Ronald P. Abeles

American Institutes for Research

*Much of the material presented in this chapter is reproduced from "Project TALENT: Studying the development of our human resources," in J.E. Milholland, Ed., New directions for testing and measurement: Insights from large-scale surveys (Jossey-Bass, 1979).

Project Talent:
A Longitudinal Study of the Development and
Utilization of Individuals' Capabilities

Lauri Steel

Laress L. Wise

Ronald P. Abeles

American Institutes for Research

Introduction

Project TALENT is, perhaps, the most ambitious effort to date to investigate the personal, educational, and experiential factors that promote or inhibit the development of human talents. A longitudinal study of over 400,000 American who were in high school in 1960, TALENT grew out of an awareness of the need for sound information about the processes by which men and women develop and use their abilities. In the words of the then-U.S. Commissioner of Education Lawrence W. Dertnick, Project Talent represented

. . . an attempt to find out more about the students' interests, their career plans, and whether the courses they take are consistent with the life objectives they have set for themselves. And above all, it is an attempt to determine why so much of the nation's human potential is lost and what schools, counselors, and parents can do to reduce this loss. (Flanagan, Dailey, Shaycoft, Gorham, Orr, & Goldberg, 1962, p. 1)

TALENT was initially conceived as a means of developing a national inventory of human resources, achieving a better understanding of how young people choose and develop in their careers, and identifying the educational and life experiences which are most important in preparing individuals

for their life work. Today, through periodic follow-ups since 1960, TALENT has become a rich source of data concerning the life development of our nation's young adults--their backgrounds, abilities, and aspirations, their educational and career development, and various aspects of the quality of their lives.

Creating this data base required a sizeable investment of funds from the federal government as well as a sizeable investment of time and effort on the part of the TALENT participants. In order to assure the maximum return on these investments, the Project TALENT data are made available for secondary analyses through the Project TALENT Data Bank. To date, over 400 studies of women's and men's educational, career, and personal development have been carried out using Project TALENT data.

The original plans for Project TALENT included seven areas of inquiry responding to particular areas of national concern.

1. Available talent. More precise information was needed as to the size of the personnel pool qualified for training in science, engineering and other professional fields.
2. Relations among aptitudes, interests, and other factors. Available knowledge was particularly deficient with respect to the interrelation patterns of aptitudes, preferences, interests, socioeconomic factors, and motivational factors.
3. Limiting effects resulting from lack of interest and motivation. Specific information was needed regarding the extent to which people qualified for training in scientific and professional fields lack interest and motivation for obtaining this training.
4. Factors affecting vocational choice. More information was needed regarding the dynamics of personal decisions and environmental factors which ultimately determine the individual's occupational career pattern.
5. Predictors of creativity and productivity. Better information was needed with respect to the relation of aptitude, interest, and motivational factors to creativity and productivity in particular professional fields.
6. Effectiveness of various types of educational experience. Basic information was needed regarding the effectiveness of various types of educational experience in developing special talents.
7. Procedures for realizing individual potentials. The final overall requirement was for sound information regarding

procedures for assisting individuals to realize their highest potential (Flanagan et al., 1962).

The design of TALENT called for the comprehensive assessment of students' characteristics while in high school (grades 9-12), with periodic follow-up surveys conducted 1, 5, 10, and 20 years following the expected year of graduation from high school. This design has been and is being implemented with only a few modifications as shown in Table 1.

Some of the original objectives (e.g., providing information on the relations among aptitudes, interests, and other factors) were completed with the initial data collection in 1960. Other objectives (e.g., determining early predictors of professional productivity) will not be achieved until after the 20-year follow-up surveys are completed. New objectives have been added during the 20 years since Project TALENT was begun, such as investigating who goes to college and how different students finance their college education and, more recently, determining the factors that comprise the subjective quality of the respondents' lives and examining the influence of educational and other experiences on these life quality factors. Still other objectives relating to adults' career and personal development were identified in conjunction with developing plans for a 17-year follow-up of the TALENT participants. These plans are described in Designing a Study of Adult Achievement and Life Quality (Steel & Wise, 1977).

The Talent Samples

A number of different overlapping samples or subsamples comprise the pool of Project TALENT participants. The general long-term objective of Project TALENT--to relate high school interests, abilities, activities, background, and environmental data to subsequent occupational choices and achievements--required a very large sample in order to include a sufficient number of students who would later pursue particular occupations. The basic TALENT sample, the Project TALENT National Probability Sample, thus consists of over 375,000 students from 987 high school systems selected to be representative of all 9th, 10th, 11th, and 12th grade students in the United States in 1960. In addition to this basic probability sample of students in grades 9-12, the 1960 testing included all 15-year-olds not in grades 9-12 in the areas served by 10 percent of the participating schools. The data on 15-year-olds in grades 9-12 plus the data on these additional 15-year-olds constitute the TALENT Fifteen-Year-Old Sample.

A number of students were tested in 1960 in addition to those included in the National Probability Sample and the Fifteen-Year-Old Sample. Many of these are in the Knox County Saturation Sample consisting of

Table 1
 Modal Age of Respondents in Relation to
 TALENT Data Collection Schedule

Data Collection Date	Grade in 1960				Data Collection Effort
	12th	11th	10th	9th	
1960	17½	16½	15½	14½	Initial Testing
1961	19				
1962		19			
1963			19	17½ ^a	1-Year Follow-Up
1964					
1965	23				
1966		23			
1967			23		5-Year Follow-Up
1968				23	
1969					
1970					
1971	29				
1972		29			
1973			29		11-Year Follow-Up
1974				29	
1975					

Note. All dates are for the Fall except the 1960 testing which was in the Spring.

a. Approximately 7,500 students retested as high school seniors

nearly every student in grades 8-12 in Knox County, Tennessee (approximately 15,000 students). The inclusion of these students was the result of the enthusiasm of local school authorities. While limited to a specific location, this sample has high potential value for certain studies because of the testing of 100 percent of the students in a five-grade range.

The initial samples were further augmented by nearly 20,000 students who were in schools where the testing of 8th grade 15-year-olds was expanded to include all in the class; in schools that were not in the national sample, but who asked to be included along with other schools in their area; or in schools selected for special sub-studies. While most analyses are limited to the probability samples, data on these additional participants may be used when the total number of cases is more important than nationally representative estimates.

In 1963, all the 12th grade students in a 10 percent sample of the TALENT schools were tested using the 1960 instruments. Most of these individuals had been previously tested in 1960 as 9th graders, and hence comprise the Retest Sample.

Each grade cohort in the high school probability sample has been followed up via mailed questionnaires 1, 5 and 11 years after its high school graduation. In each case, both the respondents to the mailed questionnaires and the members of a special survey of those who did not initially respond to the mailed questionnaires constitute the Follow-Up Sample. The three follow-ups of the four grades have generated 12 such follow-up samples. In 1975, a special-purpose follow-up study, involving intensive personal interviews, was based on a specially defined subsample of 1,000 women and men from the 15-Year-Old Sample. These individuals, who were then age 30, constitute the 30-Year-Old Sample. Further information about this sample and study is presented in An Empirical Study to Aid in Formulating Educational Goals (Flanagan & Russ-Eft, 1975).

Selection of the National Probability Sample (1960)

Since senior high schools were the primary sampling unit used in 1960, the selection of students for the national probability sample must be described in terms of the selection of the schools. Once a school was selected, an attempt was made to test every student in the school and, in the case of communities with junior high schools, to test the 9th graders in the associated junior high school(s).

The selection of schools was based on information from several sources. The primary source was file of data on public senior high schools provided by the United States Office of Education (USOE).

USOE also provided a supplementary list of schools not in the file and of private and parochial high schools. ("Parochial schools" here refers to parochial and diocesan schools run by the Roman Catholic Church. Schools affiliated with other churches were classified with other nonpublic schools as "private.") The Internal Revenue Service also provided a list of schools, which contained few not on the lists provided by USOE.

The primary stratification variables used in selecting the schools were type of school (public, parochial, or private) and geographic location. In stratifying on location, each of the five cities with populations in excess of 1.5 million (New York, Los Angeles, Chicago, Philadelphia, and Detroit) was treated as a separate stratum. The 50 states, excluding these cities, were then grouped into the 9 USOE regions.

For public schools, two additional stratification variables were used, school size and retention ratio. Four levels of school size were defined by the number of 12th grade students. The retention ratio was defined as the number of graduates divided by the number of 10th graders. Schools were divided into high and low retention ratio groups (and in a few cases a third group where the retention ratio was unknown) within each region and school size.

In general a sampling ratio of 1 in 20 was used in selecting the senior high schools, but in order to increase the number of very large schools, a sampling ratio of 1 in 13 was used for public high schools with 400 or more seniors. Correspondingly, in order to avoid spending an excessive effort covering the large number of very small schools, public high schools with fewer than 25 seniors were sampled with a ratio of 1 in 50. In all, 1,063 senior high schools were selected and 987 of these (93%) actually participated in the testing. Table 2 shows the composition of the final sample of schools by geographic region and school type.

School weights have been developed that compensate for the differential sampling ratios and for the refusal of a few schools to participate. Each school has a weight that indicates the number of schools in the population for which it stands. If there were 100 schools in one of the sampling cells and 5 were selected, then each school selected would stand for 20 schools and would thus have a weight of 20. If only 4 of these 5 schools agreed to participate, then the 100 schools in the population would be represented by 4 schools in the study so that each school would have a weight of 25. Further information on the selection and weighting of the 1960 probability sample may be found in The Project TALENT Data Bank Handbook (Wise, McLaughlin, & Steel, 1979).

Table 2

High Schools Participating in Project TALENT
by Region and School Type

	School Type						TOTAL
	Public High Schools				Parochial Schools	Private Schools	
	1-24 Seniors	25-99 Seniors	100-399 Seniors	400+ Seniors			
Regions:							
New England	4	16	12	1	13	9	55
Midwest	3	42	33	3	23	8	112
Great Lakes	15	86	30	8	21	6	166
Plains	34	63	12	3	16	5	133
Southeast	33	142	37	4	13	8	237
Southwest	24	35	10	4	7	4	84
Rockies	8	11	3	2	2	2	28
Far West	3	14	12	3	11	5	48
Non-Contiguous	0	0	2	0	0	1	3
Five Largest Cities:							
New York	0	3	32	50	3	2	90
Philadelphia	0	0	1	1	0	0	2
Detroit	0	0	0	1	3	0	4
Chicago	0	1	9	10	1	1	22
Los Angeles	0	0	1	1	1	0	3
TOTAL	124	413	194	91	114	51	987

Selection of the 15-Year-Old Sample

Most 15-year-olds were in high school, but a few had not yet reached 9th grade, a few had dropped out of school altogether, and a very few had already completed high school and gone on to college. In addition to the 15-year-olds in the high school probability sample, the probability sample of 15-year olds also includes all 15-year-olds not in high school in the areas served by a subsample of 10 percent of the senior high schools. This 10 percent subsample was selected systematically with respect to the school location and size. A more complete description of this subsample is given in Studies of a Complete Age Group: Age 15 (Shaycoft, Dailey, Orr, Neyman, & Sherman, 1963).

The Retest Sample (1963)

In 1963, the 12th graders in 118 of the TALENT high schools participated in a special retest study. Over 7,000 of these students (roughly three-fourths) had also participated in the 1960 testing. The participating schools included 17 vocational and 101 academic public high schools selected from the original sample of schools on the basis of a taxonomy code that classified schools by region, school type, and type of community. The design and findings of the 1963 retest study are described in detail in The High School Years: Growth in Cognitive Skills (Shaycoft, 1967).

The Follow-Up Samples

Each of the grade cohorts in the 1960 probability sample has been followed up 1, 5, and 11 years after the date of expected high school graduation, as shown in Table 1 above. In each case all the students tested in 1960 were asked to fill out and return questionnaires mailed to their last known addresses. The follow-ups for each grade generally included four mailings spaced about one month apart with a reminder card sent between the first and second waves.

Not surprisingly, given the size of the initial samples and the funds that could be expended in locating them, response to these follow-ups decreased over the 15 years between 1960 and the completion of the 11-year follow-up in 1975. Rates of response to the mail questionnaires ranged from 61.9 percent for the 12th grade, 1-year follow-up, to 19.6 percent for the 9th grade, 11-year follow-up. In order to overcome potential nonresponse bias, a special sample of about 2,500 of those who did not return questionnaires was drawn at the end of the follow-up of each grade cohort. An intensive effort was made to locate members of this special nonrespondent sample and to obtain follow-up data from them, generally through a telephone interview or through a personal interview by a regional follow-up coordinator. For the 11-year follow-up of the 9th grade cohort, the special sample was expanded to include

1,000 additional cases from schools with a high proportion of minority students. Additional information on the procedures used in the follow-up surveys may be found in Methodology of the Project TALENT 11-Year Follow-Up Study (Rossi, Wise, Williams, & Carrel, 1976). The procedures that have been developed to correct for possible bias due to sample attrition are discussed in a subsequent section on Technical Issues in Maintaining the Data Base.

The TALENT Test and Survey Instruments

The TALENT data base contains literally thousands of pieces of information about the backgrounds, abilities, interests, goals, opinions, educational experiences, and career outcomes of each TALENT participant. The four main sets of instruments used in collecting these data are as follows:

1. 1960 Student Data

Ability and Information Tests
Student Activities Inventory
Interest Inventory
Student Information Blank

2. 1963 Retest Data

7 overlapping batteries, each including roughly half of the 1960 instruments

3. Follow-Up Survey Data

12 separate but partially overlapping survey instruments
(3 follow-ups x 4 grades)

4. School Data

General School Characteristics Questionnaire
Guidance Program Questionnaire

The 1960 Student Data

The Project TALENT battery developed for the 1960 testing included inventories and cognitive tests in the following areas: cognitive abilities and knowledge; dispositional traits; interests; activities and background characteristics. The total package was designed to be completed in two school days of testing. Detailed information on the development of these tests and inventories, including sample items, may be found in Design for a Study of American Youth (Flanagan et al., 1962).

Ability and Information Tests. The major part of the Project TALENT battery consisted of a series of tests assessing the participants' cognitive abilities and knowledge. A 395-item information test was developed to assess the students' knowledge in both academic and non-academic areas ranging from science and literature to foods, sports, and practical knowledge. Separate cognitive tests were designed to measure specific abilities ranging from memory, visualization, and clerical aptitude to English and mathematics achievement. Table 3 lists the number of items and the testing time for the specific ability and information instruments. Table 4 lists the components of the information test.

Student Activities Inventory. This instrument was designed to assess students' consistent dispositional traits. It consisted of 150 self-descriptive terms dealing with aspects of the student's disposition. Previous work in personality measurement plus careful theoretical analysis provided a foundation for defining a number of narrow, relatively homogeneous aspects of the personality of high school students. Each trait was defined by grouping adjectives that describe similar types of behavior. For example, the trait "impulsiveness" is typified by adjectives such as "hasty," "impulsive," and "rash." Table 5 lists the 10 scales derived from this inventory.

The Interest Inventory. The TALENT Interest Inventory was developed for the purpose of surveying the interests of high school youth and obtaining data for research on subsequent educational and vocational choices. The Interest Inventory was made up of 205 items dealing with 122 occupations and 83 activities. Students were asked to respond in terms of how well they would like or dislike to work or activity, disregarding educational requirements, salary, social standing, or other factors. The individual items in this instrument have been combined into 17 interest scales which are also listed in Table 5.

The Student Information Blank. This instrument was designed to yield background data that had been hypothesized or found to relate to three key aspects of young people's lives: decisions about whether or not to stay in high school until graduation, decisions about whether or not to enter and complete college, and decisions about which occupation to pursue. Based on previous research into the background factors influencing each type of decision, 394 multiple-choice questions were developed about students' backgrounds and plans, including:

- o Family Background (parents' education and occupation, economic status, number of siblings and their educational experience),
- o Health (general health at different times, specific illnesses and disabilities),
- o Current Activities (school courses and grades, extracurricular activities, hobbies, reading and study habits, social

Table 3
Characteristics of the Ability and Information Tests

Instrument	No. of Items	Testing Time in Minutes
Information Test (see Table 5)	395	125
Language Aptitude and Ability Tests		
Memory for Sentences	16	10
Memory for Words	24	4
Disguised Words	30	3
English:		
English Total	113	52 ^a
Spelling	16	12 ^a
Capitalization	33	6 ^a
Punctuation	27	14 ^a
English Usage	25	10 ^a
Effective Expression	12	10
Word Functions in Sentences	24	15
Reading Comprehension	48	30
Complex Intellectual Aptitude Tests		
Creativity	20	20
Mechanical Reasoning	20	11
Abstract Reasoning	15	11
Visualization Tests		
Visualization in Two Dimensions	24	4
Visualization in Three Dimensions	16	9
Mathematics Tests		
Mathematics:		
Arithmetic Reasoning	16	12 ^b
Introductory	24	24 ^b
Advanced	14	16 ^b
Clerical and Perceptual Aptitude Tests		
Arithmetic Computation	72	9
Table Reading	72	3
Clerical Checking	74	3
Object Inspection	40	3

^aThe five sections of the English test were times separately and students were not allowed to work in sections other than the one being times.

^bThe three sections of the Mathematics test were times together with students being allowed to go back to prior sections or work ahead.

Table 4
Information Test Scales

	Number of Items
Part I (90 minutes, 15 derived scales)	
Screening	12
Vocabulary (Part I)	21
Literature	24
Music	13
Social Studies	24
Mathematics	23
Physical Science	18
Biological Science	11
Scientific Attitude	10
Aeronautics and Space	10
Electricity and Electronics	20
Mechanics	19
Farming	12
Home Economics	21
Sports	14
Part I Total	252
Part II (35 minutes, 24 derived scales)	
Art	12
Law	9
Medicine/Health	9
Engineering	6
Architecture	6
Journalism	3
Foreign Travel	5
Military	7
Accounting, business, sales	10
Practical knowledge	4
Clerical	3
Bible	15
Colors	3
Etiquette	2
Hunting	5
Fishing	5
Outdoor activities (other)	9
Photography	3
Games (sedentary)	5
Theater and Ballet	8
Foods	4
Miscellaneous	10
Vocabulary (Part II)	9
Part II Total	143
Combined Scales	
Vocabulary (Parts I and II)	10
Total Parts I and II	395

Table

Scales Derived from the Student Activities Inventory,
the Interest Inventory, and the Student Information Blank

<u>Instrument/Scale</u>	<u>No. of Items</u>	<u>Instrument/Scale</u>	<u>No. of Items</u>
Student Activities Inventory	150	Interest Inventory (continued)	
Sociability	12	Computation	10
Social Sensitivity	9	Office Work	7
Impulsiveness	9	Mechanical-Technical	15
Vigor	7	Skilled Trades	18
Calmness	9	Farming	7
Tidiness	11	Labor	10
Culture	10	Other	32
Leadership	5	Student Information Blank	394
Self-Confidence	12	High school courses taken	6
Mature Personality	24	High school grades	7
Conventionality	4	Guidance received in high school	10
Theoreticity	8	Guidance received elsewhere	6
Group-Centeredness	6	Study habits and attitudes	16
Interest inventory	205	Self-perception of writing skills	4
Physical Sciences	16	Self-perception of reading skills	6
Biological Science,		Amount of extracurricular reading	5
Medicine	8	Variety of extracurricular group	
Public Service	11	activities	10
Literary-Linguistic	16	Degree of participation in extra-	
Social Service	12	curricular activities	10
Artistic	7	Variety of hobbies	15
Musical	5	Degree of activity in hobbies	15
Sports	8	Participation in sports	4
Outdoor Recreation	3	Leadership roles	3
Business Management	14	Social life	5
Sales	6	Work activities (chores and jobs)	5

Table 6

Data Obtained in the Project TALENT Follow-Up Surveys

ONE-YEAR FOLLOW-UP SURVEY (1961-1964)		
EDUCATIONAL EXPERIENCE Kind and amount attained, grades, courses, plans for future, reasons for stopping, financial costs, perceived benefits	CAREER Work experience, occupation, income Job satisfaction, future plans	PERSONAL Marital status Health Military status and experience Geographical mobility Leisure and social activities Decisions regretted
FIVE-YEAR FOLLOW-UP SURVEY (1965-1968)		
EDUCATIONAL EXPERIENCE Kind and amount attained, degrees, years attended, grades, reasons for dropping out, perceived costs and benefits, financial costs	CAREER Work experience, occupation, income Job satisfaction, future plans Training received Reasons for stopping work Sources of income	PERSONAL Race, religion Marital status and history Spouse's education Number of children Health: own and parents' Geographical mobility Military status and experience Decisions regretted
ELEVEN-YEAR FOLLOW-UP SURVEY (1971-1974)		
EDUCATIONAL EXPERIENCE Kind and amount, degrees, occupational training in high school, noncollege training, colleges and dates attended, reasons for stopping, financial support during college, perceived value of education	CAREER Work experience, occupation, income Job satisfaction (15-25 dimensions), evaluation of job Career plans, reasons for stopping work	PERSONAL Race Marital status and history Spouse's education and career Childbearing history Ideal and expected family size Quality of life (15 areas) Geographical mobility Health Leisure and civic activities Military status and experience Discrimination experienced Political behavior

Note: Separate follow-up surveys were conducted for each of the four high school classes (one, five, and eleven years after their respective graduation). Some of the follow-up variables are available for one or two classes only.

activities). and

- o Plans (college, career, military service, marriage and family).

In addition to the individual item data, several composite indices have been developed for variables of frequent interest. These composite measures are also listed in Table 5. In addition to the measures listed in Table 5, a measure of individuals' background socioeconomic status was constructed because of the importance of this variable as a control measure. It is defined as a weighted combination of items on mother's and father's level of education, father's occupation, estimates of income and home value or rental, and the number of appliances and other articles in the home.

The 1963 Retest Data

With a few exceptions, the entire TALENT battery was readministered as part of the 1963 retesting. These exceptions were the Preference Test, the Memory for Sentences Test, the Student Activities Inventory, and the themes, all of which were omitted in the interest of reducing the length of the battery. Further, the Student Information Blank was abridged somewhat. In order to reduce further the required testing time, six overlapping batteries were designed, each containing approximately half of the 1960 battery. A seventh battery of TALENT tests was designed specifically for vocational schools. Each battery was completed by students from one-seventh of the schools in the retest sample.

The resulting body of data, when combined with these students' scores from 1960, gives information about changes in students' abilities, achievement levels, interests, plans, and actions during the high school years. Further, these data provide a basis for investigating factors that promote or inhibit such changes. For further information on the design and results of this study, see The High School Years: Growth in Cognitive Skills (Shaycoft, 1967).

Follow-Up Survey Data

Data for the three follow-up surveys were obtained from mailed questionnaires. In the case of the special surveys of nonrespondents to the mail survey, the questionnaires were administered via telephone or, occasionally in-person interviews. To facilitate processing of the data obtained, most of the questions were in multiple-choice format. A few questions, however, such as the title of the person's current job or the name of the college the person attended (if any) were necessarily open-ended and required coding by project staff after the completed questionnaires had been returned.

The questionnaires for the three follow-up waves varied in length and scope, ranging from 30 items for the 9th grade 1-year follow-up to 67 items for the 9th grade 11-year follow-up. However, they all included a number of items relating to individuals' educational, work, and family-related experiences. In addition, questions relating to other aspects of individuals' lives (e.g., geographic mobility, health, or political behavior) were included in some of the questionnaires (i.e., for a particular cohort or a particular follow-up wave) not necessarily all of them. The major kinds of data obtained are summarized in Table 6 and are briefly described below.

Education. In all three follow-ups large portions of the questionnaires for each cohort were devoted to assessing individuals' educational experiences. The questions covered school attendance, what kind of postsecondary institution was attended, and what degrees were obtained as well as sources of support, programs or majors, grades obtained, years attended, reasons for dropping out where appropriate, and evaluations of degree to which school had prepared the respondent for the work world.

Occupations. Several questions dealing with experiences in the work world were also included in each follow-up survey. These included the name of the job held at the time of each follow-up and the occupation planned for one's life work. Other questions asked of some, but not all, of the cohorts, covered particular aspects of the job such as rate of pay, hours worked, instances and reasons for unemployment, number of jobs held, and in the 11-year follow-up, several aspects of job satisfaction. The job satisfaction measures include both perceptions of what is important for job satisfaction (e.g., long vacations, good pay, or worthwhile work) and ratings of satisfaction with one's own job on those dimensions.

The Project TALENT coding staff assigned 4-digit occupational codes to each individual's current job based on the individual's free responses to questions of "what is your job called?" and "what do you do on it?" Occupational codes were similarly assigned for the respondent's career plan at each time. These codes were translated by computer into about 250 3-digit job codes, and these, in turn, have been grouped into 14 career groups. From the 9th grade 11-year follow-up, job codes developed by the U.S. Bureau of the Census were also used. This dual coding has provided the basis for mappings of the TALENT codes onto Census codes.

Marriage and family. The follow-up questionnaires have obtained information on current marital status, number of marriages, age at marriage, and spouse's age, level of education, and occupation. With the exception of marital status, these questions were limited primarily to the 11-year follow-up questionnaire. Questions on children were also included primarily in the 11-year follow-up. Separate questions

covered the number of children in the household and their ages; other questions dealt with expectations and desires for future children.

Other questions. Each questionnaire has included a number of unique questions on other topics. Questions on military experience, race, religion, aspects of health, attributes of one's community of residence, smoking, parental mortality, driving, Peace Corps and other volunteer work, typical time allocations for various activities, voting, and several questions asking for evaluations of one's experiences and actions (e.g., "what decisions are you sorry about?") have been included in one or more of the follow-up surveys.

The simple distributions of answers to follow-up questions, as well as other information, are contained in Project TALENT One-Year Follow-Up Studies (Flanagan, Cooley, Lohnes, Schoenfeldt, Holdeman, Combs, & Becker, 1966), The American Citizen: 11 Years After High School (Wilson & Wise, 1975), and The American Citizen: 11 Years After High School, Volume II (Wise, McLaughlin, & Gilmartin, 1977).

School Data

Considerable data were also collected in 1960 on the characteristics of the schools that TALENT participants attended. Two questionnaires were used in collecting the basic school data, one focusing on general school characteristics and the other on guidance program characteristics.

General School Characteristics. This questionnaire was designed to provide information about those characteristics of the school that might make a difference in the education of its students as compared with students at other schools. Each school principal completed a questionnaire for his/her school.

The General School Characteristics Questionnaire consisted of six parts:

1. School Policies, Practices, and Plant--the type of school, special classes, grading and advancement policies, physical condition of the school, and average class size.
2. The Size, Training, and Characteristics of the Teaching Staff.
3. The Students--enrollment, dropouts, percentage of graduates going on to college, remedial work, and the like.
4. Characteristics of the Community--PTA activity, type of area served by the school, tax rate, per-pupil expenditures, and community facilities.
5. The Principal--his/her age, experience, and training.

6. High School Courses--the courses offered in grades 9 to 12 in the school from a list of courses, grouped by broad subject matter areas.

Guidance Program/Staff Characteristics. To provide a basis for investigations of how guidance and counseling affect students, a special survey of school guidance programs was included as part of the 1960 data collection. The topics covered include:

Scope of the guidance program. Does the school have a formal guidance program? If yes, how large is the guidance staff, how adequate are referral facilities in the community?

Types of aids and guidance provided. What kinds of problems are brought to counselors by students? When and how does the counselor schedule conferences with students and parents? Other questions asked about special courses related to guidance (such as occupations), and about the means used for imparting educational and vocational information.

Past growth. To what extent has the guidance program expanded in recent years?

Testing. Does the school use nationally standardized tests and inventories--intelligence tests, aptitude batteries, interest inventories, and adjustment inventories? Are the tests administered on an individual basis or to entire classes? Are test results used to provide information to students and parents on school progress and on educational potential; to provide information to colleges and scholarship agencies; and/or to evaluate the school's progress in teaching subject matter?

Findings from Project TALENT

Much as been learned through Project TALENT regarding the characteristics, experiences, and outcomes of a very large group of American women and men. Unfortunately, it is not possible within the space of a few pages to summarize adequately the results of nearly 20 years worth of data collection and analysis. In this section, then, we will briefly describe the major survey reports that have been prepared. For a more comprehensive report of Project TALENT studies and results, the reader is referred to Publications Based on Project TALENT Data: An Annotated Bibliography (Campbell, 1979).

Major Survey Reports

The results of the initial analyses of the 1950 data were summarized in

two major volumes. Studies of the American High School (Flanagan, Dailey, Shaycoft, Orr, & Goldberg, 1962) included a description of school characteristics, the development of a taxonomy of high schools for use in TALENT studies, and reports on the status of guidance and testing programs then extant. It also discussed the characteristics of high schools which were found to be important or unimportant for educating students. In 1964, a massive report was published entitled The American High School Student (Flanagan, Davis, Dailey, Shaycoft, Orr, Goldberg, & Neyman, 1964). It focused primarily on the 1960 study, but also included results from the first (one-year) post-high school survey of 12th graders. It included the best and most comprehensive nationwide data then (and perhaps now) available on sex differences in test performance, curriculum, and life plans. These and other studies have been widely quoted in educational and guidance texts and have had a significant impact. The importance of counseling and individualized instruction was particularly underscored. The latter emphasis contributed to the development of individualized instructional programs in the '60s such as PLAN--Program for Learning in Accordance with Needs (Flanagan, Shanner, Brudner, & Marker, 1975).

Three other reports based on special subsamples of the TALENT participants should also be mentioned here. Studies of a Complete Age Group (Shaycoft et al., 1963) reports findings based on the sample of 15-year-olds that cuts across grades. This report focuses on dropouts and on age-grade relationships. The High School Years: Growth in Cognitive Skills (Shaycoft, 1967) reports the results of the 1963 retest study. This report includes findings related to the tests and findings relating to students gains and to school and other effects on student gains. An Empirical Study to Aid in Formulating Educational Goals (Flanagan & Russ-Eft, 1975) reports the basic results of the interviews of 1,000 participants at age 30.

The overall findings from the one-year follow-up studies were published in Project TALENT One-Year Follow-Up Studies (Flanagan et al., 1966). This report discusses characteristics of students pursuing post-secondary education and students who had entered the labor force. Sections are also included on the stability of career plans and predicting career group changes. The report concludes with a discussion of the implications of the findings for career guidance.

Basic findings from the five-year follow-up surveys are reported in Five Years After High School (Flanagan, Shaycoft, Richards, & Claudy, 1971). This report focuses on educational outcomes, including choice of an success in college majors, the stability and prediction of career plans, and factors relating to salary outcomes. Other, more specific analyses are also included. The 11-year survey results are summarized in The American Citizen: 11 Years After High School (Wilson & Wise, 1975) and The American Citizen: 11 Years After High School, Volume II (Wise, McLaughlin, & Gilmartin, 1977). These reports contain the response distributions to each of the survey items, their relationship to

high school abilities and socioeconomic status, and highlights of the major results. The detailed results of the 11-year follow-up surveys are presented in the individual reports of specific studies using the data.

What Lies Ahead

The original design for TALENT called for a fourth follow-up of the participants, to occur approximately 20 years following their class' graduation from high school, to obtain data on their educational, career, and personal statuses and activities. These data would be used to determine how and how well individuals' educational experiences in adolescence and early adulthood had prepared them to assume and carry out their occupational and other adult roles and responsibilities.

However, the data that would be provided by such a survey would have considerable value beyond the assessment of long-term educational benefits. In conjunction with the earlier TALENT data, they would provide a basis for investigating many important aspects of adult development and aging, such as the nature and extent of change in individuals' cognitive and social-psychological characteristics during the early adult years and the growing interrelationships between education and work (or education, work, and family responsibilities) over the life course. Unfortunately, the project's sponsor, the National Institute of Education (NIE) was unable to support this final survey and encouraged project staff to seek support from other agencies. It is thus not clear at this time when a further follow-up will be carried out.

This gives rise to an issue likely to be of concern to others involved in conducting longitudinal studies. The value of longitudinal data bases increases exponentially as additional data are obtained, even after many of the initial project objectives have been achieved. As we have found with Project TALENT, with time the value of a data base may transcend the domain of interest of its initial sponsor, or of any single potential sponsoring agency. How then can the scientific community ensure that the full potential of these data resources is realized and maintained?

While we have no quick answers to this question, we are working with other longitudinal researchers to explore this issue. Hopefully this volume on longitudinal research in the United States will serve to stimulate awareness of the value of longitudinal data and discussion about this and other issues relating to the continued development, availability, and use of longitudinal data sets.

References

- Abeles, R.P., & Wise, L.L. Coping with attrition in a longitudinal study: The case of Project TALENT. Journal of Economics and Business, 1979-1980, 32, 100-181.
- Campbell, E.A. Publications based on Project TALENT data: An annotated bibliography. Palo Alto, California: American Institutes for Research, 1979.
- Card, J.J. Consequences of adolescent childbearing for the young parent's future personal and professional life. Final Report. Palo Alto, California: American Institutes for Research, 1977.
- Card, J.J. Long-term consequences for children born to adolescent parents. Palo Alto, California: American Institutes for Research, 1978.
- Card, J.J., Steel, L., & Abeles, R.P. Sex differences in realization of individual potential for achievement. Journal of Vocational Behavior, in press.
- Card, J. J. & Wise, L.L. Teenage mothers and teenage fathers: The impact of early childbearing on the parents' personal and professional lives. Family Planning Perspectives, 1978, 10 (4), 199-205.
- Claudy, J.G., Farrell, W.S. Jr. & Dayton, C.L. The consequences of being an only child: An analysis of Project TALENT data. Final Report. Palo Alto, California: American Institutes for Research, 1979.
- Claudy, J.G., Gross, D.E. & Strause, R.D. Two population studies: I. Family size, birth order, and characteristics of young adults. II. A study of married couples in Knox County, Tennessee. Final Report. Palo Alto, California: American Institutes for Research, 1974.
- Cureton, L.W. Arrests for drunkenness of young adults and their parents and adolescent behavior problems. Palo Alto, California: American Institutes for Research, 1973.
- Flanagan, J.C. The first 15 years of Project TALENT: Implications for career guidance. Vocational Guidance Quarterly, September, 1973, 8-14.
- Flanagan, J.C. Changes in school level of achievement: Project TALENT ten- and fifteen-year retests. Educational Researcher, 5 (8), 1976, 9-12.
- Flanagan, J.C. Planning career goals based on data from Project TALENT. Vocational Guidance Quarterly, 1977a, 25, 270-273.

Flanagan, J.C. A program for career and life planning. Education, 1977b, 97(4), 318-325.

Flanagan, J.C., Cooley, W.W., Lohnes, P.R., Schoenfeldt, L.F., Holdeman, R.W., Combs, J., & Becker, S. Project TALENT one-year follow-up studies. American Institutes for Research, 1966.

Flanagan, J.C., Dailey, J.T., Shaycoft, M.F., Gorham, W.A., Orr, D.B., & Goldberg, I. Design for a study of American youth. Boston: Houghton Mifflin, 1962.

Flanagan, J.C., Dailey, J.T., Shaycoft, M.F., Orr, D.B., & Goldberg, I. Studies of the American high school. Pittsburgh: American Institutes for Research, 1962.

Flanagan, J.C., Davis, F.B., Dailey, J.T., Shaycoft, M.F., Orr, D.B., Goldberg, I., & Neyman, C.A., Jr. The American high school student. American Institutes for Research, 1964.

Flanagan, J.C. & Jung, S.M. Progress in education: A sample survey (1960-1970). Final Report. Palo Alto, California: American Institutes for Research, 1971.

Flanagan, J.C., & Russ-Eft, D. An empirical study to aid in formulation educational goals. Final Report. Palo Alto, California: American Institutes for Research, 1975.

Flanagan, J.C., Shanner, W.M., Brudner, H.J., & Marker, R.W. An individualized instructional system: PLAN. In Talmadge, H. (Ed.), Systems of individualized education. Berkeley, California: McCutchen, 1975.

Flanagan, J.C., Shaycoft, M.F., Richards, J.M., Jr., & Claudy, J.G. Five years after high school. Palo Alto, California: American Institutes for Research, 1971.

Flanagan, J.C., Tiedeman, D.V., Willis, M.B., & McLaughlin, D.H. The career data book: Results from Project TALENT's five-year follow-up study. Palo Alto, California: American Institutes for Research, 1973.

Folger, J.K., Astin, H.S., & Bayer, A.E. Human resources and higher education. New York: Russell Sage Foundation, 1970.

Gilmartin, K.J., McLaughlin, D.H., Wise, L.L., & Rossi, R.J. Development of scientific careers: The high school years. Final Report. Palo Alto, California: American Institutes for Research, 1976.

Inbar, M. The vulnerable age phenomenon. New York: Russell Sage Foundation, 1976.

Levinson, Perry. The next generation: A study of children in AFDC families. Welfare in Review, 1969, 7(2), 1-9.

McLaughlin, D.H. Career changers: A Project TALENT report. Palo Alto, California: American Institutes for Research, 1975.

McLaughlin, D.H., Fulscher, G.V., & Yen, W.M. Project TALENT's special sample: Is it necessary? Palo Alto, California: American Institutes for Research, 1974.

Rossi, R.J., Bartlett, W.B., Campbell, E.A., Wise, L.L., & McLaughlin, D.H. Using the TALENT profiles in counseling: A supplement to the career data book. Palo Alto, California: American Institutes for Research, 1975.

Rossi, R.J., Wise, L.L., Williams, K.L., & Carrel, K.S. Methodology of the Project TALENT 11-year follow-up study. Palo Alto, California: American Institutes for Research, 1976.

Sewell, W.H. Inequality of opportunity for higher education. American Sociological Review, 1971, 36(5), 793-809.

Shaycoft, M.F. The high school years: Growth in cognitive skills. American Institutes for Research, 1967.

Shaycoft, M.F. Project TALENT: A short history of a long project. Palo Alto, California: American Institutes for Research, 1977.

Shaycoft, M.F., Dailey, J.T., Orr, D.B., Neyman, C.A., Jr., & Sherman, S.E. Studies of a complete age group: Age 15. American Institutes for Research, 1963.

Steel, L., Abeles, R.P., & Card, J.J. Sex differences in the patterning of adult roles as a determinant of sex differences in occupational achievement. Manuscript submitted for publication, 1979.

Steel, L. & Wise, L.L. Designing a study of adult accomplishment and life quality: A report on Phase I of the Project TALENT 17-year follow-up survey and TALENT data bank maintenance activities. Palo Alto, California: American Institutes for Research, 1977.

Steel, L., & Wise, L.L. Origins of sex differences in high school mathematics and participation. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April, 1979.

Wilson, S.R., & Wise, L.L. The American citizen: 11 years after high school. Palo Alto, California: American Institute for Research, 1975.

Wilson, S.R., Wise, L.L., & Stancavage, F. Career development of the Project TALENT physicians. Final Report. Palo Alto, California: American Institutes for Research, 1980.

Wise, L.L. The fight against attrition in longitudinal research. Paper presented at the meetings of the American Educational Research Association, New York, April 1977.

Wise, L.L. Long-term consequences of sex differences in high school mathematics education. Paper presented at the annual convention of the American Educational Research Association, San Francisco, April, 1979.

Wise, L.L., McLaughlin, D.H., & Gilmartin, K.J. The American citizen: 11 years after high school, volume II. Palo Alto, California: American Institutes for Research, 1977.

Wise, L.L., McLaughlin, D.H., & Steel, L. The Project TALENT data bank handbook. Palo Alto, California: American Institutes for Research, 1977 (revised 1979).

Wise, L.L., & Steel, L. The effects of school quality on students' knowledge and skills. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

Yen, W.M. & McLaughlin, D.H. Post high school education and career development. Palo Alto, California: American Institutes for Research, 1974.

High Risk Early Signs for Delinquency and Related Behavioral Difficulties:

A Longitudinal Study*

George Spivack, Ph.D.

*Prepared under Grant Number 76-JN-99-0024 from the National Institute for Juvenile Justice and Delinquency Prevention, Office of Juvenile Justice and Delinquency Prevention, Law Enforcement Assistance Administration, U.S. Department of Justice. Points of view or opinions in this document are those of the author and do not necessarily represent the official position or policies of the U.S. Department of Justice.

This project could not have been conducted without the support and close collaborative efforts of the Philadelphia Public Schools, the Philadelphia Police Department, the Philadelphia Office of Mental Health and Mental Retardation, and the Archdiocese Schools of Philadelphia.

High Risk Early Signs for Delinquency and Related Behavioral Difficulties:
A Longitudinal Study

George Spivack, Ph.D.

General Purpose

The general purpose of this longitudinal study has been to discover early signs in school (from kindergarten on) which might indicate that a youngster, already part of a high risk group, sociologically speaking, is especially high risk insofar as subsequent delinquent behaviors are concerned in the general school environment, in the classroom specifically, and in the community. A closely associated secondary purpose has been to explore early high risk signs that may predict a variety of behavioral disturbances and academic failures which usually accompany delinquent behaviors.

A further purpose is to study the school history of children who become delinquent, in order to uncover patterns of behavior and achievement through time which would predict delinquency, and experiences in school that might affect the likelihood of delinquency.

The long-term objective is to be able to apply primary and/or secondary preventive intervention in the school on the basis of knowledge of these early signs of risk. If such study revealed early behavioral indicators from the measures used, intervention strategies already available could be brought to bear in preventive efforts. It would be possible for any school interested in preventive programming to identify early in the school history a subgroup of youngsters that warrant attention on the basis of objective information.

This information would derive in part from teachers' natural contacts with children, requiring no special testing or observational procedures. Finally, identification would be largely in terms of overt behavior patterns for which intervention programming could be applied.

Methodology

Data collection for this project has taken place during two periods of time: 1968-1972, and 1975 to the present. During the second period, certain information on the sample studied covering the intervening years 1972-1975 was acquired from existing records. This pattern of data collection was dictated by the funds available for the project from different sources over the total period of years.

Initial data collection methodology (1968-1972)

Initiating the data collection process throughout the first four years of the overall longitudinal study (1968-1972; kindergarten-3rd grade) required extensive groundwork, including clearances from the Superintendent's office of the School District of Philadelphia, conferences with all district superintendents and principals, periodic feedback sessions in order to review project progress and requests for new assistance, meetings with teachers, and a laborious process of tracking children in a highly mobile urban area.

Initial data collection began in October of 1968 in 29 schools from four center city school districts. These districts were selected because they serve children and families within the catchment area of Hahnemann Community Mental Health/Mental Retardation Center. This area is characterized by all of the usual signs of poverty and underprivilege found in large urban centers. Children were selected randomly, with the following constraints: there would be half boys and half girls, half would be in A.M. and half in P.M. kindergarten classes, and half or each of these would have had pre-school (Head Start) experience. It was also planned that no teacher would have more than 12 children to rate.

Having met the above criteria, all 56 kindergarten teachers from the 29 schools agreed to participate. Meeting in small groups, teachers were told this was a longitudinal study of children with the purpose of studying classroom behavior patterns and how these would relate to subsequent learning and adjustment. They were told that the long-range goal was to identify high risk behaviors, perhaps as early as in kindergarten, that call for preventive measures in the classroom. All teachers saw these purposes as reasonable, and seemed eager to participate. A brief 30 minute training period in how to use the Devereux Elementary School Behavior Rating Scale (DESB) (see Spivack and Swift, 1967), followed. After all questions were answered, each teacher was given his or her list of students to rate and asked to return completed ratings to the principal's office within two weeks.

In May of 1969 (seven months later), each teacher was again contacted for a second rating of each youngster. By this time 126 youngsters (19% of original sample) were no longer in the same kindergarten and could not be rated. Each of the remaining 533 children were rated after a brief "refresher" training meeting with the teachers. At this time, reading readiness scores were also available on each child, as were data on number of absences and

whether the child had been transferred. Early in the fall of 1969 (beginning first grade) the tracking of "lost" cases began. Each principal of the original (kindergarten) school was supplied a list of children rated, and information was sought as to each child's current first-grade whereabouts. Beginning with this query, and after numerous phone calls and correspondence, it was discovered that the children originally in 29 schools were now dispersed among over 60 schools, and in the classrooms of over 100 first grade teachers. Despite this tracking effort, and in part due to parochial school transfers (N = 35), there was continued attrition down to a sample of 443 children. The first grade teachers were met with, told the purpose of the study, and trained in the rating process.

Some teachers were unwilling to participate, and this contributed to loss of ratings. Considering the unexpected teacher resistance, funds were sought and made available by the Philadelphia Board of Education to pay teachers for subsequent ratings. Payment was made for the next ratings, in May of 1970 (end of first grade). At this point 428 children were rated in 52 schools. This sample constituted 65 percent of the original sample rated at the beginning of kindergarten, 19 months earlier. At this point, other data became available: absences and transfer information, whether or not the child was part of a "follow-through" educational program, reading and arithmetic achievement marks from first grade, and whether the child required psychological testing for any reason.

One year later (Spring of 1971) the same process was repeated to locate as many children in the sample as possible, obtain classroom behavior ratings, and collect all other records information available. At this point the same categories of data were extracted, and in addition the results of the Stanford Academic Achievement testing.

During the Fall of 1971, when the sample was entering third grade, a complete tracking search was made for all children initially involved in the study. This search was facilitated by a new computer system operated by the Division of Research of the Philadelphia Schools. With the assistance of the computer, and meetings with district superintendents, principals, and teachers, 611 (93%) of the original sample was successfully tracked. Sixty-five (10%) of the original sample were found to be enrolled in parochial schools, and 32 (5%) had left the city area. Seven percent were "lost" to the study. All remaining children, totaling 514, were rated during May of 1972 (end of third grade). Ratings obtained from 216 teachers in 91 schools located throughout the city. At this time, reading and arithmetic scores on the Iowa Test of Basic Skills were obtained, and as in previous years report cards supplied information about reading and arithmetic book level classroom achievement.

The data on each child located during this four year period, between the ages of five and eight years, may be summarized as follows:

- (1) Sex and age when child entered kindergarten
- (2) Pre-school (Head-Start) experience
- (3) Follow-through special placement following Head-Start
- (4) Racial balance of every school attended
- (5) The average academic test level of every school attended
- (6) Frequency of absences each year
- (7) Frequency of transfers each year
- (8) Whether or not the child repeated a grade
- (9) Special class placement recommended
- (10) Psychological testing prescribed
- (11) Reading readiness at end of kindergarten
- (12) Academic achievement testing in 2nd and 3rd grades
- (13) Classroom teacher marks in reading and arithmetic each year
- (14) Classroom behavior ratings (described in more detail below)

Classroom behavior ratings were obtained from teachers employing the DESB rating scales, (Spivack and Swift, 1966, 1967; Swife and Spivack, 1968). Ratings were made at the beginning and the end of kindergarten and first grade, and at the end of second and third grades. The 47-item, 11 factor scale measures how well the child adapts to the variety of task and social demands made in the usual, structured classroom. Factor scores comprise the raw data for profile type assignments which define overall level of adjustment to the demands of the classroom (see Spivack, Swift and Prewitt, 1972). The eleven factors measure:

1. Classroom Disturbance: extent to which the child teases and torments classmates, interferes with others' work, is quickly drawn into noise-making, and must be reprimanded or controlled.
2. Impatience: extent to which the child starts work too quickly, is sloppy in doing school work, is unwilling to go back over work, and rushes through the work.
3. Disrespect-Defiance: extent to which the child speaks disrespectfully to the teacher, resists doing what is asked, belittles the work being done, and breaks classroom rules.
4. External Blame: extent to which the child says the teacher does not help, never calls on him or her, blames external circumstances when things do not go well, and is quick to say the work assigned is too hard.
5. Achievement Anxiety: extent to which the child gets upset about test scores, worries about knowing the "right" answers, is overly anxious when tests are given, is sensitive to criticism or correction.
6. External Reliance: extent to which the child looks to others for direction, relies on the teacher for direction, requires precise directions, and has difficulty making his or her own decisions.
7. Comprehension: extent to which the child gets the point of what is going on in class, seems able to apply what he or she has learned, and knows material when called upon to recite.
8. Inattentive-Withdrawn: extent to which the child loses attention, seems to be oblivious to what transpires in the classroom, and seems difficult to reach or preoccupied.
9. Irrelevant-Responsiveness: extent to which the child tells exaggerated stories, gives irrelevant answers, interrupts when the teacher is talking, and makes irrelevant comments during classroom discussion.

10. Creative Initiative: extent to which the child brings things to class that relate to current topics, talks about things in an interesting fashion, initiates classroom discussion, and introduces personal experiences into class discussion.
11. Need for Closeness to Teacher: extent to which the child seeks out the teacher before or after class, offers to do things for the teacher, is friendly toward the teacher, and likes to be physically close to the teacher.

Each factor provides a continuous score, and each child's profile of factor scores was "typed," following the system devised by Spivack, Swift and Prewitt (1972). There are two basic ineffective adaptation types. One type exhibits high external reliance (factor 6) and inattentiveness (factor 8), and the other exhibits signs of poor self-control: high scores on three or more of factors 1 (classroom disturbance), 2 (impatience), 3 (disrespect-defiance), 4 (external blame), and 9 (irrelevant-responsiveness). Both types exhibit abnormally low levels of creative-initiative (factor 10), and comprehension (factor 7). Some children exhibit behaviors which when profiled reveal qualities of both maladaptive types. Successfully adaptive profiles are in general the converse of these patterns, reflecting the youngster is productively engaged and involved in the learning and social processes of the classroom, and comprehending what is going on. A third category of profile type consisted of children whose profiles were doubtful. The behavior patterns were not clearly maladaptive, but on the other hand did reveal some questionable signs. All data from these prior years were card punched for computer use. A report on the prediction of third grade achievement and behavioral adjustment from kindergarten and first grade classroom behavior has been reported, (Spivack and Swift, 1977).

Subsequent data collection methodology (1975-present)

Since a period of time elapsed between the initial data collection period and the second, the first and rather substantial task for the second data collection period was to locate as many of the original youngsters as possible through existing records. This original search and data collection of the second data collection period has in essence been repeated each year. This has meant annual repetition of certain data collection, leading to enlarged criterion groups as well as more refined and reliable measures. It has also resulted in new measures being added, as the research team gained sophistication in what is available and possible to measure, and new questions have emerged out of completed studies. The process continues at the time of this writing.

General search strategy:

The search began with the Pupil Directories of the Philadelphia Public Schools. These computerized annual directories supply information about school attended, grade, room and section, address and race of each child. Directories for 1975 and 1976 were reviewed, and prior information (between 1969 and 1972) examined via microfilm. In 118 instances children

were recorded as present in a school, but not found in that school. In these case, telephone calls were made to each school to discover where the student had been transferred to, or whether the child had left the system, and if so to where. From this procedure it was possible to group the available sample by present school location.

A list of 79 names was prepared consisting of youngsters transferred in the past to parochial schools or those the team was unable to locate. This list was used by administrative staff of the parochial schools in their search through their records to see how many of these student were enrolled there (in 1976-1977).

Of the original sample of 661, 464 were identified as currently enrolled in the public schools, and 28 currently enrolled in the parochial schools.

As part of the overall plan to maintain confidentiality, new identification code numbers were assigned to each youngster of the initial study of nine years earlier. Once new numbers were assigned by school personnel, this listing was crosshatched with the old ID numbers, and the paired numbers sent to the computer center. Names and old ID numbers were erased from the records, and replaced by the new ID numbers. Thus, only the schools retained the crosshatch, the research group retaining only the data and new ID numbers.

A list of all students in the original study, with birthdate and new ID number was sent to the Philadelphia Police Department. Police personnel searched their records for any evidence of contact with the Juvenile Aid Division. Records on identified contact cases were sent to the research team coded only by ID number, with names and other identifying information removed. In subsequent years a similar search was made of records of all Philadelphia community mental health and mental retardation centers.

Specific search strategy into school records: Having identified which children were currently enrolled in what schools, the next task was to specifically articulate what data would be of use and what informants available. Initial meetings were held with staff of the Research and Evaluation Department of the public schools to discuss school record systems, and these were followed by a meeting with a district level supervisor. Out of these a general strategy emerged.

First, a letter was sent to the principals of all schools in the study, briefly describing the nature and purpose of the study, and noting the approval of the collaborative study by the School Board. The letter was signed by appropriate school officials. Each principal was told he or she would receive a telephone call shortly to set up a visit to discuss the project further. Concurrently, a cover letter went to all district superintendents, enclosing a copy of the letter to the principals. A letter also was sent to the teacher's union explaining the nature of the study and teacher involvement, and

another to the Director of Pupil Personnel and Counseling. All of these letters succeeded in informing those whose staffs would be involved. Further material requested by the union was sent.

The principal of each school was contacted, and an interview set up. Initial interviews were exploratory to learn more about details of information available, from what sources and informants, and how best to proceed without disrupting the work of school personnel. Throughout this process, a crucial issue was how best to enlist cooperation.

A similar procedure was followed in obtaining the cooperation of the Philadelphia Archdiocesan School System. Those students not located in the records of the Philadelphia Public Schools were listed. Permission to conduct the study, using the confidentiality procedures discussed previously, was obtained from the Archdiocesan Director of Pupil Personnel. With the advice and cooperation of the assistance superintendent of schools, a letter was drafted to the principals of each school, along with a copy of the list of students not located in the Public schools. Code numbers of those students located in the parochial schools were returned to the Research team. Responses to the inquiry exceeded 90% and an additional 25 subjects were located in this manner. The principal of each school in which these 25 subjects were in attendance was contacted, and arrangements made to collect the information in a manner least disruptive to school personnel.

Specific search strategy into police records: As with the school record search, it was first necessary to determine which subjects in the sample had recorded contact with the police Juvenile Aid Division (J.A.D.) of the Philadelphia Police Department. Initial meetings were conducted with the office of the Chief Inspector to determine the specific information available and to devise a collection strategy. Strictest procedures had to be arranged to ensure confidentiality of information of such a delicate nature. In all instances police personnel removed personal ID information prior to forwarding the materials to the research team.

Working with the records division of the Police J.A.D., a search was made through police files to identify subjects who had had police contact. Of the 661 subjects for whom the search was initially conducted (1976) 45 subjects were identified as having had at least one recorded contact with the J.A.D. As of the date of this writing (1980) 129 cases with police contact have been identified.

For the purpose of categorizing the nature of the contact in a manner similar to that suggested by Sellin and Wolfgang (1969) copies of each complete J.A.D. record as well as specific incident reports for each contact were obtained. After all personal identification information had been removed, these copies were made available to the research group.

Specific search strategy into community mental health/mental retardation (CMH/MR) center records: While not in the original design of the project, a decision was made during the second year of the project to obtain data on whether project subjects had had a CMH/MR contact during their lives, and if so when and why. Initial meetings were held with research and evaluation service directors of each center, at which time their cooperation was enlisted. Clearance was obtained from the Deputy Health Commissioner of the Office of Mental Health and Mental Retardation of the City of Philadelphia, after assurance of confidentiality of information was made and procedures established similar to those employed with the police. After each R & E Director searched his or her center's records, and indicated when contact was made and the intake diagnosis, names and personal information were removed and data forwarded to the research team with ID numbers.

As of the date of this writing (1980) 66 cases with CMH/MR contact have been identified.

School data collected. From meetings with school administrators, District Superintendents, Vice Principals and Principals a strategy was developed to record that information which was both crucial to establishing the validity of the research effort and consistently accessible at each school. Following is a compilation of the school related information on which the collection effort was centered, as well as the strategies used for collection.

- (1) Hahnemann High School Behavior (HHSB) Scale: Classroom behavior ratings were obtained from teachers employing the HHSB, (see Swift & Spivack, 1969). The HHSB is a standardized and validated instrument created to provide a system for identifying and measuring classroom behaviors of junior and senior high school students. The focus is upon behaviors which interfere with, facilitate or reflect the student's level of success in adapting to the task and interpersonal demands of the classroom environment.

A separate HHSB was requested of the English and Math teachers based on the assumption that ratings would vary as a result of the differences in the nature of the coursework and the differing perceptions of the two teachers. In those instances where one teacher instructed the subject in both English and Math, the completion of separate English and Math HHSB's were requested from that teacher.

The HHSB is a 45-item scale which provides continuous scores on 13 factors. The factor scores may be used as separate scores, or the entire profile of factor may be classified into a type (Spivack and Swift, 1972).

The factor scores measure:

1. Reasoning ability - the extent to which the student grasps new ideas quickly, is able to sift through information and work out answers on his or her own, and is able to apply information and principles to new or unfamiliar problems.
2. Originality - the degree to which a student presents points of view to stimulate the thinking of others; promotes discussion in class; presents unique, yet relevant, ideas; prepares assignments and carries out tasks in an interesting, original fashion.
3. Verbal Interaction - the degree of involvement in the information flow in class.
4. Rapport with the teacher - the desire for, and willingness to relate positively to the teacher.
5. Anxious Producer - the degree to which the student feels he or she must produce and even overproduce in the classroom.
6. General Anxiety - the extent to which meeting academic and/or interpersonal demands poses problems revealed by open fearfulness, flustering, or tension.
7. Quite - Withdrawn - the level and extent of limitations in verbal interaction, communication, and involvement with peers, the teacher, and class activities.
8. Poor Work Habits - the adequacy of the student's attempts at preparation, organization, and the meeting of requirements for attendance and completion of assignments.
9. Lack of Intellectual Independence - the degree to which a student has difficulty relying upon his or her own inner intellectual resources to think, work, and persevere when confronted with teacher expectations and peer opinions.
10. Dogmatic - Inflexible - the extent to which the student is "closed" to the ideas of others and is unwilling or unable to deal with or weigh the opinions and ideas of others.
11. Verbal - Negativism - the extent to which there is a verbalized negative attitude toward all facets of the educational involvement.
12. Disturbance - Restlessness - the level of occurrence of active, obstreperous behavior.
13. Expressed Inability - the degree to which the student feels unable to cope with the work demands in the classroom.

As noted earlier, factor scores define classroom behavior profiles which may be reliably classified on the basis of prominent features. After each HHSB was scored, the profile of each was "typed" as:

1. average: scores falling within $\pm 1SD$ with reference to the scale's norms
2. high average achiever: like 1, but with slightly elevated factors 1-5
3. above average achiever: like 2, but with two or more of factors 1-5 greater than one S.D.
4. above average with acting out: elevated factors 1-5, with elevation of factors 10-12 as well
5. average with non-specific aberrance: factors 1-5 in average range, but any 2 or more other factors significantly elevated

6. underachiever: factors 1-5 low, with no other striking aberrance
7. underachiever with non-specific aberrance: factors 1-5 in average range, but any 2 or more other factors significantly elevated
8. average, acting out: factors 1-5 average, but factors 10-15 and sometimes 9 and 13 elevated
9. acting out with underachievement: same as 8, but factors 1-5 low
10. quiet-withdrawn average: factors 1-5 average, but 7 elevated
11. quiet-withdrawn, underachiever: factors 1-5 low, and 7 elevated
12. quiet-withdrawn and anxious, average: factors 6 and 7 elevated, and factors 1-5 average
13. quiet-withdrawn and anxious, underachiever: factors 6 and 7 elevated, and factors 1-5 low

While each English and Math ratings of each rated youngster could be profiled according to this system, it was also felt desirable to be able to classify as many youngsters as possible according to both English and Math profiles. This was possible in most instances by classifying each pair of profiles into the following:

1. Acting out with low achievement: type 9 on one profile, and type 6, 7, 8 or 9 on the other profile
2. Acting out with low achievement mixed: type 9, with type 1, 2, 5, 10 or 12
3. Acting out with average/above average achievement: type 4 or 8, with type 1, 2, 3, 4 or 8
4. Achievers: type 3, with type 1, 2 or 3; or type 2, with type 2
5. Average: type 1, with type 1, 2, 5, 10 or 12
6. Quiet-withdrawn mixed: type 11 or 13, with type 1, 6, 10, 11, 12 or 13
7. Underachievers: type 6 or 7, with type 6 or 7
8. Underachievers mixed: type 1, with type 6 or 7

2. The Teacher Rated Adjustment Scale (TRAS). At the same time each teacher completed the HHSB, he or she also completed ratings on 10 dimensions of adjustment on each youngster. In each instance the teacher rated the frequency of occurrence of the behavior on a 5-point scale, ratings made relative to the average youngster in such a classroom. The dimensions rated were:

1. Appear friendly and outgoing with peers.
2. Act depressed or despondent in mood.
3. Act socially withdrawn, uncommunicative, aloof, daydreaming.
4. Show positive leadership qualities.
5. Act agitated or anxious.
6. Act interested in what is going on in class or school.
7. Get overemotional; react with immediate anger or upset.
8. Act timid, shy, fearful, self-conscious.
9. Act uncooperative, disobedient, disruptive with others.
10. Act assaultive, quarrelsome, initiate fights.

A factor analysis of these ratings suggested a Conduct Disturbance Score defined by items 7, 9 and 10. A Neurotic Withdrawal Score was defined by items 1, 3 and 8.

3. The "Pink Slip" Information. Each Philadelphia High School maintains a system of "pink slips" comprised of in-school forms used by teachers and/or administrators to formally record the description of any student offense warranting disciplinary action. The file is retained only for one year, and is maintained by the vice principal or disciplinary officer. Each slip describes the offense in the recorder's own words (usually the teacher) and the action taken in response to the infraction.

Arrangements were made at each school to examine the pink slip file, and to record what was written verbatim. Only information for the previous year was available. A reliable method was developed to classify types and severity of pink slip offense. The categories of offense were:

1. Offense against the institution--those offenses that violate the property of the school or its rules or traditions
2. Offense directed at the teacher
3. Offense directed at another student

Once an offense was classified, it was further categorized as minor or major in severity. In order to create such categories, a 10 percent random sample was drawn from the complete research file, and examined. The content of these suggested it was feasible to have minor and major offense categories wherein the "victim" was another child, an adult (usually the teacher), or the institution. Some offenses involved property, and others a personal attack or affront. In instances of property violation, the magnitude of the property damage determined whether it would be major or minor. In instances of personal violation, minor offenses usually involved verbal abuse, while major offenses generally involved physical contact.

Examples of offenses and their categories are presented below:

1. Personal

A. Offenses against Child

1. MINOR (primarily verbal actions)

verbal abuse and arguments
name calling
teasing
chasing

2. MAJOR (primarily physical actions)

provoking fight
threatening with object
attacking
sexually intimidating (touching opposite sex and refusing to stop)

B. Offenses against Adult

1. MINOR (primarily verbal actions)

arguments
name calling
insulting or rude remarks

2. MAJOR (primarily physical actions)

threatening teacher or other adult
swinging at teacher
tripping teacher

II. Property

A. Offenses against Child

1. MINOR

taking books (intending to return)
taking school supplies
hiding personal property (hats, coats, etc.)

2. MAJOR

stealing personal property
destroying personal property
abusing personal property

B. Offenses against Adult

1. MINOR

taking personal property (intending to return)
hiding personal property

2. MAJOR

stealing personal property
destroying personal property
abusing personal property

C. Offenses against Institution

1. MINOR

tearing page from book
writing on desk
knocking over furniture

2. MAJOR

stealing school equipment
destroying school property
abusing school property

III. Violation of Institutional Rules

1. MINOR

class cutting
loitering
disruptive in class
late to class
ridiculed class discipline

2. MAJOR

throwing things
playing with elevator
lying to adult (authority)
refusing to do work
walking out of class
smoking in class

The scoring procedure was conservative, in that no inferences were made. Only information actually recorded was considered. Each pink slip received a total offense score, as well as set of category scores (e.g. total of minor offenses; total of major offenses). Each minor offense received a score of "1", and major offense "2". For each pink slip, each subcategory was only scored once, so that if a student pushed and then punched another student, he received a score of "2" and not "4". Incidents were scored only if an offense occurred and was described, and not if evidence indicated that the teacher was only "annoyed". History leading up to an offense was not scored. History leading up to an offense was not scored.

Since the response to each offense was recorded, it was also possible to score these. The assumption made was that, in general, the more drastic the reaction, the more serious the situation. Examination of pink slips suggested there were six classes of action which may result, singly or in combination, and that these could easily be ranked according to seriousness:

1. Verbal reprimand
2. Detention
3. Parent contact for conference
4. Suspension
5. Section or class change
6. Expulsion or transfer to disciplinary school

Each category was assigned a weighted score equal to its rank, every total action taken scored for each of its elements, each final score equalling the sum of the scores of its elements. Rater reliability rho correlations between three independent scorers all exceeded .95.

4. Counselor information. Initial conferences with school personnel had suggested it would be feasible to obtain certain information about positive school activities engaged in (e.g., clubs, athletics, orchestra or band, etc.), ratings from reliable informants about each child's extra-classroom school misbehavior (to supplement pink slips), information about the need to refer the child to various agencies for help and why, and whether any counselor contact had taken place and, again, why. Field experience with a Counselor Form indicated this was not entirely realistic. On the other hand, with good cooperation most vice principal and school counselor information about referrals was initially obtained as well as the reason for contact. Eventually only data on counselor contact have been analyzed, since the reliability of data collected about referral has been called into question, and direct data for mental health centers became available. In summary, it has been possible to record and use information from the counselor form indicating whether contact with the child had taken place, when, and whether for conduct, emotional disturbance, academic difficulty, or other (at times positive) reasons.

5. Student history and achievement data. After examination of the student records of the public schools, a Student History Data Sheet was designed to capture a wealth of information covering each student's complete school history. Achievement data were collected from the 4th through 7th grades, since similar data from earlier grades were available on each student from prior data collection, (1969-1972).

If a grade had been repeated, information was used only from the year of the repeated grade, since the fact of repeated grade was recorded and considered elsewhere in the data collection system. In public school grades 4-6, levels of achievement in Math are recorded rather than grades, and since these are not convertible the information could not be used. Letter marks were recorded for Written Expression, Social Studies, and Science. Since parochial schools employ a number rather than letter grading system, a conversion table was created to transfer the numbers into letters. Since the public schools have curriculum changes after the 6th grade, there were minor changes in subjects for which marks were recorded. Despite these and other adjustments made necessary by the data, conversions were possible so as to make subject marks comparable and transformable into a graded number system for analysis.

California Achievement Test (CAT) scores were available from the 4th grade on, adding to earlier test data in 2nd grade existing in prior data. Both the percentile score and developmental score were recorded in Reading, Math, Language, and Spelling.

All department classroom marks were recorded, whether called "behavior grades" or "citizenship practices." These were averaged when necessary, and converted to a number system similar to that created for data from earlier grades in order to maintain continuity and consistency.

Rates of absenteeism were recorded from public school files and considered in three categories: excused absence, unexcused absence, and total number of absences. Only the last was available in parochial school files.

At the time of this writing, a system to record school "dropout" is being created, as well as means of identifying school "non-attenders."

Finally, it was possible to record number of changes in school and residence of each child, each year over the total school history of the child. These data were transcribed from school records directly.

Most of the school data were converted into scores for data analysis which would allow correlational analysis since data were continuous. Other data were coded nominally and required different methods of analysis. All data were coded for time of occurrence so that a cross-hatch of data from the same points in time could be created if necessary.

6. Psychological testing information. Each subject in the study was sought in the Special Education file of the public schools to see if testing was ever done. If so, information from each testing was coded according to a system devised from pilot work. The coding system allowed for language of each psychological testing report to be coded as follows:

1. Conduct/behavioral disturbance/management difficulties (e.g., "poor impulse control"--"acting out")
2. Emotional/social adjustment difficulties (e.g., "feels insecure"--"anxious and immature"--"gets easily upset"--"is dependent, fearful, has low self-concept")
3. Academic talent
4. Academic difficulties (e.g., "learning slowly"--is "retarded in learning ability")
5. Academic difficulties, specialized deficiency (e.g., has a specific reading handicap)
6. Dangerous to self or others
7. Neurological indications

Any record could be categorized as falling into more than one category. Reports of 10 cases were examined and the reports coded independently by two researchers. Of 80 possible decision, there was agreement on 76, or 95 percent. Total IQ scores were also recorded.

Police Data Collected: Initial meetings were held with police officials and officers in the department's Record Division to define exactly what data were available and in what form, and how it could be extracted and used while maintaining confidentiality.

Juvenile police records in Philadelphia are segregated from those of adult offenders, and maintained in an automated system until the age of 18. Two forms contained the desired information for the present study. Form 75-163 is a cumulative record, one of which is maintained for each youth who has had contact with the Juvenile Aid Division. In addition to identifying information, it lists each alleged offense in chronological order, the date on which the offense occurred, whether the youth was arrested in

connection with the offense, the 2-digit Philadelphia Police district in which the offense occurred, and the 4-digit complaint or incident report number described below.

The 75-48 Philadelphia Police Department Complaint or Incident Report is generally completed in writing by the officer(s) on duty in whose patrol area an offense occurs. It is standard procedure that this form must be completed each time an officer initially answers a complaint regardless of the nature or outcome of the complaint. The form contains information which identifies: (1) the precise area in which the offense allegedly occurred, giving police district number, sector of that district, car number of the investigating police vehicle(s), street location and whether complaint involved an indoor or outdoor incident; (2) precise time of occurrence, including date (day, month, year), time car left to investigate the complaint, time car departed the scene of the alleged incident, and day code; (3) complainant and his or her address; (4) details of the alleged incident including whether the complaint is "founded" (i.e., was the reported incident found to have actually occurred?), and the specific nature of the incident or complaint; and (5) information from the officer(s) answering the complaint and, in the case of juvenile offenders, the identification of the Juvenile Aid Division officer who is assigned responsibility for investigating the incident.

The 75-48 is completed each time a complaint is filed; however, since subsequent forms utilized by the Police Department record the same information as well as details of subsequent investigation of the complaint, the 75-48 was examined by the research team only in those instances where no subsequent, more detailed information was available.

The police records thus supplied information on every incident in which a study subject was involved, and the following information for each subject was obtained:

1. Date
2. Gang affiliation, if any
3. Whether violation occurred alone or with others
4. Disposition
5. Classification of crime

Confidentiality

Considering the nature of the present study, requiring the gathering and analysis of information from school, police and psychiatric records, procedures were established early to guarantee confidentiality and meet the specific regulations of all agencies involved. The process has involved:

1. Elimination of names from preexisting computerized tapes;
2. Substitution of new numbers for names and old identification numbers, with knowledge of the cross-hatch of all 661 names to new ID numbers known only to the public schools;

3. A method of searching police, school and community mental health/mental retardation center files so that only police, school and mental health personnel know names of youths involved, with data about these youths passed on to research staff with only new ID numbers to link data together for analysis;
4. A method of data collection such that the schools do not know what youths have had police or mental health contact, and the police know nothing about the school adjustment and achievement of the youth involved or mental health contact, and mental health facilities know nothing about police or school histories and data.

The procedures have been approved by the legal staffs of all involved, and have worked smoothly.

Results

Since the purpose of the present chapter is to provide the methodology and not results of the study in question, the reader is referred to reports describing results which are available from the author. These are listed in the references. However, for the sake of closure, a listing of major findings relative to early high risk signs for delinquency are provided below.

1. The original selection process aimed at identifying a high risk group for study was successful: approximately one-quarter of the subjects have had at least one police contact, and over half have had to be referred at least once within the schools for disciplinary action during 8th and 9th grades; approximately 10 percent have had a CMH/MR contact to date, and 28 percent were left back in grade or placed in a special class for "retarded educables" at some time during their school histories; median academic achievement test scores fall between the 25th and 35th percentiles on national norms, and approximately one-half were exhibiting classroom behavior patterns that interfere with learning by the end of their elementary school years.
2. Conduct disturbance in the school is related to police contact in the community; one out of every five police contact subjects had at least one contact with a CMH/MR facility; the youngster who displays misconduct in one classroom is likely to in another, and from one year to the next once adolescence is reached.
3. By the end of kindergarten, a high risk classroom behavior pattern emerges that predicts all indices of delinquency employed, in and out of school; this pattern is stronger in grade 3; the pattern predates academic exposure and thus can not be ascribed specifically to academic failure; the behavior pattern suggests children who are less able than their peers

to tolerate the tension of waiting, or (to be more parsimonious) have not learned to wait. Their children exhibit an inappropriate drive to enter into and complete tasks before having a clear conception of where to go, or what must be done. They surge forth before getting directions straight, preferring not to look back or think much about the quality or implications of what they are doing. They may interrupt others, or blurt out what is on their minds as though oblivious of others and their reaction or feelings. Because of this, their behavior is often intrusive, as well as disruptive and annoying to others. This unreflective, forward driving style does not necessarily imply hostile intent or negative affect. In fact, there is often an associated tendency for active social involvement--to keep the action going and to be where the noise is--with no evidence of need or tendency to injure or attack others.

4. Classroom academic achievement, test scores, absenteeism rate and mobility in early grades in and out of themselves do not define high risk for police contact or conduct disturbance in the school, nor do they add predictive power during early grades once the child begins to exhibit the high risk pattern.
5. However, subjects doing more poorly academically or who are exhibiting higher school absenteeism at the time of first police contact are more likely to have a subsequent police contact than subjects doing better at school at the time of their initial police contact; early school failure may bring out a high risk behavior pattern.
6. Subjects with histories of being left back or having been placed (usually during early grades) in special classes for retarded educables are significantly more likely than peers to have subsequent police contact and/or in-school conduct referral; among subjects exhibiting the high risk pattern at the end of kindergarten, subsequently being in a special class or being left back in grade experience adds to the likelihood of later delinquency, and this finding is not due to differences in intellectual ability between the police contact and non-police contact groups.
7. A high risk child at the end of kindergarten is at especially high risk if the high risk classroom behavior pattern persists during grades 1-2; among such initially high risk children, information about school achievement mobility or attendance during grades 1-3, does not add to the predictive picture.
8. A low risk child at the end of kindergarten who develops the high risk pattern during grades 1-3 is more likely to commit delinquencies than peers; however, with this child poor academic achievement and classroom academic behaviors may accompany and possibly pre-date the emergence of the high risk pattern.

Project Reports

Spivack, G., Rapsher, L., Cohen, A., and Gross, R. High risk early signs for delinquency and related behavioral difficulties: The first nine years of a longitudinal study. Interim Report I. Phila.: Department of Mental Health Sciences, Hahnemann Medical College and Hospital, April 1978.

Spivack, G., Rapsher, L., Cohen, A., and Gross, R. High risk early signs for delinquency and related behavioral difficulties: A ten year longitudinal study. Interim Report II. Phila.: Department of Mental Health Sciences, Hahnemann Medical College and Hospital, May 1979.

Spivack, G. and Rapsher, L. Special class placement and being left back in school as elements in a high risk delinquency history. Phila.: Department of Mental Health Sciences, Hahnemann Medical College and Hospital, November 1979.

Spivack, G. High risk events and the emergency of the high risk child in the primary grades, Phila.: Department of Mental Health Sciences, Hahnemann Medical College and Hospital, December 1979.

References

- Sellin, T. & Wolfgang, M. (Eds). Delinquency, New York: Wiley 1969.
- Spivack, G. & Swift, M. The Devereux Elementary School Behavior Scale: A study of the nature and organization of achievement related disturbed classroom behavior, Journal of Special Education, 1966, 1, 71-91.
- Spivack, G. & Swift, M. Devereux Elementary School Behavior Rating Scale Manual, Devon, Pa.: The Devereux Foundation, 1967.
- Spivack, G., & Swift, M. The Hahnemann High School Behavior Rating Scale (HHSB). Philadelphia, Pa., 1972.
- Spivack, G. & Swift, M. "High risk" classroom behaviors in kindergarten and first grade. American Journal of Community Psychology, 1977, 5, 385-397.
- Spivack, G., Swift, M., & Prewitt, J. Syndromes of disturbed classroom behavior: A behavioral diagnostic system for elementary schools, Journal of Special Education, 1972, 5, 269-292.
- Swift, M. & Spivack, G. Achievement-related classroom behavior of secondary school normal and disturbed students. Exceptional Children, 1969, 36, 99-104.
- Swift, M. & Spivack, G. The assessment of achievement-related classroom behavior. Journal of Special Education, 1968, 2, 137-154.

The National Longitudinal Surveys of Labor Force Behavior

Herbert S. Parnes
and
Kezia Sproat

The National Longitudinal Surveys
of Labor Force Behavior

Herbert S. Parnes
and
Kezia Sproat

I. Introduction

The National Longitudinal Surveys (NLS) were initiated at the Ohio State University Center for Human Resource Research in 1965 under contract with the U.S. Department of Labor. Originally planned as a five-year survey of the work experience of older men in the U.S. civilian non-institutionalized labor force, they have been enlarged in scope and extended in time repeatedly. The NLS now include data from 32,000 respondents in five age/sex cohorts. The older and younger men's cohorts (45 to 59 years of age and 14 to 24 years of age, respectively) were first surveyed in 1966; a mature women's (30 to 44 years of age) in 1967; young women (14 to 24) in 1968; and a youth cohort (males and females 14 to 21 years of age) in 1979. An interdisciplinary team of researchers at the Ohio State University Center for Human Resource Research analyzes the data and prepares policy-oriented research reports. Data are provided at cost on computer tapes for the public.

This chapter will give a broad picture of the data available in the NLS. Further information is provided by The NLS Handbook, which may be obtained free of charge from the NLS Users' Office at The Center for Human Resource Research, 5701 North High Street, Worthington, Ohio 43085.

II. Sampling, Interviewing and Estimating Procedures

The sampling and field work for the four original NLS cohorts have been the responsibility of the Census Bureau while The National Opinion Research Center of the University of Chicago (NORC) has these responsibilities for the new cohort of youth. The sampling methods employed by each of these organizations are described below.

A. Census Bureau (Four Original Cohorts)

Each of the four original NLS samples was designed to represent the civilian noninstitutional population of the United States at approximately the time of the initial survey. Because there were no additions to the samples over the years of the surveys, they cannot be construed to be precisely representative of the civilian population in any year after the first. The departure is most significant in the case of the young men, since those who were in the military service as of 1966 but who subsequently returned to civilian life are not represented. Conversely, young men who entered the service subsequent to 1966 are unrepresented during their military service, since no attempt was made to survey members of the sample during such periods. The New Youth Cohort includes approximately 1,300 respondents who were serving in the armed forces at the time of the first interview.

Sampling Procedure

Each of the four age-sex cohorts is represented by a multi-stage probability sample located in 235 sample areas comprising 485 counties and independent cities representing every state and the District of Columbia. The 235 sample areas were selected by grouping all of the nation's counties and independent cities into about 1,900 primary sampling units (PSU's) and further forming 235 strata of one or more PSU's that are relatively homogeneous according to socio-economic characteristics. Within each of the strata a single PSU was selected to represent the stratum. Within each PSU a probability sample of housing units was selected to represent the civilian noninstitutional population.

Since one of the survey requirements was to provide separate reliable statistics for blacks, households in predominantly black enumeration districts (ED's) were selected at a rate between three and four times that for households in predominately white ED's. The sample was designed to provide approximately 5,000 interviews for each of the four cohorts--about 1,500 blacks and 3,500 whites. When this requirement was examined in light of the expected number of persons in each age-sex-color group, it was found that about 42,000 households would be required in order to find the requisite number of blacks in each age-sex group.

An initial sample of 35,360 housing units was selected and a screening interview took place in March and April, 1966. Usable information was collected for 34,662 households, a completion rate of 98.0 percent.

Men 45-59. Following the initial interview and screening operation, 5,518 males age 45-59 were designated to be interviewed. These were sampled differentially within four strata: whites in white ED's (i.e., ED's which contained predominantly white households), blacks in white ED's, whites in black ED's, and blacks in black ED's.

Young men 14-24. The original plan called for using the initial screening to select all four samples. On reflection it was decided to rescreen the sample in the fall of 1966 prior to the first interview of young men, since a seven-month delay between screening and interview seemed inordinate in view of the mobility of young men in their late teens and early twenties.

To increase efficiency, it was decided to stratify the sample for the rescreening by the presence or absence of a 14- to 24-year-old male in the household. The probability is high that a household that contained a 14- to 24-year-old in March would also have had one in September. However, to insure that the sample also represented persons who had moved into sample households in the intervening period, a sample of addresses that previously had no 14- to 24-year-old males was also included in the screening operation.

This phase of the screening began in early September, 1966. Since a telephone number had been recorded for most households at the time of the initial interview, every attempt was made to complete the short screening interview by telephone. Following the screening operation, the same differential sampling process that had been used in the case of the older male cohort produced 5,713 males aged 14 to 24 who were designated to be interviewed.

Women 30-44 and 14-24. The rescreened sample of households from which the sample of young men was drawn was used also to obtain the sample of women 30 to 44 years of age. Using the same type of differential sampling ratios described above, 5,393 women were selected for interview. Finally, from the same group of households, a sample of 5,533 young women aged 14 to 24 as of January 1, 1968, was designated.

Multiple-respondent households. The total number of households represented in the four NLS samples of individuals is 13,582. Thus, a number of households have yielded more than one respondent, frequently across cohorts. By cross-indexing households and respondents, the Census Bureau has provided a means of identifying respondents having common households when the samples were originally drawn. The data tapes made available by the Center show for each respondent (1) an identification number within the particular cohort, (2) an identification number for the household from

which the respondent was originally selected, and (3) the identification numbers of any other respondents (in the same or other cohorts) who were selected from the same household. The Center makes available lists of households from different cohorts, e.g., husband-wife, father-daughter, brother-sister, etc.

Sampling Weights. In each survey year after the initial interview, the sample was reduced for reasons of noninterview. In order to compensate for these losses the sampling weights of the individuals who were interviewed had to be revised. This revision was done in two stages. First, the out-of-scope noninterviews in each of the years were identified by the Bureau of the Census and eliminated from the sample of noninterviews. This group consisted of individuals who were institutionalized, had died, were members of the Armed Forces or who had moved outside the United States--i.e., individuals who were no longer members of the non-institutional, civilian population of the United States.

The second stage in the adjustment acknowledged the non-random characteristics of the noninterviews attributable to refusals or disappearance of respondents. For each of the survey years the eligible noninterviewees and those interviewed were distributed into strata (cells) according to their race, years of school completed, and the number of years in their 1966 (1967) place of residence (cohorts of men and women) or their race, education of their father, and years in 1966 (1968) place of residence (boys and girls). Within each of the cells the base year sampling weights of those interviewed were increased by a factor which measured the relative size (using base year weights) of the noninterviewed population to the interviewed population in that year.

One further adjustment was needed because the young men who were in the Armed Forces at the time the sample was selected were excluded from the sample. Since most of these men subsequently returned to the civilian non-institutional U.S. population, the population estimates in any given year based on the revised sampling weights of those interviewed in our sample are understated. To correct for this understatement the Bureau of the Census has provided an adjustment factor for each of the cells. The revised sampling weights of the young men incorporate this additional adjustment.

B. National Opinion Research Center (Cohort of Male and Female Youth

14 - 21 Years of Age)

Sampling Procedure. The target population for this survey consists of ten groups between the ages of 14 and 21 on January 1, 1979:

1. Hispanic males

2. Hispanic females
3. Black, non-Hispanic males
4. Black, non-Hispanic females
5. Economically disadvantaged non-Black, non-Hispanic males
6. Economically disadvantaged non-Black, non-Hispanic females
7. All non-Black, non-Hispanic males
8. All non-Black, non-Hispanic females
9. Male military personnel
10. Female military personnel

For each of these groups, individuals were considered in the population if they were living within the 50 states or if they were on active military duty outside the United States. Individuals living in institutions on a permanent basis were excluded.

With the exception of individuals on active military duty, all sample selection was accomplished through a multistage, stratified area probability sample of dwelling units and group quarter units. A screening interview was administered at approximately 75,000 dwellings and group quarters distributed among 1,818 sample segments in 202 Primary Sampling Units. A primary sampling unit is composed of either a single county or group of counties (SMSA). In certain special situations, state-defined units are termed "independent cities," "parishes." In these instances, such units are used in the definition of primary sampling units. Included in this screening interview was information which would allow the identification of persons eligible for sample membership.

Approximately 18,000 of the screening interviews were carried out among 918 sample segments in the 102 Primary Sampling Units constituting the NORC Master Probability Sample of the United States. This sample is designed to maximize the statistical efficiency of samples which are "cross-sectional" with respect to the general population. Specifically, through the several stages of sample selection (counties, enumeration districts, block groups, sample listing units), probabilities of selection are based upon either total population or total housing units.

The remaining 57,000 screening interviews were carried out among 900 sample segments in a 100 PSU sample specifically designed to produce statistically efficient samples of Groups 1-6. All stages of sampling, except the final stage, were carried out with probabilities proportional to a

linear combination of population size for these groups. The effect of this procedure is to produce sample listing segments which vary significantly in terms of total population size but tend toward equality with respect to the target groups.

In the final stage of sample selection (i.e., dwelling units within sample listing segments), a moderate degree of oversampling was employed in order to increase the sample composition with respect to Groups 1-6. Since the use of oversampling tends to decrease sample efficiency, attempts were made to hold required oversampling to a minimum.

At all selected dwellings, attempts were made to obtain appropriate classification information for all persons living in the dwelling. In order to minimize the potential for "interviewer effect," survey interviewers were not informed about specific groups that would be included in the subsequent interviews.

Since certain classes of group members do not usually reside in dwelling units, procedures were also used to establish "linkages" between dwellings and these individuals. As part of the initial screening, household respondents were asked if there were any primary family connections to household members who were currently away from the household. Included in this group were college students, persons in the military, persons in prisons and other institutions. Household respondents were also asked to name persons who might occasionally stay at the dwelling who did not have any other "usual place of residence." For each individual identified in this process an attempt was made to determine whether or not the individual would be "linked" to some other household (e.g., college students living off campus in their own dwelling units). All individuals without other linkages were included in the household composition for purposes of subsampling. The only exception to this linkage occurred in the case of active military personnel who were sampled separately.

Selecting sample respondents for base year interviewing was based upon the individuals identified at the household screening phase. Base year samples for Groups 1-6 were selected from individuals identified in both screening samples (i.e., 102 PSU Cross-sectional Sample and 100 PSU Special Purpose Sample). To the extent that individuals identified in the screening phase were obtained with different probabilities of selection (because of selective oversampling), the selection of base year samples will attempt to minimize these probability differences.

Sample respondents in Groups 7 and 8 were selected from the 102 PSU Cross-sectional Sample. Restriction of subsampling for these groups to the 102 PSU National Sample is based upon considerations of per element statistical efficiency.

Members on active military duty were sampled from rosters provided by the Department of Defense. Sample selection was accomplished in two stages. In the first stage, a sample of approximately 200 "military units" was selected. These units were selected with probabilities proportional to the number of persons 14-21 within the unit. Within selected units, persons 14-21 were subsampled with probabilities inversely proportional to the first-stage selection probability. Separate samples were selected of males and females. A small number of military personnel living in their own households, not on military bases, were included in the household screening too.

III. Data Collection Waves

Table I presents years and types of interview for each of the five NLS cohorts.

Table I: NLS Data Collection Sequence

YEAR	MEN 45-59	BOYS 14-24	WOMEN 30-44	GIRLS 14-24	YOUTH 14-21
1966	interview	interview			
1967	interview	interview	interview		
1968	mail survey	interview	mail survey	interview	
1969	interview	interview	interview	interview	
1970		interview		interview	
1971	interview	interview	interview	interview	
1972			interview	interview	
1973	telephone survey	telephone survey			
1974			telephone survey		
1975	telephone survey	telephone survey		telephone survey	
1976	interview	interview	telephone survey		
1977			interview	telephone survey	
1978	telephone survey	telephone survey		interview	
1979			telephone survey		interview
1980	telephone survey	telephone survey		telephone survey	interview
1981	(final) interview	(final) interview	telephone survey		interview
1982			(final) interview	telephone survey	interview
1983				(final) interview	interview
1984					interview

IV. Description of the NLS Data

A. The Four Original Cohorts

The National Longitudinal Surveys were designed primarily to analyze the sources of variation in the labor market behavior and experience of the four age-sex subsets of the U.S. population represented by the samples. Thus, the information collected from the respondents in the several samples relates to variables that either represent significant aspects of labor market activity and labor market status or that are hypothesized to influence such activity or status.

In order to know precisely what information is available and how it has been coded, there is no substitute for a careful examination of the interview schedules and codebooks. Nevertheless, it is possible to describe here in reasonable detail the major variables that are included in the data tapes. Although the classification is not entirely satisfactory, the variables will be categorized in the following pages under the major headings of "labor market experience variables," "socio-economic and human capital variables," and "environmental variables." Where information has been collected on a particular topic, for convenience, the four original cohorts will be identified as M (men 45-59 years of age), W (women 30-44 years of age), B ("boys," i.e., young men 14-24), and G ("girls," i.e., young women 14-24). Where reference is to a particular survey year, the last two digits of the year will be used. Thus, "W67" refers to the 1967 interview of the woman's cohort. "Survey week" will be used to refer to the calendar week preceding the interview with each respondent, as this is the reference week for questions relating to labor force and employment status.

1. Labor Market Experience Variables

Labor Force Participation. There are several variables used to measure the extent of labor market participation of the respondents. These are reasonably uniform across cohorts and across surveys.

Survey week labor force status. This is the conventional measure based on the individual's activity in the survey week. The interview questions and the coding procedure for classifying individuals are identical to those used in the Current Population Survey (CPS).

Number of weeks worked. In each survey, information was obtained on the number of weeks during the preceding year in which the respondent had done any work for pay or profit.

Numbers of hours per week usually worked. Respondents were asked the number of hours they "usually" worked in the weeks in which they worked.

Number of hours worked in survey week. This is based on questions identical to those in the CPS.

Work experience prior to initial survey. In W7 only, respondents were asked the number of years since leaving school in which they worked at least six months. For married women without children this was divided between (a) the period between school-leaving and first marriage and (b) the period since first marriage. For married women with children, there were three periods covered: (a) school-leaving and first marriage, (b) first marriage and birth of first child, and (c) since birth of first child.

Unemployment. Several measures of unemployment have been used that are uniform across cohorts and across survey years.

Survey week employment status. Employment status ("at work," "with a job but not at work," and "unemployed") is defined and measured exactly as it is in the CPS. For those unemployed, this includes a measure of duration and a specification of method(s) of job search.

Number of weeks unemployed. For weeks in which they did not work, respondents were asked in how many they were on layoff from a job or were looking for work.

Job history prior to initial survey. In the initial survey of each cohort, a number of questions are asked relating to particular jobs held by the respondent in the past. These are less extensive in the case of the two younger cohorts than in the case of the women and men. Generally speaking, a job is defined as a continuous period of service with a given employer or in a self-employed status. Information is collected on length of service in the job and on the reason for having left it. Class of worker is also coded, as are both occupation and industry of employment. The latter two variables are coded in terms of the 1960 Census classification system at the three-digit level. In addition, most occupational assignments are coded according to "prestige" level as measured by the Duncan socio-economic index. (For a description of the Index see Duncan, 1961.) Finally, location of job is coded in such a way as to permit identification of cases of geographic mobility. The specific retrospective job history information that has been collected for each cohort is described below.

Men. Respondents were asked in the 1966 survey about their first job after leaving school, the longest job of their entire career, and, in cases where they had been in their current job less than five years, about their immediately preceding job. In addition, they were asked to identify the "best" occupational assignment of their career, and, if no longer in it, the reason and whether they would like once again to be practicing that occupation.

Women. In the 1967 survey, each married respondent who had ever borne children was asked about three jobs held in the past: longest job between time of school-leaving and first marriage, longest job held between time of first marriage and birth of first child, and longest job held since birth of first child. Married respondents who had never had children were asked to identify the longest job held between school and first marriage and the longest held since that marriage. Never-married respondents with no children were asked about the first job after leaving school that lasted at least six months and the longest job ever held. Never-married respondents who had borne children were asked about the longest job held between the time of leaving school and birth of first child and the longest job held since the birth of that child.

Girls and Boys. In the initial surveys, respondents were asked about two previous jobs: the first in which they had served at least a month after leaving full-time school and the job, if any, they had held during their last full year in high school. In addition to the usual questions about occupation, industry, length of service, etc., respondents were asked how they had found each of these jobs.

Job history since the initial survey. A fairly detailed record of the jobs held by respondents has been maintained during the years covered by the surveys, although the degree of detail varies somewhat among cohorts and is considerably less in the telephone surveys than in the longer face-to-face interviews.

Survey week job. In all surveys of all cohorts, information has been collected on the job held by the respondent in the survey week or, if not employed, on the most recent job. The questions are identical to those used in the Current Population Survey and permit the coding of occupation, industry, class of worker, and location. Duncan socio-economic status is also coded for the occupation. In addition to the CPS questions, respondents are also asked how much they usually earn at their current job and how many hours per week they usually work at it.

Jobs held between interviews. The extent of detail recorded for jobs held between survey weeks varies, but generally includes as a minimum an indication of any job change since the preceding interview and the reason therefore.

Interfirm mobility. The occurrence of an interfirm job change is coded as between every possible pair of survey dates.

Migration. Although the location (city and state) of each job is contained in the interview schedule, Census rules of confidentiality preclude coding a geographic location in any greater detail than Census Division. Nevertheless, comparisons of location from one survey year to the next have been made and individuals are coded according to whether they are in the same county of SMSA as that of the preceding interview. Thus, the data permit an analysis of the incidence and the extent of geographic movement but do not permit an analysis of the direction of movement among areas smaller than a Census Division.

Commuting time and cost. Commuting time and costs were obtained for the job held in the survey week at least once for each of the four cohorts.

2. Socioeconomic and Human Capital Variables

Early formative influences. In the initial survey of each of the cohorts, information was collected on a number of aspects of the early background of the respondent. Among the characteristics included in all of the cohorts are nationality; type of residence as a teenager (farm, rural nonfarm, urban, suburban, etc.); father's educational attainment and occupation; mother's occupation and educational attainment; and family structure when respondent was 14 or 15 years of age. In addition, for the two cohorts of youth, another dimension of the socioeconomic status of family of origin is provided by an index indicating the extent of reading materials in the home when respondent was 14 years of age.

Education. Information on educational attainment is more complete for the two cohorts of youth than for the two older cohorts. For the latter, information was collected in the initial survey on highest year of regular school completed as well as type of high school curriculum. For the two cohorts of youth, information on current enrollment status has been obtained in each survey, as well as highest year of regular school completed. High school curriculum, field of specialization in college, and the existence and amount of scholarship aid in college are illustrative of the additional kinds of information available for the young men and women. Also, from

questionnaires sent to the last high schools attended by the respondents, there are data on school characteristics that form the basis for an index of high school quality.

Training. Information has been collected in every survey on extent and character of training outside the formal education system. In the initial survey of each cohort, questions were asked about training received in company training programs, in business college, technical institutes, apprenticeship programs, or from any other source. Additionally, in each subsequent survey a comparable question was asked concerning any training received since the previous interview.

Health and physical condition. In the initial survey of each cohort, respondents were asked whether health problems or physical conditions either prevented working or limited the amount or kind of work they could do. If so, the nature and duration of the limitation were described. These questions were repeated in a number of subsequent surveys. In the 1971 surveys of all four cohorts and once again about five years later, these questions were supplemented by a set of detailed questions about the precise nature of functional limitations and the kinds of environmental working conditions that had to be avoided because of health problems.

Marital and family characteristics. In each of the surveys a considerable amount of information has been collected about the relationship between the respondent and other members of his or her family, as well as on the characteristics and activities of other family members. Such information includes current marital status of respondent, health condition of spouse, number of children, number of dependents, whether parents and parents-in-law are living, and (for each household member) age, school enrollment status and last year of school completed, weeks worked during the past 12 months, usual number of hours per week, and occupation of longest job. For female respondents who are employed and who have children there are also questions on type of costs of child care arrangements. Those unemployed and out of the labor force are asked about the need for child care should the respondent find a job or decide to work outside the home. The young women have also been asked at several points in time their views on "the ideal number of children for a family" and how many (more) children they expected to have.

Financial characteristics. Rather detailed information relating to the wealth and the income of respondents and members of their families exists for all cohorts over the several years covered by the surveys. Respondents have been asked periodically about assets owned and liabilities owed, including ownership of home, farm, business or professional practice, other real estate, automobile(s), savings or checking accounts, U.S. Savings Bonds, stocks, bonds, or shares in mutual funds, mortgages or other loans due, and several other types of debt.

In each survey there are fairly detailed questions on income received by respondent and spouse, and (in the case of the women's and men's cohorts) by other family members during the preceding year. There is variation among cohorts in the specific sources of income that are specified. An illustrative list includes wages, salaries, commissions, and tips; (net) income from a business or profession; (net) income from operating a farm; unemployment compensation; rental income; interest or dividends or income from estates or trusts; disability benefits from a variety of sources; (other) Social Security benefits; (other) public assistance or welfare payments; food stamps; government pensions; and all other sources.

Military service. For the older cohort of men, information was obtained in M69 on whether respondent ever served in the armed forces and, if so, when. For the younger cohort, the initial survey likewise ascertained whether the respondent ever served in the armed services and, if so, duration of service and age at separation. Respondents who had not served were asked whether they had tried to enter the service and, if so, why they had not been accepted. In subsequent surveys of the young men additional questions relating to military service were included.

Job attitude. A variety of measures of attitudes toward current job have been included in the surveys of all four cohorts. For example, in every survey employed respondents have been asked how well they like their current jobs. Responses are classified into four categories: "like very much," "like somewhat," "dislike somewhat," and "dislike very much." In the initial survey of each cohort and occasionally thereafter, employed respondents were asked to indicate what factors they like least. Job attachment has been measured at several points in time by asking respondents to react to a hypothetical job offer in their line of work with another employer (a) if it is in the same local labor market area, and (b) if it is elsewhere. In each case the respondent is asked how much the job would have to pay in order for him to be willing to take it. For both female age cohorts, the initial survey instrument asks respondents how long they will continue to work at their present job and, if the answer is less than five years, what their plans are immediately thereafter.

Work attitudes. As distinguished from attitudes toward specific jobs, the surveys have also attempted to tap a variety of attitudes toward work in general. Work commitment has been measured by asking respondents whether they would continue to work if, by some chance, they were to get enough money to live comfortably without working. As an indication of work motivation respondents were asked in the initial survey of each cohort, "What would you say is the more important thing about any job: wages or liking the kind of work you are doing?" In every survey, unemployed workers have been asked what kind of work they are looking for, whether there are any restrictions on their availability for work, and what wage rate they require. Workers out of the labor force have been asked whether they would take a job if offered one by an employer. In cases of affirmative answer, the type of work required and the wage rate are explored.

In each of the surveys of the older group of men, respondents have been asked at what age they expect to retire from their regular job and what their plans are thereafter. There are also questions relating to whether there is a compulsory retirement age and whether the respondent will be eligible for benefits under a pension plan other than Social Security. In later surveys there are also questions on attitude toward retirement.

In the initial surveys of the women and the girls, and periodically thereafter, there were questions (which differed somewhat between the two cohorts) designed to measure the respondent's attitude to the propriety of labor market activity on the part of married women with children. Also, each married respondent has been asked about her husband's attitude toward her working. There are five categories of response, ranging from "like it very much" to "dislike it very much."

Aspiration and expectation. In the surveys of the two cohorts of youth, questions have been asked concerning both educational and occupational aspirations. In each case, where the response in a given year has differed from that of the previous year, reasons for the change have been explored.

Attitudes toward high school and college. In the initial interview with the two young cohorts those who were attending or had attended high school were asked how well they (had) liked their high school experience, the subjects they liked best, and the subjects they liked least. A comparable set of questions relating to college was asked of those with college experience.

Other social psychological variables. There are three additional social psychological variables used in the NLS that are of particular importance. The first of these is available for all four cohorts, while the other two relate only to the two cohorts of youth. In the 1968 survey of the boys, the 1969 surveys of the men and women, and the 1970 survey of the girls an abbreviated version of the Rotter I-E scale was administered. The Rotter test purports to measure perceived "locus of control," i.e., whether an individual believes that he or she is in charge of his or her own destiny (internal) or that what happens is determined by forces over which the individual has little control (external). The identical test was administered on two subsequent occasions for each cohort. For each of the two cohorts of youth, scores on mental ability tests were obtained from the last high school attended by each respondent. In B66 and G69 there were sets of questions designed to measure the extent of respondents' occupational knowledge. In the case of the young men, respondents were asked to identify ten occupations, to indicate the amount of education the typical incumbent of each occupation had, and to indicate for each of eight pairs of occupations the one in which earnings are higher. For the young women the occupational information test consisted solely of ten occupational identification questions.

Retrospective evaluation of labor market experience. At the end of each five-year period there is a series of questions relating to the respondent's perception of his or her labor market experience over the recent past. For example, respondents were asked whether they believed they had "progressed, moved backward, or just about held your own" over the period, so far as work was concerned. The bases for the response to this question were then explored.

3. Environmental Variables

Several characteristics of the local area in which the respondent resides have been coded on the data tapes. In all cases these derive from sources other than the National Longitudinal Survey instruments themselves. The local area to which the data relate is, with very few exceptions, one of the 235 Primary Sampling Units (PSU's) from which the samples were drawn, and almost invariably consists of either an SMSA or one or more counties. These include size of labor force in the local area, the local area unemployment rate, an index of demand for female labor, an index of demand for teenage male labor, and the existence and type of accredited colleges in the local area.

B. The 1979 Youth Cohort

The variables that have been included in the first two surveys are shown in Table II. New classes of variables include those on government jobs and training programs, military service, significant others, school discipline, delinquency and drugs, and reported police contacts. New attitudinal variables on self-esteem and education and new questions on religion and ethnicity have also been added.

Table II: Variables Collected for the New Youth Cohort

Variable	1979	1980
1. <u>Labor Market Experience Variables</u>		
a. Current Labor Force and Employment Status and Characteristics of Current Job		
Survey week labor force and employment status	*	*
Occupation (Census code)	*	*
Occupation (DOT code)	*	*
Industry	*	*
Class of worker	*	*
Starting date	*	*
Number of hours worked in survey week	*	*
Number of hours per week usually worked	*	*
Shift worked	*	*
Hourly rate of pay	*	*
Commuting time to current job	*	*
Covered by collective bargaining?	*	*
Is respondent union member?	*	*
Availability of vacation and insurance benefits on current job	*	*
Reason for leaving last job	*	*
Job satisfaction	*	*
Job characteristics inventory	*	*
Size of employer	*	*
Minority status of co-workers and supervisor	*	*
b. Work experience since January 1, 1978, or Previous Survey		
Number of weeks worked	*	*
Number of hours usually worked per week	*	*
Number of weeks unemployed	*	*
Spells of unemployment	*	*
Number of weeks out of labor force	*	*

Variable	1979	1980
c. Characteristics of Jobs with More than 20 Hours per Week and More than 9 Weeks in Duration since January 1, 1978		
Occupation	*	*
Industry	*	*
Class of worker	*	*
Starting date	*	*
Number of hours per week usually worked	*	*
Hourly rate of pay	*	*
Covered by collective bargaining?	*	*
Is respondent union member?	*	*
Reason for leaving job	*	*

2. Human Capital and Other Socioeconomic Variables

a. Early Formative Influences

Nationality and birthplace	*	
Ethnic self-identification	*	
Person(s) respondent lived with at age 14	*	
Occupation of primary male adult and primary female adult when respondent was 14	*	
Highest grade of school completed by father	*	
Highest grade of school completed by mother	*	
Birthplace of parents	*	
Were magazines available in home at age 14?	*	
Were newspapers available in home at age 14?	*	
Was library card available in home at age 14?	*	
Current and past religion	*	
Whether language other than English spoken when respondent was a child	*	

b. Migration

Counties of residence since January 1, 1978	*	*
---	---	---

Variable	1979	1980
c. Education		
Current enrollment status	*	*
Highest grade of school completed	*	*
Reason stopped attending high school	*	*
Date of last enrollment	*	*
Is (was) school public or private?	*	*
High school curriculum	*	*
College degree received	*	*
Type of college attending	*	*
Field of specialization in college	*	*
College tuition	*	*
Financial aid in college	*	*
Attitude toward selected aspects of high school	*	*
Courses taken during last year of high school	*	*
d. Vocational Training outside Regular School		
Type(s) of training	*	*
Duration of training programs	*	*
Whether training was completed	*	*
Whether obtained degree, certificate or journeyman's card	*	*
Usual hours per week spent in training	*	*
e. Government Jobs and Training Programs		
Participation in programs	*	*
Type of program	*	*
Satisfaction with program	*	*
Participation in program of aid on subsequent jobs	*	*
Services provided by program	*	*
Hours per week and per day spent in program	*	*
Aspects liked most and least about programs	*	*
f. Health and Physical Condition		
Does health limit work?	*	*
Duration of health limitation	*	*
Type of health problem (ICD-9 code)	*	*

Variable	1979	1980
g. Marital and Family Characteristics		
Marital status	*	*
Number of dependents	*	*
Are respondent's parents living?	*	*
Occupation of father and mother in past year	*	*
Number of weeks worked by father and mother in past year	*	*
Education of family members	*	*
Occupation of spouse	*	*
Periods of living away from parents	*	*
Extent of work of spouse in survey year	*	*
Age at which respondent expects to marry	*	*
Changes in marital status since January 1, 1978	*	*
Number and duration of marriages	*	*
Number and age distribution of children living in household	*	*
Expected number of children	*	*
Number of children respondent considers ideal	*	*
h. Financial Characteristics		
Total family income in previous year	*	*
Income of respondent (and spouse) from farm or own business in previous year	*	*
Income of respondent (and spouse) from wages or salary in previous year	*	*
Income of respondent (and spouse) from unemployment compensation in previous year	*	*
Income from public assistance in previous year	*	*
Income from food stamps in previous year	*	*
Income from pensions and Social Security	*	*
Income from other sources in previous year	*	*
i. Military Service (Current or Past)		
Branch of armed force	*	*
Months spent in armed force	*	*
Military occupation(s)	*	*
ROTC or officer training	*	*
Reserve or guard activities	*	*
Pay grade and income	*	*
Type and amount of military training	*	*
Formal education received while in service	*	*
Future military plans	*	*
Why entered military	*	*
Reason for leaving military	*	*

Variable	1979	1980
j. Work Attitudes		
Would respondent continue to work if he/she had enough money to live on?	*	*
Characteristics of job respondent is willing to take (respondents unemployed or out of labor force)	*	*
Reaction to hypothetical job offers to specific types of work	*	*
k. Educational and Occupational Aspirations and Expectations		
Would you like to receive more education or training? Type?	*	*
How much education do you wish and do you think you will actually get?	*	*
What kind of work would you like to be doing at age 35?	*	*
Expectation of achieving occupational goal	*	*
l. Other Social/Psychological Variables		
Knowledge of world of work score	*	*
Rotter Internal-External Locus of Control score (4 times)	*	*
Attitude toward women working	*	*
Rosenberg self-esteem (10 times)	*	*
m. Retrospective Evaluation of Labor Market Experience		
Perception of age discrimination	*	*
Perception of race discrimination	*	*
Perception of sex discrimination	*	*
Reason for problems in obtaining employment	*	*
n. Significant Others		
Persons having most influence on respondent	*	*
Responses of significant other to various situations	*	*
o. School Discipline		
Ever been suspended from school? Date.	*	*
Ever been expelled from school? Date.	*	*

Variable	1979	1980
p. Delinquency and Drugs		
Activities within last year (20 items)		*
Income from illegal activities within last year		*
q. Reported Police Contacts		
Number of times stopped by police		*
Number of times booked or arrested		*
Number of convictions		*
Charges respondent has been convicted		*
Number of times incarcerated. Date of release.		*

V. Results of NLS-Based Research

Over twenty volumes of research results and over four-hundred articles have been based on the NLS, and it is impossible to summarize these findings briefly. (A bibliography of NLS-based publications appears in the NLS Handbook.) They address a wide variety of subjects in diverse fields, including many areas of economics and sociology and diverse topics, including unionism and fertility. To illustrate the kinds of policy-relevant findings for which the NLS was initiated, however, we present below two brief summaries of research results.

A. Teenagers: What are their choices about work?

1. School completion is a major determinant of labor market success.
2. Young people don't know enough about the labor market.
3. Post-school training confers definite advantages to both drop-outs and graduates.
4. For some youth the problem of unemployment is not temporary.
5. Job shopping, even if it means higher rates of teenage unemployment, may be desirable.
6. Created public service or subsidized private sector jobs for teens can be a good means of providing labor market exploration.
7. Teenage pregnancy is an increasingly serious barrier to young women's long-term career success, and it seriously jeopardizes the welfare of the new generation.
8. Transportation inadequacy is an impediment to teenage employment.
9. Discrimination still accounts for a substantial part of the labor market problems of youth, but it is not inevitable.

B. Middle-aged and Older Men in the United States Labor Force

1. A larger proportion of blacks than whites leave the labor force due to disability, but when controls are introduced for earnings relative to disability benefits, the difference disappears.
2. In addition to forcing men into early retirement, health problems have other serious consequences for the welfare of middle-aged men and their families.

3. Although it is decreasing, race discrimination still causes severe labor market disadvantages for blacks.
4. While not common, as many as one in twelve men lose their jobs in mid-life after building up substantial equities in them; rarely do they regain equally advantageous positions.
5. Most men who retire for reasons other than health are "very happy" in retirement and would, if they had it to do over again, retire at the same age. A minority, however, recognize that they made a mistake in opting for retirement.
6. Most middle-aged and older men who remain in the labor force enjoy relatively favorable positions.

References

- Andrisani, P.J. An Empirical Analysis of the Dual Labor Market Theory. Columbus: Center for Human Resource Research, 1973.
- Ashenfelter, O. Estimating the Effect of Training Programs on Earnings. The Review of Economics and Statistics, 1978, LX (1), 47-57.
- Becker, B.E., & Hills, S.M. Teenage Unemployment: Some Evidence of the Long-run Effects on Wages. Journal of Human Resources, in press.
- Borus, M.E. Indicators of CETA Performance. Industrial and Labor Relations Review, 1978, 32 (1), 3-14.
- Cooley, T.F., McGuire, T.W., & Prescott, E.C. The Impact of Manpower Training on Earnings: An Econometric Analysis. Final Report MEL 76-01 to Office of Program Evaluation. Pittsburgh: Employment and Training Administration, U.S. Department of Labor, 1975.
- Davis, H.E. Growth of Benefits in a Cohort of Pension Plans. Monthly Labor Review, 1971, 94 (5), 48.
- Daymont, T.N. Changes in Black-White Labor Market Opportunities. In Herbert S. Parnes (Ed.), Work and Retirement, a Longitudinal Study of Men. Cambridge: The MIT Press (forthcoming, 1981).
- Duncan, O.D. A Socioeconomic Index for All Occasions. In A.J. Reiss, Jr., et al. (Ed.), Occupations and Social Status. New York: Free Press, 1961.
- Employment and Training Report of the President. Washington, D.C.: Department of Labor Employment and Training Administration, 1979.
- Flanagan, R.J. Labor Force Experience, Job Turnover, and Radio Differentials. Review of Economics and Statistics, 1974, 56, 521-529.
- King, R.H. The Labor Market Consequences of Dropping out of High School. Columbus: Center for Human Resource Research, 1978.
- Kohen, A.I. Determinants of Early Labor Market Success Among Young Men: Race, Ability, Quantity and Quality of Schooling. Columbus: Center for Human Resource Research, 1973.
- Leighton, L.S., & Mincer, J. Labor Turnover and Youth Unemployment. In The Youth Employment Problem: Its Nature, Causes, and Consequences, in press.

- Mincer, J., & Polachek, S. Family Investments in Human Capital: Earnings of Women. Journal of Political Economy, 1974, 82, 576-608.
- Moore, K.A.; Waite, L.J.; Caldwell, S.B., & Hofferth, S.L. The Consequences of Age at First Childbirth: Educational Attainment. Washington: The Urban Institute, 1978.
- Mott, F.L., & Shaw, L.B. Work and Family in School Leaving Years: A Comparison of Female High School Graduates and Dropouts. Columbus: Center for Human Resource Research, 1978.
- Oaxaca, R.L. Theory and Measurement in the Economics of Discrimination. Equal Rights and Industrial Relations. Madison: Industrial Relations Research Association Series, 1977.
- Parnes, H.S., & Kohen, A.I. Occupational Information and Labor Market Status: The Case of Young Men. Columbus: Center for Human Resource Research, 1973. (Also in Journal of Human Resources, 1975, 10, 44-55.)
- Sandell, S.H., & Shapiro, D. The Theory of Human Capital and the Earnings of Women: A Reexamination of the Evidence. Journal of Human Resources, 1978a, 13 (1), 103-119.
- Sandell, S.H., & Shapiro, D. Work Expectations, Human Capital Accumulation, and the Wages of Young Women. Columbus: Center for Human Resource Research, 1978b.
- Shapiro, D., & Carr, T.J. Investments in Human Capital and the Earnings of Young Women. In F.L. Mott (Ed.), Years for Decision, Columbus: Center for Human Resource Research, 1977.
- Statistical Abstract of the United States 1977. Washington, D.C.: Bureau of the Census, 1977.
- U.S. Department of Labor, Office of Youth Problems. Interim Report on Analysis of Cognitive and Attitudinal Change Among VEP and SPEDY Enrollees. In Compilation of Reports on the 1978 Summer Youth Employment Program, 1979, I.

The Southern Youth Study:
Longitudinal Survey Research In
The Deep South*

J. Steven Picou
 John K. Thomas
 Arthur G. Cosby

*J. Steven Picou is associate professor of Sociology and Arthur G. Cosby is professor of Sociology and Rural Sociology at Texas A&M University, College Station, Texas. John K. Thomas is Social Climate Analyst with Shell Oil Company, Houston, Texas. He was formerly a research associate at Texas A&M University. Support for this research was provided by USDA (CSRS) Regional Research Project S-114, The Texas Agriculture Experiment Station Research Project H-3141, and the National Institute of Education Research Project NIE-G-76-0072.

The Southern Youth Study:
Longitudinal Survey Research In
The Deep South

J. Steven Picou

John K. Thomas

Arthur G. Cosby

In this chapter we will provide an overview of the Southern Youth Study (SYS). Included in this overview will be a concise review of the origins, sponsoring agencies and participants in the SYS. Next, we will identify the data sources and specify demographic characteristics of the sample, along with comparative information regarding the participating states. We will also address data collection procedures for all waves of this cooperative longitudinal research effort. A general outline of instruments and measurements utilized, and a brief specification of research outcomes studied by SYS social scientists will follow.

Overview of SYS

The Southern Youth Study data originated from a series of research projects that began in 1958. The first project was a regional research study entitled, "Factors in the Adjustment of Individuals and Families in Low-Income Farm Areas of the South." Although this study generated descriptive information concerning family conditions in low-income rural areas, it failed to produce information regarding career development and decision-making processes of southern youth. Consequently, a second regional project was proposed in 1966 to specifically examine such processes and identify career achievement influences. The 1966 study was entitled, "Human Resource Development and Mobility in the Rural South", and incorporated as much as possible the study design of the earlier 1958 research. More specifically, the objectives of the 1966 research were: (1) to analyze changes in selected areas of development--occupation, level of living, family structure, functions and resources, social participation and attitudes; and (2) to analyze the relationships of occupational and social

aspirations to decision-making, mobility, and development of youth (Jacobs, 1973).

Funding for this study was provided for by the Cooperative State Research Service in the United States Department of Agriculture and the individual state Agricultural Experiment Stations. A total of six southern states participated in the longitudinal research survey.

The states participating in both the 1966 and 1968 waves were Alabama, Georgia, South Carolina, and Texas; Mississippi took part in 1966 only, while Louisiana joined in 1968. A follow-up to this project in which all six states participated was conducted in 1972. This second project was entitled, "Development of Human Resource Potentials of Rural Youth in the South and Their Patterns of Mobility" (White, 1974). The Southern Youth Study (SYS) refers to the total three-wave longitudinal research effort.

Source of Data

A major concern of the original 1966 study was the selection of a sample of youth residing in counties which were rural and economically disadvantaged (Thomas, 1979). Each of the participating states generated sampling procedures to insure such representativeness among its respondents.

Counties in Alabama and Georgia were selected because of their location near Appalachia, which traditionally has been regarded as a high poverty area. In Mississippi, all counties were ranked on the basis of the following three factors: (1) a weighted index of education, income, and occupation; (2) percent change in population from 1960 to 1963; and (3) percent in wages paid under the Mississippi Employment Security Commission. One high and one low ranking county were randomly selected. In Texas, counties were chosen because they had a high poverty index and were over 30 percent Black (White, 1974).

A nonprobability, or purposive sampling procedure, was characteristic for most of the SYS states, except in the cases of South Carolina and Louisiana. In South Carolina, a stratified random sample based on size of school and racial predominance criteria was obtained from all high schools. This sample resulted in 24 counties being selected. When Louisiana entered the SYS in 1968, another sampling procedure was introduced. The state was divided into four geographical areas, with each area being socioeconomically homogeneous. Schools were randomly selected from rural parishes to obtain a racially proportionate sample.

Although different sampling procedures were used, collectively they resulted in the selection of high schools and counties that were socioeconomically similar. Thomas (1970) and Lever (1969) have presented information derived from the 1960 census which showed that the SYS counties selected in the 1966 wave had a lower mean family income per year than for the United States as a whole. White (1974) has also reported racial and rural similarities between parishes selected in Louisiana and the counties of the other states. A socioeconomic profile of five of the SYS states and their counties participating in the present study is presented in Table 1. This profile reflects general population, education, income, and occupation characteristics as they appeared in 1970.

Population Characteristics of Sample States

Despite the fact that nine counties had over 50,000 residents, most of the SYS counties were characterized by small populations with a high percentage of Blacks and rural residents (Table 1). For example, about a third of the study counties had populations of less than 20,000. The racial composition of SYS county populations varied with Alabama and Georgia being the most atypical in the study. Six of their eight participating counties withstanding, 25 of the 47 SYS counties were more than 30 percent Black; eleven of these 25 counties had greater than a 50 percent Black population.

Education Characteristics of the Sampling States

Educational attainment levels are also compared in Table 1 for the SYS counties, their respective states, and the United States. All the SYS counties and states are below the national median levels of educational attainment than their respective states. At the other extreme, no SYS county or state, except Texas, had more than the national (10.7%) percentage of college educated persons. With 18 counties reporting less than five percent, the SYS counties had a lower range of about three to ten percent.

Income Characteristics of the Sample States

The SYS counties and states were low when contrasted with the median family income figure of \$9,957 for the United States. The medians of the study counties range from a low of \$3,173 to \$8,754.

Occupational Characteristics of Sample Counties

The occupational characteristics of the SYS states were consistent with those patterns observed for the other socioeconomic characteristics. This is to say, 33 counties reported lower percentages (based on respective total county populations) of persons 16 years of age who are employed than the 37.7 percent figure for the United States. The range for all counties, however, was about 24 to 44 percent. Breaking down the occupations of those persons employed revealed a high concentration of semi- and unskilled occupations in the SYS counties. Compared with the national figure of 36.1 percent, all but one county and one state, as a whole, had higher concentrations in these occupations.

In summary, the observed similarities of the study counties in the states participating in the SYS indicate a homogeneous socioeconomic context for interviewing high school adolescents, who generally come from a rural population, characterized by a high percentage of Blacks, low educational attainment levels, low family income levels, and a predominance of semi- and unskilled occupations.

Data Collection Procedures

The series of interviews began in the spring of 1966. At this point in time the respondents were sophomores in high school. After the schools had been selected from rural and socioeconomically "disadvantaged" counties, students were administered questionnaires in a group situation. This method of administering the research instrument involved either interviewing sophomores collectively at one time or interviewing them while they attended classes of a common course (White, 1974). Group interviews were conducted by project directors and graduate students. No attempts were made to contact students who were absent the day of the interview. Determining the number of sophomore students not contacted was hampered by poor, inaccurate school records (in some cases, no records were made available) and by students with double statuses (classified as juniors but taking sophomore courses). In total, information was collected from 7,972 sophomore students. This total figure includes Mississippi's participation and non-whites other than blacks.

Wave II

In 1968, senior students were interviewed in the same selected high schools. The approximate size and number of senior classes are shown in Table 2 for the five states which participated in the second data collection. Although no senior class had an enrollment above 200, South Carolina had the largest class size averaging 125 students per class in its 38 selected schools. These large classes inflated the average size for all the SYS schools. In spite of the total average of approximately 79 students, half of the school's senior enrollments were less than 49 students (Table 2).

Within each of the 96 schools of the SYS, students were again administered questionnaires similar to the 1966 schedule. After the group interviews were conducted, students were matched according to their participation in both waves of the study. The matching process sorted out both students who had transferred in the interim and original students who failed to be interviewed. Student attrition between 1966 and 1968 was attributed to: (1) out-of-state migration; (2) transferal to a school system that was not part of the SYS sample; (3) school drop-outs; and (4) absenteeism the day of the interview.

Wave III

A second follow-up was conducted in the Summer, 1972 in an effort to interview respondents four years after they were expected to have graduated from high school. A sample was drawn from the original pool of students. This sample procedure was necessitated by the anticipated wide dispersment of the respondents following high school graduation and by limited resources available to the participating states for locating and interviewing these students. Consequently, a decision was made by the SYS Technical Committee to select 250 respondents from the larger pool of students for each state. The panel was to be stratified 50:50 by race (Black:White) and 60:40 by sex

(Male:Female). Thus, in each state, 75 Black and 75 White males and 50 Black and 50 White females were randomly selected from the pool of high school students who had participated in the first two waves of the study. Hypothetically, a panel of 1250 persons was possible with the collective effort (White, 1974). However, Louisiana and Mississippi were to select respondents from the wave in which they had participated. Although Louisiana had not taken part in 1966, social origin and school tracking information, which had been collected for the other SYS states in 1966, was sought in 1968. Mississippi had only 234 original participants in 1966. Consequently, when the state joined the 1972 follow-up effort, its total pool of students was needed to approximate the ideal SYS quota of 250. With the Mississippi sample, the total SYS panel potential was 1484.

After demanding sample ratios and selecting panel members, an extensive locating procedure was initiated. Initially, verification of student addresses was made. Parents or relatives were contacted regarding the current address of particular panel members. Individuals who had remained in-state were personally interviewed.

The research instrument employed consisted of two major sections. The first section contained questions chosen from instruments and measures used in prior waves. These questions, which the respondent completed, related to aspirational and expectational attitudes toward a number of social status issues. The second part of the instrument contained new questions designed to measure current activities and attainments of respondents. Questions of this nature were presented by field interviewers who recorded the replies. In a few instances, questionnaires were mailed or telephone interviews were conducted for panel members who had migrated long distances. If an individual had left the state of origin, but remained within the SYS region, a cooperative effort was made between the research teams of the participating states to administer the personal interviews. The research instrument employed in both cases was identical to that used for interviewing the majority of the panel. However, for mail-out situations, a detailed set of instructions accompanied the questionnaire.

The total number of respondents interviewed by the fall, 1972 was 1052 for five of the six participating states (Mississippi included, 1268). In Table 3 a breakdown of the SYS sample size are presented for each wave and state. When compared with the intended stratification ratios, the final panel was: by race--42.6 percent Black and 57.4 percent White; by sex 58.0 percent Male and 42.0 percent Female (Table 3).

Wave IV

At the time of this writing in the Fall of 1979, a Wave IV recontact was in progress in each of the six SYS states. The strategy of this fourth interview was to reinterview each of the youths who had participated in the 1972 third wave. Approximately 75 percent of the 1228 students in the early periods had been contacted and interviewed at this interim phase. It is anticipated that by early 1980 over 80 percent recontacts will have been achieved.

The focus of this 1979 study was to take repeated measurements again on the core attitudinal and behavioral variables utilized in earlier SYS surveys.

Types of Research Outcomes Studied

Although a variety of topics have occupied the concerns of SYS researchers, most studies have focused on outcome variables that reflect specific stages of this cooperative longitudinal project. The first series of data collected resulted in information on career-related attitudes, values and aspirations. Most studies attempted to determine the consequences of socioeconomic status, residence, race, and other background characteristics for the occupational and educational desires and plans of the respondents (e.g., see: Kuvlesky & Ohlendorf, 1968; Cosby, 1971; Picou & Cosby, 1971; Cosby & Picou, 1971). During this initial phase of the project, a variety of conceptual studies also emerged which structured future research directions in terms of measurement strategies (Kuvlesky & Bealer, 1966). Additionally, the first wave of the SYS project produced empirical studies of residence projections (Kuvlesky & Pelham, 1970), military service values (Kuvlesky & Dameron, 1971), marriage and fertility plans (Kuvlesky & Obordo, 1972), language patterns (Patella & Kuvlesky, 1973), school desegregation (Falk & Cosby, 1974), and interscholastic athletics (Picou & Curry, 1974). These studies and many others (e.g., see: Cosby et al., 1974) attempted to determine the consequences of a variety of social structural, attitudinal, and behavioral determinants of career-related attitudes and values.

The second data collection (Wave II) provided information on career-related attitudes and values for a later time period in the respondents' high school careers. This data collection allowed for analyses of changes and stability in career and residential aspirations during high school (Cosby, 1974; Cosby & Howard, 1976). Throughout this time period additional theoretical development regarding the occupational choice process emerged from SYS researchers (Falk, 1975; Falk & Cosby, 1975) as well as research reports which replicated and expanded previous studies (Cosby, et al., 1974).

The third data collection (Wave III) allowed for the analysis of educational and occupational achievement of Southern rural youth (e.g., see: Cosby & Charner, 1978). With the completion of this latest data collection effort, detailed findings of the consequences of social origins, race, sex, school context, ethnicity, aspirations, values, etc. for career achievements was possible. A variety of studies have been completed and, at present, research articles and reports continue to be produced from this three-wave data collection effort. Longitudinal research is a very complex process which must be carefully monitored, coordinated, and organized in all phases. The SYS has been an important source of information on the career achievement process of a "forgotten minority," i.e., rural youth residing in the deep-south. The collection, coding, and analyses of these data required resources from contributing agencies and much hard work on the part of many social science researchers over the years. Hopefully, the SYS will continue to follow respondents throughout their early adult careers so that empirical information will be available for addressing policy issues concerning members of our society who originate from the rural South.

In addition a greatly expanded array of attainment and occupational history information was requested. The major sections of the interview schedule were as follows: (1) Work History Section; (2) Education and Training Section; (3) Family Section; (4) Residence Section; and (5) General Information Section.

Enumeration of Instruments and Measures

A variety of questionnaire items were utilized throughout the SYS project. Listed below are those variables employed in at least one SYS survey. It should be noted that many variables were repeated in subsequent data collection efforts. A special emphasis of the SYS project was the measurement of attitude change and stability of rural youth regarding career-related orientations. A listing of the variables included in the SYS is provided below:

I. Social Origins and Backgrounds (Wave I and II)

- A. Parental breadwinners' occupations
- B. Fathers' education
- C. Mothers' education
- D. Residence
- E. Sex
- F. Race
- G. Religion
- H. Age

II. School Context (Wave I and II)

- A. High school programs
- B. Vocational training
- C. School clubs and activities

III. Interpersonal Relations (Wave I and II)

- A. Sources of occupational encouragement
- B. Role Models for occupational choice

IV. Career Goal Blockages (Waves I, II, III and IV)

- A. Availability of educational facilities
- B. Parental interest
- C. Occupational opportunities
- D. Personal handicaps
- E. Race
- F. Sex
- G. Personal ability

V. Work Values (Waves I, II and III)

- A. To make money
- B. To help other people
- C. To become an important person
- D. Steady employment
- E. To be your own boss
- F. Chance for excitement

VI. Career Orientations (Waves I, II, III and IV)

- A. Occupational aspirations
- B. Occupational expectations
- C. Certainty of career plans
- D. Educational aspirations
- E. Educational expectations
- F. Occupational goal deflection

VII. Career Related Orientations (Waves I, II, III and IV)

- A. Residential aspirations
- B. Residential expectations
- C. Marital plans
- D. Fertility
- E. Plans for military service
- F. Desires for future mates' occupations

VIII. Adult Attainment (Waves III and IV)

- A. Occupational attainment
- B. Occupational satisfaction
- C. Income
- D. Educational attainment
- E. Participation in training programs

IX. Other Adult Behaviors (Waves III and IV)

- A. Place of residence
- B. Marriage
- C. Fertility
- D. Military service

X. Miscellaneous Variables (Wave IV)

- A. Physical handicaps
- B. Birth control practices
- C. Desires for childrens' education

References

- Cosby, A.G. Black-White differences in aspirations among deep-south high school students. Journal of Negro Education, 1971, XXXX, 17-21.
- Cosby, A.G. Occupational expectations and the hypothesis of increasing realism of choice. Journal of Vocational Behavior, 1974, 5, 53-65.
- Cosby, A.G., Byrd, W.L., Dowdell, R.S., Middleton, S., & Kuvlesky, W.P. Youth status projections in the south: Structural annotations of research literature from regional project S-61. Department of Rural Sociology, Texas A&M University, September, DIR. No. 74-78.
- Cosby, A.G., & Charner, I. (Eds.). Education and work in rural America: The social context of early career decision and achievement. Houston: Stafford-Lowden, 1978.
- Cosby, A.G., & Howard, W.G. Residential preferences in America: The growing desire for rural life. Paper prepared for the Rural Development Seminar Series Extension Service, United States Department of Agriculture, (March), 1976.
- Cosby, A.G., & Picou, J.S. Vocational expectations of adolescents in four deep-south states. Vocational Guidance Quarterly, 1971, (March), 177-182.
- Falk, W.W. An integrative model of the occupational choice process. In J.S. Picou & J.E. Campbell (Eds.). Career behavior of special groups. Columbus, Charles E. Merrill Publishing, 1975, 62-92.
- Falk, W.W., & Cosby, A.G. School desegregation and educational projections of rural black youth. Rural Sociology, 1974, 39, 28-41.
- Falk, W.W., & Cosby, A.G. Women and the status attainment process. Social Science Quarterly, 1975, (September), 307-314.
- Jacobs, N.L. The influences of occupational value orientations on the vocational development of East Texas youth: A longitudinal study. Unpublished master's thesis, Texas A&M University, 1973.
- Kuvlesky, W.P., & Bealer, R.C. A clarification of the concept "occupational choice." Rural Sociology, 1966, 31, 265-276.
- Kuvlesky, W.P., & Dameron, J.S. Adolescents' perceptions of military service as a vehicle for social mobility: A racial comparison of rural youth. Journal of Vocational Behavior, 1971, 1, 57-67.
- Kuvlesky, W.P., & Obordo, A.S. A racial comparison of teenage girls' projections for marriage and procreation. Journal of Marriage and the Family, 1972, 34, 75-84.

Kuvlesky, W.P., & Ohlendorf, G.W. A rural-urban comparison of the occupational status orientations of negro boys. Rural Sociology, 1968, 33, 274-283.

Kuvlesky, W.P., & Pelham, J. Place of residence projections of rural youth: A racial comparison. Social Science Quarterly, 1970, 50, 166-176.

Lever, M.F. Selected occupational status projections of southern youth: An analysis by sex, race, and socioeconomic status. Unpublished master's thesis, Texas A&M University, 1969.

Patella, V., & Kuvlesky, W.P. Situational variation in language patterns of Mexican American boys and girls. Social Science Quarterly, 1973, 54, 855-864.

Picou, J.S., Tracy, G.S., & Hernandez, .P.F. Occupational projections of Louisiana black high school seniors. Education and Urban Society, 1970, 2, (August), 459-468.

Picou, J.S., & Howard, W.G. Determinants of career orientations and early educational attainment: Evaluation of a causal model. In A.G. Cosby & I. Charner (Eds.). Education and work in rural America. Houston, Stafford-Lowden, 1978, 138-186.

Picou, J.S., & Cosby, A.G. Social origins, occupational goals and southern youth. Youth and Society, 1971, 2, (March), 307-322.

Picou, J.S., & Curry, E.W. Residence and the athletic participation-educational aspiration hypothesis. Social Science Quarterly, 1974, 55, (December), 768-776.

Thomas, K. Educational orientations of southern rural youth: An analysis of socioeconomic status and racial differences. Unpublished master's thesis, University of Kentucky, 1970.

Thomas, J.K. Ambition and status attainment in the early career development of rural young adults. Unpublished doctoral dissertation, Texas A&M University, 1979.

White, C.D. The social mobility-fertility hypothesis: A racial and class competition among southern females. Unpublished master's thesis, Texas A&M University, 1974.

The Florida Longitudinal Project:

A Review

Jack M. Fletcher

Texas Research Institute of Mental Sciences

Houston, Texas

Paul Satz

University of Victoria

Victoria, B. C., Canada

Robin Morris

University of Florida

Gainesville, Florida

The Florida Longitudinal Project:

A Review

Jack M. Fletcher

Texas Research Institute of Mental Sciences

Houston, Texas

Paul Satz

University of Victoria

Victoria, B. C., Canada

Robin Morris

University of Florida

Gainesville, Florida

Introduction

Learning disabilities have been recognized as a major social and educational problem. Prevalence estimates in North America indicate that approximately 10-15% of school children with at least average intelligence suffer from severe handicaps in reading, writing, spelling and/or math. Follow-up studies have suggested that learning problems which persist into late childhood and adolescence may lead to serious emotional and behavioral disturbances (Eisenberg, 1966; Gates, 1978; Kline, 1972; Schonhaut & Satz, 1980). It has also been suggested that some of the aggressive, antisocial disturbances associated with learning disabilities, particularly in adolescence, may precede the development of schizophrenia in the adult (Menkes, Rowe & Menkes, 1967;

this area promise a tremendous payoff. (p. 1710)

With statements of this sort, it is not surprising that both Task Panels recommended as major research priorities early detection and intervention studies. However, longitudinal research in school age samples must be completed if prevention is to become a reality. Large samples of children must be obtained early in their school career and followed longitudinally for several years (into adolescence). Such a study could include research on issues such as the incidence and long-term outcome of early learning problems. In addition, the developmental precursors of these problems and their correlative outcomes could be studied. Programs could be addressed to the prevention of the primary consequences of learning problems (school failure) and their secondary consequences (emotional and adjustment problems). Finally, clues as to the etiologies of learning problems could result from appropriately designed longitudinal research.

One study which attempted to address some of these issues was the Florida Longitudinal Project (Satz & Friel, 1973; Satz, Friel & Rudegair, 1976; Satz, Taylor, Friel & Fletcher, 1978). This project, which began in 1970, comprised a 9 year longitudinal study of several samples of kindergarten children in Alachua County, Florida and elsewhere. The overall purpose of the Project was to determine some of the developmental precursors, correlates and outcomes of academic failure. Three major objectives were addressed by the Project. Objective I concerned the standardization and validation of a kindergarten screening battery that could predict achievement outcomes later in school. Validation of the battery depended partly on the longitudinal follow-up. Of equal importance was the sensitivity of the battery to cognitive and neuropsychological precursors of learning disabilities. This need constituted Objective II, which concerned developmental processes associated with learning disabilities. Objective II focused on (a) developmental changes in the cognitive correlates of learning disabilities; and (b) a multivariate search for subtypes of learning disabilities. The former research addressed a theory which formed the guidelines for the project (Satz & Sparrow, 1970; Satz & Van Nostrand, 1973; Satz et al., 1978). In recent years, follow-up research within Objective II (developmental changes and subtypes) has forced revision of this theory (Satz & Fletcher, 1980; Satz & Morris, 1980). The subtype research now constitutes the primary thrust of the project. Objective III addressed the incidence and prognosis of learning problems. We would have liked to extend this phase of the research into adolescence. With this extension, issues concerning the relationship of learning problems and other emotional and adjustment problems could be explored. Unfortunately, this was not possible because of funding cut-backs and termination of the project in Year 9 (1978).

Methodological Overview

The Florida Longitudinal Project began in 1970, with collection of data essentially terminating in 1978 (nine years). It comprised detailed neuropsychological, cognitive, neurological, family and behavioral studies of 3 primary samples of children in Alachua County, Florida. As a way of introduction, Table 1 presents a temporal overview of the project and its different components. In the next few sections the composition of the sample and probes will be described. It should be noted that samples in Australia (White, Batini, Satz, & Friel, 1979), Illinois, and Maryland were studied with project methodology, along with additional samples in Alachua County, Florida. Only the 3 basic groups of children comprising the primary thrust of the Project will be presented.

Samples

Sampling methods were designed to circumvent many of the problems characteristic of previous longitudinal research. This was accomplished, in part, by selecting a total population of children who would be at higher risk for learning disabilities (boys) and excluding those children (blacks) who would more likely be culturally disadvantaged and representative of the larger population of general reading and learning disabilities. This research strategy, while biased in terms of population selection, was felt to provide a more homogeneous sample of children, of largely middle class background, who met at least minimal potential criteria for the investigation of specific reading and learning disabilities (Eisenberg, 1966). At the same time, the population was large enough to ensure the identification of subgroups of failing children in later years. The identification of these subgroups could only be obtained in later years after more comprehensive achievement tests, family interviews and neurological examinations were completed. Furthermore, the selection of an additional kindergarten population of white boys in the Fall of 1971 provided a necessary fulfillment for a cross-validation sample which also increased sample size for subsequent attempts to stabilize equations for predictive validation. The selection of a cross-validation group had two other advantages for a long-term prospective study. First, by increasing sample size, it provided a hedge against attrition in later years. Secondly, it ensured a larger pool of children from which subgroups of disabled readers could later be selected. In general, the sampling and follow-up methods were successful, with attrition negligible even after 7 years (10%).

Standardization population (1970). The original Standardization Population consisted of 497 white male kindergarten pupils in Alachua County,

Table 1
Temporal Overview of the Florida Longitudinal Project (1970-1976)*

Sample	1970 Year 1	1971 Year 2	1972 Year 3	1973 Year 4	1974 Year 5	1975 Year 6	1976 Year 7
Standardi- zation N=497	Test Battery (N=497) TCHR CRIT. (N=474)	TCHR Criterion (N=473)	Test Battery (N=420) TCHR CRIT. (N=463) Combined Crit. (N=420)	Neurolo- gical & TCHR. CRIT. (N=459)	Family ** Studies TCHR CRIT. (N=455)	Test Battery (N=181) TCHR CRIT. Combined Crit. (N=175)	TCHR CRIT. (N=)
Cross- Valida- tion I N=181		Test Battery (N=181) TCHR CRIT. (N=)	TCHR CRIT. (N=173)	Test Battery (N=151) TCHR CRIT. (N=175)	Neurological & Family Studies ** TCHR CRIT. (N=168)	TCHR CRIT. (N=159)	Test Battery (N=55) TCHR CRIT. (N=161) Combined Cr (N=55)
Cross- Valida- tion II N=132					Test Battery (N=132) TCHR CRIT. (N=128)	TCHR CRIT. (N=114)	TCHR CRIT. (N=107) Combined Cr (N=)

* All initial test batteries were obtained in the fall. Other test batteries and criteria were obtained in the spring.

** Total N (Standardization and Cross-Validation I) = 175.

Florida. This figure represented 96 percent of the white male population enrolled in the 10 county schools (14 urban, 6 rural) at the time of the testing (1970). Data were collected on 86 days between October and March. In order to determine whether discrepant amounts of schooling affected the test results, Day of Testing designations, ranging from 1 through 86, were assigned to each child. Subjects were tested on an individual basis in an 8' x 35' trailer (equipped with 4 testing modules) which was transported to the grounds of each school. Typically, 4 children were tested in the morning, and 4 in the afternoon. The mean age (in mos.) was 66.2; the range, 57 through 78. Low socio-economic status characterized 9.7 percent of the families of these children. In other words, approximately 90 percent of the Standardization group came from families in the middle to upper socio-economic levels.

Cross-Validation I Sample (1971). In addition to the standardization population, a cross-validation group of white boys (N = 181) was tested in the Fall of 1971 at the beginning of kindergarten. This sample was drawn from the 5 largest urban elementary schools in the county (1970 statistics) and comprised the total population of white male children in these schools. Their mean age (in mos.) was 64.4; the range, 53 through 73. The percentage of low socioeconomic status was smaller in this sample (5 percent) because of the exclusion of the 6 rural elementary schools. Data were collected in this cross-validation sample on 27 school days between October and December.

Cross-Validation II Sample (1974). This sample consisted of 28 black children (13 boys, 15 girls) and 104 white children (54 boys, 50 girls) for a total of 132 children. This group represented all of the children who entered kindergarten in September, 1974, at Stephen Foster Elementary School (mean age = 54.1 mos.). Inclusion of this sample represented an attempt to pilot the battery for large scale screening in groups heterogeneous for sex and race. Consequently, some 20% of this sample fell in the lower socioeconomic classes.

Testing Procedures

As Table 1 reveals, the children received 3 systematic assessments of cognitive and neuropsychological functioning; at the beginning of kindergarten (Year 1), end of Grade 2 (Year 3), and end of Grade 5 (Year 6). The test batteries comprised measures of neuropsychological and cognitive processing skills that are acquired at different rates during childhood. The initial kindergarten battery was modified to ensure adequate test ceilings and floors at Grades 2 and 5 for the Standardization Population and Cross-Validation samples. However, the factor structure was relatively constant at each probe (cf. pp. 10-13). The following variables comprised the basic test

battery:

- (1) Peabody Picture Vocabulary Test (Dunn, 1965). A measure of comprehensional vocabulary requiring a child to point to 1 of 4 pictures, 1 of which represents a word read to him. Score: Mental Age.
- (2) Recognition-Discrimination (Small, 1968). Visual-perceptual (matching to sample) task requiring the child to identify a geometric stimulus design among a group of 4 figures, 3 of which were rotated and/or similar in shape to the stimulus figure. Score: percent of 15 trials correct.
- (3) Beery Visual-Motor Integration (Beery, 1967). An age normed perceptual-motor copying task. Score: age equivalents in months.
- (4) Finger Localization. (Benton, 1959). Somatosensory test consisting of 5 levels of performance, 4 of which (1, 2, 4, and 5) were presumed to assess increasing levels of complexity. (1) Shielded unilateral stimulations were made to the fingertips; The shield was removed between stimulations and the child required to point to the finger touched with the index finger of his free hand. Five trials per hand, starting with the preferred hand. (2) Shielded unilateral stimulations made to the fingertips, The child identified each stimulated finger on a corresponding diagram of an opened hand. Five trials per hand, starting with a preferred hand. (3) Shielded, randomized series of 3 bilateral and 10 unilateral stimulations made to the backs of the child's hands; The child waved hand(s) stimulated. Only bilateral trials were scored. (4) Shielded unilateral stimulations made to the fingertips; The child recalled the number of the finger stimulated. Instructions in the numbering of each hand given immediately before the stimulations to that hand. Five trials per hand, starting with preferred hand. (5) Shielded simultaneous bilateral stimulations made to pairs of fingertips; The child recalled the number of the finger stimulated on each hand. Five pairs of stimulations, starting with preferred hand. Score: percent correct across all 5 levels.
- (5) Alphabet. Recitation of ABC's. Score: number of letters named, regardless of order in which given.
- (6) Auditory-Discrimination. Shortened, taped, version of the Wepman Auditory Discrimination Test (Wepman, 1958). The child was required to recognize on 20 trials whether pairs of words heard through earphones were the same (a single word repeated) or different (two different, but similar sounding words). Score: sum of the ratio of correct "same" responses to the total number of "same" responses and the ratio of correct "different" responses to the total number of "different responses."

(7) Dichotic Listening. (Satz, 1968). Measure of ear asymmetry in which the child was presented with disparate pairs of numbers arriving simultaneously via stereo headphones every half-second. The child was required to recall numbers heard. Version consisted of 30 trials of three-pair digit sequences with an inter-trial interval of 10 seconds for recall. Scores: total recall from both the right and left channels.

(8) Verbal Fluency. Modified version of Verbal Fluency Test described by Spreen and Benton (1965). The child is required to name in 1 minute as many objects as possible in each of three rooms in his home.

(9) Finger Tapping. Timed measure of fine motor movement (Reitan, 1964) requiring the child to rapidly depress a metered key with his index finger; four 10 second trials per hand, starting with preferred hand. Scores: total performance based upon the average sum of preferred and nonpreferred hand performance.

(10) Embedded Figures. Visual-perceptual task requiring the child to identify in which of 3 choice designs a stimulus figure was embedded. Score: percent of 15 trials correct.

(11) Similarities. Verbal reasoning subtest of the Wechsler Pre-school and Primary Scale of Intelligence (Wechsler, 1967). Score: scaled scores.

(12) Auditory-Visual Integration (Birch and Belmont, 1964). Taped test requiring the child to detect the visual equivalents (dots on display cards) of patterned auditory sequences (clicks). Score: percent of 10 trials correct.

(13) Right-Left Discrimination. Ten questions measuring the child's ability to make right-left discriminations on his person. Score: percent correct, half-credits for correct responses following hesitations.

(14) Socio-economic status. Rated by teacher on dichotomized scale as "low," or "average or above."

(15) Day of Testing. Day of evaluation period child was tested.

(16) Behavioral Checklist. Ratings by the examiner on 10 5-point scales of the child's emotional reactivity; degree of irritability, cooperation, and dependency; duration of attention span; goal orientation; response to directions; level and nature of activity; and nature of communication. Score: absolute total of the ratings on the 10 item scales.

(17) Maturity. Child's physical, emotional, and social maturity each rated by teacher as "low" (1), "average" (2), or "high" (3). Score: sum of individual ratings.

(18) Activity level. Rated by teacher as "low" (1), "average" (2), or "high" (3).

This basic battery was administered to the Standardization Population and Cross-Validation I sample at each of the 3 probes. The Cross-Validation II sample received an abbreviated battery (Tests 1-8 and SES) that was not repeated in subsequent years. Some changes were made in individual tests over the years of the project. For the Grade 2 (Year 3) repeat testing, a more advanced form of Verbal Fluency was used. At Grade 5 (Year 6), more advanced forms of the following measures were used to adjust for possible ceiling effects: Similarities, Verbal Fluency, Embedded Figures and Recognition-Discrimination. The Goldman-Friscoe-Woodcock Test of Auditory-Discrimination was substituted for the Wepman. Cross-sectional studies were conducted to ensure the appropriateness of these measures across all ages (Kindergarten-Grade 8). Two tests with ceiling effects, Alphabet Recitation and Finger Localization were not changed because of their use in pediatric neuropsychological examination. All Year 6 tests were validated independently to ensure appropriate ceilings by our colleague, H. Gerry Taylor.

Table 1 shows that the battery was administered to each available child at the beginning of kindergarten and end of Grade 2 for the entire Standardization Population ($N_{kg} = 181$; $N_{G2} = 151$). For the Grade 5 (Year 6) follow-up, only a selected portion of these populations were retested ($N = 236$). These children comprised all of the learning disabled boys from the 2 original groups (1970, 1971) who continue to reside in Alachua County, Florida 6 years later, plus their respective age and school matched controls (primary and often secondary). In this respect, only the learning disabled boys were unselected in this second follow-up evaluation (Year 6). In addition to the 8 variable abbreviated battery, the Cross-Validation II sample received 4 measures designed to measure more specific psycholinguistic functions, a major weakness of the original battery. These measures, which included the Verbal Fluency along with the ITPA Gramatic Closure subtest, the Berry-Talbot Test of Morphological knowledge, and the Double Object Comprehension Test, have been described elsewhere (Fletcher, Satz & Scholes, 1980).

Achievement Criteria

Two types of achievement tests were administered throughout the follow-up years. They comprised teacher-based and test-based criterion assessments. The teacher-based criterion was termed Classroom Reading Level and was administered at the end of each academic year in the project.

It was not a subjective teacher judgement. Rather, the teacher was asked to check on a list the book in which the child was reading (e.g. primer, pre-primer, different grade levels). As such, Classroom Reading Level constituted the teacher's specification of the child's instructional level in reading.

Two different instruments were used for the test-based assessment of achievement. Both were administered in 3 year intervals. For the Grade 2 follow-up, the test-based assessment consisted of the Iota Word Test (Monroe, 1928), a standardized instrument with 53 words read aloud by the child. For the Grade 5 follow-up, the test-based assessment consisted of the Reading Recognition, Spelling and Arithmetic subtests of the Wide Range Achievement Test (Jastak, Bijou & Jastak, 1965). In addition, the Vocabulary and Comprehension subtests of the Gates-McGinitie (1965) were obtained in Grade 2 on 358 children to provide additional validation of the teacher-based and test-based criteria.

The teacher-based criteria, which were obtained annually on each child, permitted follow-up of children even when they moved to other states. The use of the test-based assessment permitted validation of these teacher-based assessments. Correlational analyses revealed a robust relationship between measures (.78-.88) thus lending support for the validation of the Classroom Reading Level criterion (Satz et al., 1978).

The criterion measures were used in different ways to form reading groups. Children were classified into 4 groups (Severely Disabled, Mildly Disabled, Average and Superior) on the basis of Classroom Reading Level alone or on the basis of both Classroom Reading Level and either the IOTA or WRAT (Combined Criterion). In the former, groups were formed by devising cutting lines such that Severely Disabled readers were approximately 2 years below the mean and Superior readers were 2 years above the mean. For the Combined Criterion, teacher- and test-based assessments were averaged into T-scores. Groups were formed according to standard deviation cut-offs of the distribution on T-scores. Severely Disabled readers were below 1 standard deviation of the mean. Mildly Disabled readers scored between 1 and .4 standard deviations below the mean. Average readers scored from .4 below to 1 standard deviation above the mean. Superior readers scored above 1 standard deviation. For Grade 2, 1 standard deviation below the mean corresponded to a reading level of 12 months below age level on the Iota Word Test and a pre-primer instructional level in reading. One standard deviation above the mean corresponded to a reading level 12 months above grade level and a Grade 4 instructional level in reading.

Neurological and Family Studies

In Grades 3 and 4, neurologicals were completed on a subset of the Standardization and Cross-Validation I groups. The blood parents

of these children were also administered reading and spelling tests to assess familial incidence of learning disability.

The neurological examinations were conducted by pediatric residents under the supervision of Dr. John Ross (University of Florida) to learning disabled children and their matched controls in mobile laboratories provided by the project. The examination consisted of the following: (a) a general exam, which assessed cranial nerves, motor responses, sensation, stereognosis, two-point simultaneous discrimination, reflexes, and cerebellar functions; (b) a special exam, which consisted of additional items evaluating fine and gross motor functioning (e.g., alternating hand movements, tapping, associated movements), right-left discrimination, and eye tracking; and (c) an examination of gross body anomalies, or stigmata, of the head, eyes, ears, mouth, and feet. In evaluating the three exams, each of 129 subtest items was rated as normal (0), equivocal (1), or abnormal (2). In addition, the examining neurologist rendered 2 overall clinical judgements for each child, 1 based on the general and special exams and the other on stigmata. Each child was then assigned 1 of 3 neurological designations. Children whose numerical scores, summing over subtest items, were deviant relative to the total group of children who received neurologicals (at least 1 standard deviation from the mean) and who were judged as abnormal on both clinical ratings were designated as affected; those whose total scores were nondeviant and who were judged as abnormal on neither clinical rating were designated as normal; and those who failed to meet criteria for either of the above designations, and who thus received mixed clinical judgements, were designated as borderline-equivocal. It should be noted that this study was single-blind. The residents had no knowledge of group identity nor did they use reading and writing measures for the examination.

Parental competencies in reading and spelling were obtained on blood parents of 175 learning disabled children and 175 controls. 112 mothers and 97 fathers participated. The WRAT Reading and Spelling subtests were employed. Spelling as well as reading was of interest due to the close association of these skills and to the special persistence of spelling problems among adults who have otherwise compensated for early reading disorders (Owen, Adams, Forrest, Stolz, & Fisher, 1971; Rutter, 1978).

Conclusion

It is apparent that more follow-up research on learning disabled children is needed. The ideal follow-up study should use a longitudinal design, starting with children at an age prior to the development of learning disabilities (e.g., starting no later than the beginning of the first grade). The advantages of this type of design have been well

documented by researchers in the area of schizophrenia (Mednick and MacNeil, 1968; Garnezy, 1974). These advantages include: 1) the absence of experimenter bias based on knowledge of which subjects are in the largest group, 2) availability of an ideal control group, 3) standardized measurement on selected variables of interest at the age of interest, and 4) optimal data for making causal inferences.

Future follow-up studies should use a wide array of outcome measures to examine the prognosis of learning disabled children in a variety of areas (e.g., academic, occupation, social, emotional). Structured interviews may be used to collect a broad range of data. Spreen's (1978) procedure of interviewing both children and parents is recommended as a means of getting a maximum amount of information and also as a check on data reliability. Variables such as school suspension, arrests, convictions and drug abuse need to be examined so one can determine that the empirical relationship is, if any, between childhood learning problems and adolescent adjustment (e.g., antisocial behavior). Personality measures should be gathered to see whether learning disabled children do have a higher incidence of emotional problems as adults. Studies of this sort would help to unravel the complex life history question of the relationship between childhood learning problems and primary and secondary outcomes during adolescence.

References

- Ames, L. B. Learning disabilities: the developmental point of view. In H. R. Myklebust (Ed.), Progress in learning disabilities: volume one. New York: Harper & Row, 1966.
- Applebee, A. N. Research in reading retardation: two critical problems. Journal of Child Psychology and Psychiatry, 1971, 12, 91-113.
- Baltes, P. B. Longitudinal and cross-sectional sequence in the study of age and generation effects. Human Development, 1968, 11, 145-171.
- Beery, J. Matching of auditory and visual stimuli by average and retarded readers. Child Development, 1967, 38, 827-833.
- Benton, A. L. Right-left discrimination and finger localization: development and pathology. New York: Hoeber-Harper, 1959.
- Benton, A. L. Developmental dyslexia: neurological aspects. In J. Friedlander (Ed.), Advances in neurology, 1975, 7, 1-41.
- Benton, A. L. Developmental dyslexia. Journal of Pediatric Psychology, 1976, 1, 28-31.
- Benton, A. L. Some conclusions about dyslexia. In A. L. Benton and D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge, New York: Oxford University Press, 1978.
- Bentzen, F. Sex ratios in learning and behavior disorders. American Journal of Orthopsychiatry, 1963, 23, 92-98.
- Birch, H. G. & Belmont, L. Auditory-visual integration in normal and retarded readers. American Journal of Orthopsychiatry, 1964, 34, 852-861.
- Bloom, S. Stability and change in human characteristics, New York: John Wiley & Sons, Inc., 1964.
- Boder, E. Developmental dyslexia: a diagnostic approach based on three atypical reading-spelling patterns. Developmental Medicine and Child Neurology, 1973, 15, 663-687.
- Bradshaw, J. L. Three interrelated problems in reading: a review. Memory and Cognition, 1975, 3, 124-134.

- Bryant, T. G. Report of the president's commission on mental health, Washington, D. C., U. S. Government Printing Office, 1978.
- Caldwell, B. M. The usefulness of the critical period hypothesis in the study of filiative behavior. In N. S. Endler, L. R. Boulter & H. Osler (Eds.), Contemporary issues in developmental psychology. New York: Holt, Rinehart & Winston, Inc., 1968.
- Critchley, M. Dysgraphia and other anomalies of written speech. Pediatric Clinics of North America, 1968, 15, 639-650.
- Cronbach, L. J. & Meehl, P. G. Construct validity in psychological tests. Psychological Bulletin, 1955, 52, 281-302.
- Divoky, D. Education's latest victim: the "L. D." kid. Learning, 1974, 3, 20-25.
- Doehring, D. G. The tangled web of behavioral research on developmental dyslexia. In A. L. Benton & D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge, New York: Oxford University Press, 1978.
- Doehring, D. G. & Hoshko, I. M. Classification of reading problems by the Q-technique of factor analysis. Cortex, 1977, 13, 281-294.
- Dunn, L. Peabody Picture Vocabulary Test, American Guidance Service, 1965.
- Eisenberg, L. The epidemiology of reading retardation and a program for preventative intervention. In J. Money (Ed.), The disabled reader: education of the dyslexic child, Baltimore, Md.: John Hopkins Press, 1966.
- Everitt, B. Cluster analysis. London: Heineman Educational Books, 1974.
- Fennell, E. B. A longitudinal test of the lag theory of developmental dyslexia. Doctoral Dissertation, University of Florida, 1978.
- Feshback, S., Adelman, H. & Fuller, W. W. Early identification of children with high risk of reading failure. American Educational Research Association, New Orleans, 1973.
- Fiske, J. L. & Rourke, B. P. Identification of subtypes of learning disabled children at three age levels: a neuropsychological, multivariate approach. Journal of Clinical Neuropsychology, 1979, 4, in press.

Fletcher, J. M. Linguistic factors in reading acquisition: evidence for developmental changes. In F. J. Pirozzolo & M. C. Wittrock (Eds.), Neuropsychological and cognitive processes in reading. New York: Academic Press, in press.

Fletcher, J., Rice, W. J., & Ray, R. R. Linear discriminant function analysis in neuropsychological research: some uses and abuses. Cortex, 1978, 14, 564-577.

Fletcher, J. M. & Satz, P. Developmental changes in the neuropsychological correlates of reading achievement: a six year longitudinal follow-up. Journal of Clinical Neuropsychology, in press.

Fletcher, J., Satz, P. & Scholes, R. J. Developmental changes in the linguistic correlates of reading disabilities. Brain and Language, in press.

Fletcher, J., Smidt, R. & Satz, P. Discriminant function strategies for the kindergarten prediction of reading disabilities. Journal of Clinical Neuropsychology, 1979, 1, 151-166.

Gaddes, W. Learning disabilities: prevalence estimates and the need for definition. In R. Knights & D. J. Baker (Eds.), The neuropsychology of learning disorders. Baltimore, Md.: Univ. Park Press, 1976.

Gallagher, J. J. & Bradley, R. H. Early identification of developmental difficulties. In I. J. Gordon (Ed.), Early childhood education. Chicago: Univ. of Chicago Press, 1972, 291-303.

Garnezy, N. Children at risk: the search for the antecedents of schizophrenia. Schizophrenia Bulletin, 1974, 1, 14-90.

Gates, A. I. The role of personality maladjustment in reading disability. In G. Natchez (Ed.), Children with reading problems. New York: Basic Books, 1968.

Gates, A. I. & McGinitie, W. Gates-McGinitie Reading Tests. New York: Teachers College Press, 1970.

Gibson, E. J. & Levin, H. The psychology of reading. Cambridge, Mass.: MIT Press, 1975.

Golden, C. & Anderson, S. Learning disabilities and brain dysfunction. Springfield, Ill.: Charles C. Thomas, 1979.

Guthrie, J. T. Reading comprehension and syntactic responses in good and poor readers. Journal of Educational Psychology, 1973, 56, 294-299.

Guthrie, J. T. & Seifert, M. Letter-sound complexity in learning to identify words. Journal of Educational Psychology, 1977, 69, 686-696.

Guthrie, J. T. & Tyler, S. Psycholinguistic processing in reading and listening among good and poor readers. Journal of Reading Behavior, 1976, 8, 415-526.

Hallahan, D. P. Comparative research studies on the psychological characteristics of learning disabled children. In W. M. Crunckshank & D. P. Hallahan (Eds.), Perceptual and learning disabilities, Vol. I. Syracuse, N. Y.: Syracuse Univ. Press, 1975.

Huberty, C. J. Discriminant analysis. Review of educational research, 1975, 45, 543-598.

Jansky, J. J. A critical review of "Some developmental and predictive precursors of reading disabilities: a six-year follow-up". In A. L. Benton & D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge. New York: Oxford Univ. Press, 1978.

Jastak, J., Bijou, S. W. & Jastak, S. R. The wide range achievement test. Wilmington, Del.: Guidance Associates, 1965.

Jastak, J. F. & Jastak, J. The wide range achievement test manual. Wilmington, Del.: Guidance Associates, 1976.

Keeney, A. H. & Keeney, V. T. Dyslexia: diagnosis and treatment of reading disorders. St. Louis: C. V. Mosley, 1968.

Kline, C. L. The adolescent with learning problems: how long must they wait? Journal of Learning Disabilities, 1972, 5, 127-144.

Labouvie, E. W., Bartsch, T. W., Nesselroade, J. R. & Baltes, P. B. On the internal and external validity of simple longitudinal designs. Child Development, 1974, 45, 101-113.

Lieberman, I. Y. & Shankweiler, D. Speech, the alphabet and teaching to read. Conference on the theory and practice of beginning reading instruction, University of Pittsburgh, 1976.

Mattis, S. Dyslexia syndromes: a working hypothesis that works. In A. L. Benton & D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge. New York: Oxford Univ. Press, 1978.

Mattis, S., French, J. H. & Rapin, I. Dyslexia in children and adults: three independent neuropsychological syndromes. Developmental Medicine and Child Neurology, 1975, 17, 150-163.

- Mednick, S. A. & McNeil, T. F. Current methodology in research on the etiology of schizophrenia: serious difficulties which suggest the use of the high-risk group method. Psychological Bulletin, 1968, 70, 681-693.
- Meehl, P. E. & Rosen, A. Antecedent probability and the efficiency of psychometric signs, patterns of cutting scores. Psychological Bulletin, 1955, 52, 194-216.
- Meerlo, J. A. Reading block and television apathy: an alarm for parents. Mental Hygiene, 1962.
- Mendelson, W., Johnson, N. & Stewart, M. A. Hyperactive children as teenagers: a follow-up study. Journal of Nervous and Mental Disease, 1971, 153, 273-279.
- Menkes, M. J., Rowe, J. S. & Menkes, J. H. A twenty-five year follow-up study on the hyperkinetic child with minimal brain dysfunction. Pediatrics, 1967, 39, 393-399.
- Money, J. (Ed.) Reading disability: progress and research needs in dyslexia. Baltimore: Johns Hopkins Press, 1962.
- Monroe, R. Monroe Diagnostic Reading Test. Chicago: Stoelling, 1928.
- Muehl, S., Forell, E. R. A follow-up study of disabled readers: variables related to high school reading performance. Reading Research Quarterly, 1973, 9, 110-123.
- Murray, C. A. The link between learning disabilities and juvenile delinquency. Washington: LEAA, 1976.
- Owen, F. W. Dyslexia-genetic aspects. In A. L. Benton & D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge. New York: Oxford University Press, 1978.
- Owen, F. W., Adams, P. A., Forrest, T., Stolz, L. M. & Fisher, S. Learning disorders in children: sibling studies. Monographs of the society for research in child development. Chicago: Univ. of Chicago Press, 36, 1971.
- Petrauskas, R. & Rourke, B. P. Identification of subgroups of retarded readers: a neuropsychological, multivariate approach. Journal of Clinical Neuropsychology, 1979, 1, 17-37.
- Porter, R. D. & Cattell, R. B. Handbook for the children's personality questionnaire (The CPQ). Champaign, Ill.: Institute for Personality and Ability Testing, 1972.

- Rayner, K. Developmental changes in word recognition strategies. Journal of Educational Psychology, 1976, 68, 323-329.
- Rabinovitch, R. D., Drew, A. L., DeJong, R., Ingram, W. & Withey, L. A. A research approach to reading retardation. Association for Research in Nervous and Mental Disease, 1954, 34, 363-396.
- Rawson, M. Task force on learning failure and unused potential. In T. Bryant, Report of the President's Commission on Mental Health. Washington, D. C.: U. S. Government Printing Office, 1978.
- Reitan, R. Relationship between neurological and psychological variables and their implications for reading instruction. In H. Robinsin (Ed.), Meeting individual differences in reading. Chicago: Univ. of Chicago Press, 1964.
- Robins, L. Deviant children grown up. Baltimore: Williams and Williams, 1966.
- Ross, A. O. Psychological aspects of learning disabilities and reading disorders. New York: McGraw-Hill, 1976.
- Rourke, B. P. Brain-behavior relationships in children with learning disabilities: a research program. American Psychologist, 1975, 30, 911-920.
- Rourke, B. P. Reading retardation in children: developmental lag or deficit. In R. Knights & D. Bakker (Eds.), The neuropsychology of learning disorders: theoretical approaches. NATO Conference, Baltimore: Univ. Park Press, 1976.
- Rutter, M. Prevalence and types of dyslexia. In A. L. Benton & D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge. New York: Oxford Univ. Press, 1978.
- Satz, P. Laterality effect in dichotic listening. Nature, 1968, 218, 277-278.
- Satz, P. Learning disabilities: a "state of the art" paper. Prepared for the President's Commission on Mental Health, 1978.
- Satz, P. Cerebral dominance and reading disability: an old problem revisited. In R. Knights and D. J. Bakker (Eds.), The Neuropsychology of learning disorders: theoretical approaches. Proceedings of NATO Conference, Baltimore: Univ. Park Press, 1976.
- Satz, P., Fennell, E. & Reilly, C. The predictive validity of six neurodiagnostic tests: a decision theory analysis. Journal of Consulting and Clinical Psychology, 1970, 34, 375-381.

- Satz, P. & Fletcher, J. M. Early screening tests: some uses and abuses. Journal of Learning Disabilities, 1979, 12, 43-50.
- Satz, P. & Fletcher, J. M. Minimal brain dysfunction: appraisal of research, concepts and methods. In H. Ric & H. Ric (Eds.), Handbook of minimal brain dysfunction. New York: Wiley Interscience, 1980.
- Satz, P. & Friel, J. Some predictive antecedents of specific reading disabilities: a preliminary one-year follow-up. In P. Satz & J. Ross (Eds.), The disabled learner: early detection and intervention, Rotterdam: Univ. Press, 1973.
- Satz, P. & Friel, J. Some predictive antecedents of specific reading disability: a preliminary two-year follow-up. Journal of Learning Disabilities, 1974, 7, 437-441.
- Satz, P. & Friel, J. The predictive validity of an abbreviated screening battery: a three-year cross validation follow-up. Journal of Learning Disabilities, 1978.
- Satz, P., Friel, J. & Rudegair, F. Some predictive antecedents of specific reading disability: a two-, three- and four-year follow-up. In J. T. Guthrie (Ed.) Aspects of reading acquisition. Baltimore: The Johns Hopkins Press, 1976.
- Satz, P. & Morris, R. Learning disability subtypes: a review. In F. J. Pirozzolo & M. C. Wittrock (Eds.), Neuropsychological and cognitive processes in reading. New York: Academic Press, in press.
- Satz, P., Rardin, D. & Ross, J. An evaluation of a theory of specific developmental dyslexia. Child Development, 1971, 42, 2009-2021.
- Satz, P. & Sparrow, S. S. Specific developmental dyslexia: a theoretical formulation. In D. J. Bakker & P. Satz (Eds.), Specific reading disability: advances in theory and method. Rotterdam: Rotterdam Univ. Press, 1970.
- Satz, P., Taylor, G., Friel, J. & Fletcher, J. Some developmental and predictive precursors of reading disabilities: a six-year follow-up. In A. L. Benton & D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge. New York: Oxford Univ. Press, 1978.
- Satz, P. & Van Nostrand, G. K. Developmental dyslexia: an evaluation of a theory. In P. Satz & J. Ross (Eds.), The disabled learner: early detection and intervention. Rotterdam: Rotterdam Univ. Press, 1973.

- Schonhaut, S. & Satz, P. Prognosis of the learning disabled child: a review of the follow-up studies. Unpublished paper, 1980.
- Schmideberg, M. Search light on delinquency. New York: International Univ. Press, 1949, 174-189.
- Schauer, C. A. The developmental relationship between neuropsychological and achievement variables: a cluster analytic study. Doctoral Dissertation, University of Florida, 1979.
- Scott, J. P. Critical periods in behavioral development. In N. S. Endler, L. R. Boulter & H. Osser (Eds.) Contemporary issues in developmental psychology. New York: Holt, Rinehart & Winston, 1968.
- Silberman, C. Crisis in black and white. New York: Random House, 1964.
- Silver, A. A. Prevention. In A. L. Benton & D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge. New York: Oxford Univ. Press, 1978.
- Small, N. Levels of perceptual functioning in children: a developmental study. Unpublished Masters Thesis, Univ. of Florida, 1969.
- Spreen, O. The dyslexias: a discussion of neurobehavioral research. In A. L. Benton & D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge. New York: Oxford Univ. Press, 1978.
- Spreen, O. & Benton, A. L. Spreen-Benton language examination profile. Iowa City, Iowa: Univ. of Iowa, 1965.
- Stevenson, H. W., Parker, T., Wilkinson, A., Hegion, A. & Fish, E. Longitudinal study of individual differences in cognitive development and scholastic achievement. Journal of Educational Psychology, 1976, 68, 377-400.
- Strag, G. A. Comparative behavioral ratings of parents with severely mentally retarded, special learning disability, and normal children. Journal of Learning Disabilities, 1972, 5, 52-56.
- Taylor, H. G., Satz, P. & Friel, J. Developmental dyslexia in relation to other childhood reading disorders: significance and utility. Reading Research Quarterly, 1979, 15, 84-101.
- Thurstone, L. L. The differential growth of mental abilities. Chapel Hill, N. C.: Univ. of North Carolina Psychometric Lab, No. 14, 1955.

USDE. First Annual Report of the National Advisory Committee on Handicapped Children. Washington, D. C.: U. S. Dept. of Health, Education and Welfare, 1968.

Usprich, C. The study of dyslexia: two nascent trends and a neuropsychological model. The Bulletin of the Orton Society, 1976, 26, 34-48.

Vellutino, F. R. Toward an understanding of dyslexia: psychological factors in specific reading disability. In A. L. Benton & D. Pearl (Eds.), Dyslexia: an appraisal of current knowledge. New York: Oxford University Press, 1978, 61-111.

Waites, L. World federation of neurology: research group on developmental dyslexia and world illiteracy, 1968.

Wechsler, D. Wechsler preschool and primary scales of intelligence. New York: Psychological Corporation, 1967.

Weitz, J. R. & Zigler, E. Cognitive development in retarded and non-retarded persons: Piagetian tests of the similar sequence hypothesis. Psychological Bulletin, 1979, 86, 831-851.

Wepman, J. M. Auditory discrimination test. Chicago: Language Research Associates, 1958.

White, M., Batini, T., Satz, P. & Friel, J. Predictive validity of a screening battery for children "at risk" for reading failure: a 3-year cross-cultural follow-up. British Journal of Educational Psychology, 1979, 49, 132-137.

Wohlwill, J. F. The study of behavioral development. New York: Academic Press, 1973.

Wright, P. W. Reading problems and delinquency. World Congress on Dyslexia, Mayo Clinic, 1974.

Yule, W. & Rutter, M. The epidemiology and social implications of specific reading retardation. In R. Knights & D. J. Bakker (Eds.), The neuropsychology of learning disorders: theoretical approaches. Baltimore: Univ. Park Press, 1976.

Zigler, E. Developmental versus difference theories of mental retardation and the problem of motivation. American Journal of Mental Deficiency, 1969, 73, 536-556.

LONGITUDINAL HEALTH RESEARCH
IN THE U.S. NAVY

E.K. Eric Gunderson, Ph.D.
Head, Environmental and Social Medicine Division
Naval Health Research Center
San Diego, CA 93138

Robert E. Mitchell, M.D.
Captain, Medical Corps, United States Navy
Commanding Officer, Naval Aerospace Medical Research Laboratory
Naval Air Station
Pensacola, FL 32508

and

Robert J. Biersner, Ph.D.
Commander, Medical Service Corps, United States Navy
Program Manager for Human Performances
Naval Medical Research and Development Command
National Naval Medical Center
Bethesda, MD 20014

*The opinions, conjectures and interpretations contained herein are those of the authors only and do not necessarily represent the views, policies or endorsement of the U.S. Navy, the Department of Defense, or any other government agency

Longitudinal Health Research

In the U.S. Navy

E.K. Eric Gunderson, Ph.D.

Robert E. Mitchell, M.D.

Robert J. Biersner, Ph.D.

Introduction

Longitudinal health research has many uses in military medicine. Because disease and injury reduce military performance or the readiness to perform, military commanders must be concerned about health risks to members and casualties that could impair unit effectiveness. The primary mission of the Navy Medical Department is to keep as many active duty personnel physically and mentally fit for duty as possible. Longitudinal research methods provide the means to estimate the distribution and relative risks of disease and injury in the military population so that appropriate medical support can be provided in various settings and to discover the determinants or causes so that prevention and control can be implemented. Thus, longitudinal health research is an essential component of any large-scale medical program to define health care needs and to allocate scarce resources in an optimal fashion.

The Navy has originated and conducted several longitudinal research programs during the last four decades, beginning with the Thousand Aviator Program in 1940. Two other programs, the Active Duty Enlisted Personnel Program and the Submariner/Diver Program, were later innovations. These programs involve different populations, designs and methodologies, and each was initiated to provide data about distinctive health problems. With diminishing resources and closer scrutiny by program managers over the years, the findings obtained from these programs have become more discrete and refined, with less applicability to conditions and groups that are found commonly in the general American population. (The three programs are, however, directed toward answering particular questions about the health effects that may result from having been exposed to specific military situations.) Therefore, the earliest research, the Thousand Aviator Program, was established to determine the etiology of cardiovascular disease among naval aviators, a disease that occurs frequently among American civilians, especially among a civilian group that has much in common with naval aviators

(i.e., commercial airline pilots). The most recent program, the Submariner/Diver Program, involves an analysis of the health effects that may be associated with environmental conditions (e.g., low level radiation, high carbon dioxide levels, absence of sunlight, fresh fruits and vegetables, hyperbaric pressure and exposure to high levels of trace metals) that will probably never be experienced by any American civilian group. In addition, the characteristics of the samples used in these programs restrict broad applications to the civilian sector. The samples are weighted in favor of males who have above average intelligence. (Some correction for these biases is possible, however, in the Active Duty Enlisted Program described below.)

The three programs differ in the techniques used for acquiring and analyzing data. The Thousand Aviator Program adheres to a traditional methodology, representing a prospective analysis of primary (medical examination) data collected directly from project volunteers. The Active Duty Enlisted Personnel Program and the Submariner/Diver Program collate secondary data (personnel and hospitalization records, environmental measures, deployment schedules, and so forth) that are usually obtained indirectly (through archival sources such as computer or microfiche records). The Active Duty Enlisted Personnel Program, however, has the capability of being either prospective or retrospective, while the Submariner/Diver Program is retrospective only.

The Navy Inpatient Medical Data System. Beginning with data for 1965, the Naval Health Research Center, San Diego, has constructed a computer file of inpatient medical data for active duty Navy personnel. This file contains hospitalization records, medical board actions (to determine fitness for duty and/or compensation for disability), and deaths. Each hospitalization record contains demographic and identifying information, admitting facility, dates of admission and discharge, primary and secondary diagnoses, whether or not the condition existed prior to enlistment, occupational specialty and pay grade, marital status, and medical disposition. The file is updated periodically as copies of new records are received from the Naval Medical Data Services Center, Bethesda, Maryland. A unique feature of this file is that the various record systems are compiled into individual medical histories that represent morbidity and mortality over entire naval careers and during retirement. In addition to the medical history file, a parallel service history file has been constructed for enlisted personnel. This file contains basic personnel data such as age, sex, race, marital status, education, and aptitude scores, and a chronological history of important changes in personnel status, such as promotions, demotions, unauthorized absences, desertions, discharges and reenlistments. With this system, disease and disability histories for individuals or groups can be documented continuously and collated with military service histories.

Archival medical data used in the Active Duty Enlisted Personnel Program and Submariner/Diver Programs are a unique and important source of health information. Inpatient data are especially useful because of (a) the

completeness, uniformity, and generally high accuracy of the diagnostic data, (b) linkage to personnel records, and (c) ease of following service members throughout entire careers (20 to 30 years) and the retirement period. The limitations of such data for general epidemiologic purposes are (a) restrictions as to age and sex, and (b) sample biases resulting from induction screening standards, early discharge for diseases and injuries that interfere with military performance, and early attrition of non-career personnel. Record linkage of archival medical data to retired or active duty personnel (service history) records is made possible by routine assignment of a unique number, the Social Security Number, to the pertinent records.

The reliability of the medical data system can be verified by matching and cross-checking medical files against the personnel data files. Records that match on Social Security Number also must agree with respect to demographic and military service information. Errors found in this matching process can often be corrected on the basis of the internal consistency of information.

The integration of medical and personnel files has made possible comparative analyses of incidence, days lost, disposition, recurrence of illness or injury, disability, and death for various disease and accident categories. Risks of morbidity and mortality can be determined for any Navy occupation, work environment, or segment of the Navy population. For example, the file permits comprehensive and long-term analyses of disease and injury rates by age, sex, race, occupational specialty, job experience, pay grade, aptitude level, education, and duty assignment/geographic location. The relationship of dynamic etiological factors such as change in occupational specialty, promotions, demotions, disciplinary problems, marriage and children, isolated or unusual duty assignments, training failure, and so forth with health and morbidity can be evaluated by means of cohort analyses conducted over a decade or more of naval service. While this system is unique to the Naval Health Research Center, the data are available for use by, or integration with, other longitudinal health systems (e.g., submarine personnel--see below).

Active Duty Enlisted Personnel Program

The first example of longitudinal research using active duty enlisted personnel is shown in Table 1. This table provides a comparison of overall inpatient morbidity by sex (Hoiberg, 1978). The comparison for men and women shows generally higher rates for women for most of the disease categories -- about 2.5 to 1 overall. The largest differences among categories applicable to both groups were found for neoplasms, mental disorders, and the infectious/parasitic disease. The leading causes for hospitalization of women during this period were pregnancy-related conditions. The admission

rate for pregnancy-related conditions rose dramatically from the 1960s to the late 1970s -- from about 1,100 per 100,000 in 1966-1969 to 6,400 in 1976. This non-cohort epidemiologic study has implications for re-directing health care resources during the next decade if the proportion of women in the Navy continues to grow.

The next example to be presented is a prospective longitudinal cohort analysis that determines whether or not the presence of hemoglobinopathies diagnosed during Navy recruit training is related to subsequent health and performance among Black sailors. Before describing this analysis, distinguishing features of longitudinal cohort analyses will be explained further. The essential characteristics of the longitudinal cohort analysis are: (a) the cohort group or groups are defined before the occurrence of disease, and (b) the groups are observed over a specific period in order to determine frequencies of disease occurrence. A distinction is further made between prospective and retrospective cohort analyses. In the prospective cohort analysis, the incidence of the subsequent disease is not known prior to defining the cohort. In the retrospective analysis, relevant events (causes and effects) have already occurred before the analysis is initiated (MacMahon & Pugh, 1970). Differences do not exist, however, between retrospective and prospective cohort analyses in terms of methodology, scientific interpretations and practical outcomes, assuming the investigator adopts a fixed design or plan for testing hypotheses before beginning the analysis. Generally, for longitudinal analyses using the Navy Inpatient Medical Data System, labeling the analysis prospective or retrospective has little significance because the categories and measures of disease and performance to be used are usually standardized or fixed before the analysis begins. However, this distinction is important in defining statistical methods to be used in determining the significance of differences between samples.

In this prospective cohort example, 8,725 Black Navy recruits were screened for hemoglobin variants and glucose-6-phosphate dehydrogenase (G-6-PD) deficiency in 1972. Based on the results of these blood analyses, the recruits were grouped into four categories: Sick trait (N = 599), G-6-PD deficient (N = 1,003), both trait anomalies (N = 73), and neither trait or normal (N = 7,050).

In 1978, the Naval Health Research Center medical and service history files of these recruits were searched for appropriate follow-up data, and the four groups were compared on a number of demographic characteristics, indices of military performance and health (Hoiberg, Ernst & Uddin, 1979). Comparisons of the four groups on demographic and pre-service characteristics -- age, marital status, verbal intelligence (as measured by General Classification Test scores), and education -- indicated that differences among groups were not significant for any of these variables. Table 2 presents data on indicators of military performance and medical disability and death. Again, significant differences were not apparent. Finally, rates of total hospitalization for the four groups were compared and found

not to differ. Only for the category "Diseases of the Blood and Blood-forming Organs" (including hemoglobinopathies -- a confounding variable) was a significant effect found. Men with both traits were represented disproportionately in this diagnostic category.

The conclusion of this longitudinal cohort analysis was that examination of the subsequent military performance and health status of the three trait groups compared with a normal sample of Black sailors failed to differentiate the groups in terms of illness rates or impaired performance. This evidence indicated that these genetic trait anomalies were benign. A similar analysis is planned for Marine Corps personnel. Inasmuch as physical exertion is more severe and frequent for Marines than for Navy personnel, the results could conceivably be different than those presented above.

This example demonstrates clearly that cohort analyses provide the basis for more exact estimates of health risks than non-cohort analyses. Results such as those for the sickle cell analysis can usually be translated directly into recommendations and decisions for establishing or modifying policies related to occupational health and safety if significant effects are found. Other examples of longitudinal cohort analyses of Navy enlisted personnel are those conducted by Hoiberg on Vietnam combat personnel (Hoiberg, 1979a), pregnant women (Hoiberg & Ernst, 1979b) and selected occupational groups (Hoiberg & Pugh, 1978).

Submariner/Diver Program

Submariners. A systematic analysis of morbidity and mortality patterns among submarine personnel has been of fairly recent origin. With the introduction of ballistic missile nuclear submarines (SSBNs) in the mid- to late 1950's, interest developed in possible health effects associated with prolonged exposure to this closed environment (50-75 days for each patrol/two or three patrols a year). Nuclear submariners are exposed to a number of unusual conditions, including high carbon dioxide levels (an average level about 1% through the mid-1960's, to a current average of about 0.5%), absence of sunlight, lack of fresh fruits and vegetables, and exposure to high levels of some trace metals. These conditions are present 24 hours a day throughout the patrol period. In addition, chronic exposure to low levels of ionizing radiation has emphasized the importance of longitudinal analyses of health.

Initial analyses of health patterns among active duty submarine personnel during the 1960's demonstrated that submariners were as healthy as, or healthier than, other Navy groups (Austin, 1964; Wilken, 1969). These morbidity rates were determined from the medical section of patrol reports written during a seven-year period (1960-1967). (These reports do

not document illnesses or injuries during off-patrol periods, which account for about two thirds of the yearly cycle.) While morbidity varied substantially across major diagnostic categories, the rate within each category was low. (Comparison groups were not used.) A later analysis of Navy inpatient medical records for the years 1966-1969 (Hester, 1971) showed that submariners had lower morbidity rates for most major diagnostic categories than other Naval personnel (including Marine Corps). Submariners were similar to the comparison sample in rates of tuberculosis, malignant and benign neoplasms, metabolic disorders (including obesity) and diseases of the peripheral nervous system. These findings were essentially replicated in a later analysis (Tansey, Wilson & Schaeffer, 1979) which showed that submariners had lower rates of respiratory, traumatic, gastrointestinal, dermal, infectious and miscellaneous illnesses, and higher rates of genitourinary, systemic (including mononucleosis), cranial, and neuropsychiatric illnesses compared to personnel stationed onboard surface vessels (e.g., destroyers). These data were also collected from sick call records maintained while on patrol. While differences in reporting procedures, periods of reporting (1968-1973 for submariners, 1973 only for surface personnel), and age differences between the samples (surface personnel were several years younger than the submariners) could have mitigated these differences, the overall effects are nonetheless consistent with previous results.

The major problem with the above analyses is that the effects have been derived from acute situations among young, active duty personnel. None of these analyses have involved samples in which long-term, latent health effects could be manifest. To correct this deficiency, a prospective survey of submariner health was established in 1969. The survey involved collecting multiphasic physical and psychological examination data on over 1,000 active duty submariners from the Groton, Connecticut area. While the protocol and techniques used to collect these data have been described thoroughly (Sawyer & Baker, 1972; Tansey, 1974), the rationale for the survey was ambiguous. The only possible health consequences of interest that were described were those involving cardiovascular disease (Sawyer & Baker, 1972). Indeed, most of the physical examinations were centered around these effects, with most of the remaining tests used to document sensorimotor status.

In order to obtain a sample of personnel who have had long-term exposure to nuclear submarine environments, and among whom latent health effects, if any, have had a sufficient period in which to develop, only data for retired and Fleet Reserve personnel were included for analysis. (The Fleet Reserve consists of enlisted personnel who are placed in a retainer status after 20 years of active service, and remain in this status until 30 years of service have been completed, after which they are fully retired.) Major identifying information (name, address, Social Security Number, age submarine qualification and rating/pay grade) will be obtained from the Naval Reserve Personnel Center. Early estimates are that this file will contain approximately 40,000 former submarine personnel. In addition,

records of several thousand submariner deaths (Casualty Reports) will be available from the Naval Manpower and Personnel Command. These records (retired personnel and deaths) will be collated with hospitalization data, including Navy inpatient medical data, as well as inpatient data from Army, Air Force and Veterans Administration hospitals (which retired Navy personnel may use for convenience). Inasmuch as some retired personnel may not have access to military hospital facilities, they may obtain private inpatient care through the Civilian Health and Medical Plan for the Uniformed Services (CHAMPUS). These admissions, too, will be included as part of the project. In addition, Navy Medical Board and Physical Evaluation Board reports will be used for information about personnel with disability retirements.

Length of nuclear submarine service and records of the cumulative low-level radiation dose over an entire submarine career will be obtained from magnetic tape files of dosimetry (measures of radiation exposure) records documented on a form called NAVMED 6470/3, and maintained at the Navy Bureau of Medicine and Surgery. Possible relationships between disease-specific morbidity/mortality rates and selected occupational specialties (e.g., those who have access to the nuclear reactor) will be explored.

Disease-specific mortality rates will be developed initially from Casualty Reports and Naval Reserve Personnel Center master records for inactive personnel. Preliminary exploration indicates that diagnostic information in the Casualty Reports is of generally high quality. These records are, however, subject to the common limitations of death certificates. For deaths among retired and Fleet Reserve personnel associated with admissions to Army/Air Force, Veterans Administration or CHAMPUS hospitals, information relevant to cause of death beyond that appearing in the Casualty Reports will be available.

The data can be classified into three types: outcome variables, exposure variables, and extraneous or confounding variables. Outcome variables can relate to either morbidity or mortality. For each individual, mortality variables would be the presence or absence of death from a specific disease or condition, or perhaps a continuous response such as level of blood pressure.

Inasmuch as this project is a recent development, data are not yet available. Reports of the initial analyses are expected in early 1982.

Drivers. While a formal program to analyze health patterns among members of the Navy diving community does not exist presently, the data resources for such a program are readily available. While the Active Duty Enlisted Personnel and Submariner programs described above are severely constrained by the availability of detailed exposure data (either because exposure has not been documented routinely or because such data are classified - especially in the case of submariners), the occupational activities of Navy divers

re well documented. Highly detailed and reliable descriptions of diving exposures have been maintained since 1970 by the Naval Safety Center, Norfolk, Virginia. In that year, both routine and experimental dives were documented on standard record sheets using simple alphanumeric entries. If the dive resulted in an accident (i.e., decompression sickness, air embolism, drowning, and so forth), this information is also entered onto the same form using a standard format. This record, called the "Combined Diving Log - Accident/Injury Report" (OPNAV Form 9940/1), replaces two previous forms containing extensive narrative descriptions and non-standard data entries. The present diving record is designed to be easily keypunched and entered into computer files.

Information is provided about diver characteristics, the diving environment, protective clothing and life support equipment, the dive schedule and profile, and accident parameters (if an accident occurred in association with the dive). Subsets of information within these categories are codified and entered onto the diving record sheet and mailed quarterly to the Navy Safety Center. (Detailed descriptions of these subsets may be obtained from the overlays--OPNAV Form 9940/1A--used in conjunction with OPNAV 9940/1 or from the U.S. Navy Diving Manual, volume 1, "Air Diving", both of which may be obtained from the Supervisor of Diving, Naval Sea Systems Command, Washington, D.C. 20362.)

While little use has been made of the above data for longitudinal analyses, such use is planned in the near future.

Thousand Aviator Program

The oldest longitudinal health program still ongoing in the United States is the Thousand Aviator Program. In addition to being the senior research program of this type, the Thousand Aviator Program had the distinction (and advantage) of involving participants who were comparatively young, and who were a more homogeneous group medically, psychologically and socially than is true of most other longitudinal programs. Measurement variability is therefore restricted, resulting in improved reliability and validity of the findings. In addition, most of the participants were career officers, thereby ensuring adequate long-term follow-up.

While in the later years of this program emphasis has been on the medical consequences (especially the cardiovascular effects) of aging, the program was initiated originally to select aviator candidates for flight training because such training was lengthy and expensive, and reliable and easily administered screening and selection methods had long been sought. After initial interest during the first World War, little additional medical screening research was conducted until 1939. In that year, the Committee on Selection and Training of Civilian Aircraft Pilots of the National

Research Council received funds from the Civil Aeronautics Authority (now Federal Aviation Agency) for use in planning and supervising aviation medical research. In the summer of 1940, the Council expanded this charter to include military aviation and, in cooperation with the U.S. Navy, began the program that became known as "The Pensacola Study of Naval Aviators," or the "Thousand Aviator Program."

In 1940 program explored the use of psychological and physiological testing in prediction of flight training success. Criteria included (a) passing or failing the flight course, and (b) negative actions by the Commandant's Board (a panel of officers that reviewed the course grades and performance of marginal candidates). The program was designed to determine as quickly as possible (in anticipation of combat needs) the measures deemed most promising for selection of candidates for flight training. The original plans called for participant follow-up only through the flight training course.

The first group to be examined consists of incoming cadets and officers in each flight class at Pensacola during the period from July 16 through September 20, 1940 (classes 147 through 151; N = 479). Twelve cadets and officers were tested each day during the ground school period prior to flight training. The officers in this sample were Naval Academy graduates who were waiting to be commissioned.

Because of the small number of failures in the above group, a decision was made subsequently to extend these examinations. From October 1 to December 15, 1940, a representative sampling comprising about one-fifth of each incoming class (classes 152 through 159) was tested. During this period, only five participants could be tested each day because of reductions in the research staff. From January 1 to May 15, 1941, cadets from classes 160 through 165 were examined who appeared before the Commandant's Advisory Board (for a screening interview). This second sample (classes 152-165; N = 750) consisted of men who had either a high school diploma or two years of college, but who would be commissioned on completion of flight training. The members of both samples were between 20 and 30 years of age and had passed some form of medical screening prior to admission to the flight program. In addition, they had completed a preliminary ten-hour flight training course, including solo flight, before being sent to Pensacola for more extensive training. Personnel who failed the medical screening and preliminary flight training were not examined for the Thousand Aviator Program. A group of 83 instructors at the Naval Air Station (Pensacola) were also examined so as to obtain normative data for pilots known to be successful. The average age of this group was 27 years, and they had accumulated an average of 1,500 flight hours. The physical biochemical and psychological measures administered to these groups during this initial evaluation, as well as measures used during subsequent evaluations (in 1951, 1957, 1963 and 1969) are presented in Table 3.

Following the end of World War II, the usefulness of this large, homogenous group of healthy young men for a unique, prospective analysis of the aging process became apparent. Re-evaluation of the living members of the group who had undergone evaluation in 1940 became the major goal of the program. A follow-up analysis of this group after an interval of ten or more years could be a source of much useful information for development of future diagnostic and treatment methods. The first follow-up analysis (in 1951) was therefore designed to estimate the current physical status of the group with particular emphasis on the cardiovascular system, morbidity and mortality rates, and whether or not aviation experience was related to these effects. Findings were to be compared with the data collected in 1940-1941. Of the 1312 participants examined in 1940-1941, positive identification and/or location was made of 1056 participants in 1951. The vast majority of those not positively located or identified had failed aviator training and follow-up information was unavailable.

Identification of important physiologic precursors of disease, especially cardiovascular disease, requires longitudinal analysis. Although several epidemiologic programs with these goals have been initiated since 1940, these programs involve primarily middle-age samples. Hence, an important phase in the pathogenesis of the disease has necessarily been neglected -- namely, subtle physiologic differences during the young adult years. These early physiologic differences, and associated environmental factors, cannot be determined in programs confined to older age groups. In addition, only from long-term programs involving young, healthy participants can normal standards be developed that permit diagnosis of asymptomatic, sub-clinical disease. The results of serial examinations provided a unique opportunity to determine the relationship of blood pressure and electrocardiograms to aging and to the onset of cardiovascular disease (Packard, Graettinger & Graybiel, 1954; Harlan, Osborne & Greybiel, 1962; Harlan, Osborne & Graybiel, 1964). These research findings, as well as program redirection, were the major outcomes associated with the 1951 follow-up analysis. These men had now reached an age at which direction of latent disease was highly probable. With the interest and support generated by previous evaluations, thorough examinations could now be conducted at the Naval School of Aviation Medicine, in contrast to the previous follow-up examinations that were conducted by a team of Navy physicians who traveled about the country in a mobile laboratory unit. This capability enabled the research staff to perform a more detailed physiological appraisal not possible in the previous two evaluations. The participants were provided commercial air transportation to Pensacola for two days of extensive testing, during which every important physiological measurement included in the earlier examinations was repeated. In addition, tests that were not possible to conduct in the field, or that were not previously available, were used. Much of this additional support and interest came from the Heart Disease Control Program of the U.S. Public Health Service.

Preliminary questionnaires were sent to 815 participants during the latter part of 1962 requesting information concerning recent health,

occupation, flying and military status. Actual testing began in January, 1963 and continued until early 1965. Tests and procedures were standardized as much as possible. Many of the forms were designed for computer analysis. The major findings of this evaluation, used in conjunction with previous data, were a determination of the (a) serum lipid status of the group and the usefulness of this measure as an indicator of cardiovascular health (Harlan, Osborne & Graybiel, 1963), (b) independent contributions made by various physiologic measures in prediction of cardiovascular health (Oberman, Doll & Greybiel, 1964; Harlan, Graybiel & Osborne, 1965), (c) correlations of ballistocardiograms and systolic blood pressure with precursors of cardiovascular disease (Jackson, Oberman, Mitchell & Greybiel, 1967; Oberman, Lane, Harlan, Graybiel & Mitchell, 1967), and (d) distribution of vestibular performance across various age groups (Fregly, Oberman, Graybiel & Mitchell, 1968).

The most recent evaluation occurred in 1969-1971. In this 29th year of the program, most of the participants were entering the sixth decade of life. The 1969 program was conducted in a fashion similar to the 1963-65 program in order to maintain uniformity and standardization of procedures. Beginning in 1968, participants were contacted, sent preliminary questionnaires, and subsequently tested for two days at the Naval Aerospace Medical Institute in Pensacola, Florida. The findings of this evaluation were used to describe pulmonary function and blood pressure patterns over the course of 30 years (Mitchell, 1972; MacIntyre, Mitchell, Oberman & Graybiel, 1977). As of 1977, only 95 of the participants had died of non-military illnesses and injuries, a rate less than half that expected of a random sample of white American males born during the same period (MacIntyre, Mitchell, Oberman, Harlan, Graybiel & Johnson, 1978). Of the 1,056 original participants positively identified in 1951, 682 were known to be alive in 1977. An additional 23 were located through VA or IRS sources. The status of 28 could not be determined. A total of 228 had died of military-related aviation accidents during World War II and the Korean conflict. The death rate was especially low for non-aviation accidents, about 10% of the expected number. The low death rate among the participants was attributed to a higher than normal socio-economic background, better than average intelligence, superb health maintenance and physical fitness, high intelligence, and a genetic predisposition to longevity.

Future research in this program will be directed toward redefining the medical screening standards for older (45+), active duty aviators in order to determine if continued flight duty is possible. This extension of duty is necessitated by a severe shortage and poor retention of younger aviators. These new standards, however, must be based not only on future health risks, but on the ability of these personnel to perform a number of aviation tasks satisfactorily. The Federal Aviation Administration may use these standards for screening commercial pilots for retirement as well. In future years, physical and psychological examination of the participants in the Thousand Aviator Program will be curtailed, but morbidity and mortality data will continue to be collected routinely through questionnaires.

Brief mention should also be made of another longitudinal health program involving mostly aviators--that directed toward the repatriated prisoners of the Vietnam War (RPOWs). The RPOWs, who were released from incarceration in 1973, consisted of 141 Navy and 37 Marine Corps men, 140 of whom were aviators. The aviator subgroup is being examined annually using the physical and psychological measures listed in Table 4. In addition, a comparison group of aviators matched for age, rank, and flight experience is being tested using the same measures. To date, the two groups have been found not to differ significantly on a number of personality measures (as determined from responses to the Jackson Personality Research Form A). The comparison group was, however, found to suffer a more significant high frequency hearing loss than the RPOW group. A number of other comparison data will be available in the near future.

Future Directions

Longitudinal research in the U.S. Navy will most likely become dominated by computer-based technology during the next five to ten years. This direction is necessitated by scarce resources, including both skilled personnel and fiscal support. Additionally, such direction makes optimum use of unique Navy data files and the availability of long-term follow-up medical and personnel information. The Navy is most likely to be interested in determining the incidence and prevalence of rare diseases contracted as a result of unique occupational exposures, especially those diseases associated with ionizing, microwave and extremely low frequency radiation, heavy metal contamination, long-term exposures to carbon monoxide and dioxide, as well as atmospheric hydrocarbons and fluorides, and disease states related to chronic hyperbaric exposures (particularly the medical effects of decompression sickness). These rare conditions necessitate the use of computer-based epidemiology instead of serial medical examinations. The causal factors involved in these diseases can be determined more precisely through environmental monitoring and medical examination after the computer-based system has identified significant diseases and the major characteristics of those who are at risk. In this manner, computer-based longitudinal research will serve as a planning and screening technique, permitting improved allocation and direction of resources in determining disease casualty.

Cost-effective use of computer-based epidemiology requires that corrections and improvements be made to the present system. New epidemiologic methods are needed to provide the basis for more comprehensive and dynamic models of morbidity. Fundamental difficulties with the most commonly used epidemiologic method, the case-control method, have been noted above. Another limitation of classical epidemiologic methods is the tendency to limit analyses to the effects of single variables. Methods are needed that estimate the effects of a number of causal variables or risk factors simultaneously. Also, methods should document trends over long periods,

including variations in both risk factors and illness rates.

The limitations of classical epidemiology become most apparent in trying to account for the many possible risk factors that are now presumed to be related to health and longevity. Many more factors could probably be added from the epidemiology literature. Unfortunately, methods and techniques are not available to represent these complex conditions in any valid, integrated way.

Specific examples of dynamic risk factors in naval settings include transfers from one occupation to another or from one duty station to another; implementation of effective occupational health and safety programs; variations in smoking or drinking habits; and participation in weight reduction or physical conditioning programs. Aging, of course, represents a ubiquitous and important risk factor. For example, on one current longitudinal health analysis, illness incidence was found to be markedly higher in the third decade of naval service than in the second decade among career enlisted personnel (Hoiberg, 1979b). Higher risk with age is more pronounced in occupational specialties such as Construction, Engineering, and Service ratings for disease categories such as Circulatory Diseases, Neoplasms, and Endocrine, Nutritional, and Metabolic Disorders, than in other occupations or diseases.

Important tasks ahead include better delineation of causal factors (including dynamic factors) in disease and injury incidence, and the construction of more comprehensive and valid morbidity models to represent the health of groups or populations in terms of dynamic risk factors and subsequent illness.

Table 1

Incidence Rates for Major Disease Categories by Sex^a

<u>Diagnostic Category</u>	<u>Men</u>	<u>Women</u>
Infectious and Parasitic	770	2,673
Neoplasms	155	738
Endocrine, Nutrition, and Metabolic	98	208
Blood, Blood-forming Organs	31	117
Mental Disorders	1,539	2,754
Nervous System, Sense Organs	411	416
Circulatory System	372	444
Respiratory System	1,499	3,426
Digestive System	1,034	1,702
Genitourinary System	633	2,288
Pregnancy-related Conditions	--	4,433
Skin and Subcutaneous Tissue	754	867
Musculoskeletal System	869	1,337
Congenital Anomalies	135	292
Symptoms and Ill-defined Conditions	489	1,747
Accidents, Poisonings, Violence	2,010	2,904
Supplementary Classifications	321	1,247
Total incidence rate	11,010	27,693
Average population	464,582	13,143

^aIncidence rates are the numbers of hospital admissions per 100,000 enlisted population per year for the period 1973-1975.

Table 2
Comparison of Hemoglobinopathy Groups on Military Performance and Health Indicators

Performance Variables	Sickel Cell Trait (SCT)		G-6-PD Deficient		SCT & G-6-PD Deficient		Normal	
	Mean	S.D.	Mean	S.D.	Mean	S.D.	Mean	S.D.
Promotions	1.78	1.28	1.79	1.25	1.68	1.28	1.74	1.26
Demotions	.30	.59	.30	.55	.38	.62	.30	.56
Unauthorized Absence	.47	1.01	.49	1.09	.38	.76	.49	1.10
Desertions	.10	.39	.10	.41	.04	.20	.12	.46
Highest Pay Grade	2.74	1.28	2.68	1.30	2.64	1.41	2.68	1.33
<u>Effective Status</u>	<u>f</u>	<u>%</u>	<u>f</u>	<u>%</u>	<u>f</u>	<u>%</u>	<u>f</u>	<u>%</u>
On active duty/reenlisted	97	16.2	159	15.8	15	20.5	1,152	16.3
Completed enlistment successfully	245	40.9	400	39.9	30	41.1	2,603	36.9
<u>Noneffective Status</u>								
Not recommended for reenlistment	97	16.2	156	15.6	8	11.0	1,182	16.8
Unsuitability Discharge	59	9.8	104	10.4	9	12.3	753	10.7
Disciplinary separation or in deserter status	76	12.7	143	14.3	10	13.7	999	14.2
<u>Medical Separations</u>	20	3.3	39	3.9	1	1.4	328	4.7
<u>Deaths</u>	5	0.8	2	0.2	0	0	33	0.5
Total	599	99.9	1,003	100.1	73	100.0	7,050	100.1
Hospitalization rate (per 100,000 per annum)	12,211		11,375		11,207		12,615	

Table 3
Medical Evaluation Tests Used for the Thousand Aviator Program^a

Tests	Evaluation Period				
	1940	1951	1957	1963	1969
<u>Interview: Personal and medical histories^b</u>	*	*	*	*	*
<u>Physical examination</u>	c	*	*	*	*
<u>Cardiovascular</u>					
routine electrocardiogram	*	*	*	*	*
startle electrocardiogram	*				
computer processed electrocardiogram				*	*
exercise electrocardiogram			*	*	*
ballistocardiogram			d	*	*
vectorcardiogram				*	*
plethysmogram				*	*
cold pressor test	*		d		*
other	*			*	
<u>Laboratory determinations</u>			*	*	*
<u>Pulmonary and metabolic</u>					
spirometry	*			*	*
basal metabolic rate	*				
other	*			*	
<u>Anthropometry</u>					
somatotype	*			*	*
measurements (in addition to height & weight)			e	*	*
<u>Teleoroentgenograms</u>		*	*	*	*
<u>Psychologic-Psychomotor</u>					
Builford-Zimmerman Temperament Survey				*	*
ataxia test	*			*	
tilt chair	*			*	
other	*			*	
<u>Vision</u>	*			*	*
<u>Neurophysiologic</u>					
electroencephalogram	*			*	
skin resistance	*				
<u>Audiometry</u>				*	*

^aCompletion of the tests is noted by an asterisk; if a procedure was not performed during an evaluation, the appropriate column is blank.

^bPersonal and medical histories include alcohol, smoking, and exercise histories.

^cOnly blood pressures were recorded because each member had qualified medically before inclusion in the study.

^dExaminations performed on less than 25 percent of the study group.

^eArm circumference only.

Table 4
Medical Evaluation Tests Used for the RPOW Program

Tests	Evaluation Period					
	1974	1975	1976	1977	1978	1979
<u>History (including demographics)</u>	*	*	*	*	*	*
<u>Physical examination</u>	*	*	*	*	*	*
anthropometries	*	*	*	*	*	*
<u>Laboratory</u>						
hematology	*	*	*	*	*	*
total eosinophile counts	*	*	*	*	*	*
urinalysis	*	*	*	*	*	*
SMAC 17 or 24	*	*	*	*	*	*
glucose tolerance test	*	*	*	*	*	*
stool studies	*	*	*	*	*	*
O&P, occult blood	*	*	*	*	*	*
total lipids (electrophoresis)						
phenotype				*		
HDL					*	
total proteins (electrophoresis)	*	*	*	*	*	*
immunoglobulins	*					
thyroid profile	*					
<u>X-rays</u>						
cardiac series	*					
PA, lateral (chest)		*	*	*	*	*
flat plate abdomen	*					
hands	*	*				
<u>Electrocardiograms</u>						
routine fasting	*	*	*	*	*	*
stress (Bruce protocol)	*	*	*	*	*	*
<u>Vectrocardiograms</u>	*	*	*	*	*	*
<u>Ballistocardiograms</u>	*	*	*	*	*	*
<u>Pulmonary function studies</u>	*	*	*	*	*	*
<u>Audiograms</u>	*	*	*	*	*	*
<u>Vestibular and special visual studies</u>	*					
<u>Special consultation^a</u>	*	*	*	*	*	*
<u>Psychiatric interview</u>	*	*	*	*	*	*
<u>Psychological testing</u>	*	*	*	*	*	*

^aAs indicated.

References

- Austin, D.A. A review of the medical experiences of seventy Polaris patrols, November 15, 1960 - November 15, 1963. Submarine Medical Officer Qualification Thesis, Naval Submarine Medical Center, Groton, Connecticut, 1964.
- Biersner, R.J. Factors in 171 Navy diving decompression accidents occurring between 1960-1969. Aviation, Space, and Environmental Medicine, 1975, 46, 1069-1073.
- Biersner, R.J., Dembert, M.L. & Browning, M.D. The aging diver: Do the older become bolder? Faceplate, 1978, 9, 19-21.
- Biersner, R.J., Dembert, M.L. & Browning, M.D. The antisocial diver: performance, medical and emotional consequences. Military Medicine, 1979, 144, 445-448.
- Biersner, R.J., Dembert, M.L. & Browning, M.D. In search of super-diver. Faceplate, in press.
- Fregly, A.R., Oberman, A., Graybiel, A. & Mitchell, R.E. Thousand aviator study: Non-vestibular contributions to postural equilibrium functions. Aerospace Medicine, 1968, 39, 33-37.
- Harlan, W.R., Jr., Osborne, R.K., & Graybiel, A. A longitudinal study of the blood pressure. Circulation, 1962, 26, 530-543.
- Harlan, W.R., Jr., Osborne, R.K. & Graybiel, A. Factors determining serum lipid concentrations. Report Number 869, Naval School of Aviation Medicine, Pensacola, Florida, 1963.
- Harlan, W.R., Jr., Osborne, R.K. & Graybiel, A. The prognostic value of the cold pressor test and the basal blood pressure based on an eighteen year follow-up study. American Heart Journal, 1964, 13, 683-387.
- Harlan, W.R., Jr., Graybiel, A. & Osborne, R.K. Determinants of cardiovascular disease in a young population. American Journal of Cardiology, 1965, 15, 1-12.
- Hester, R. Provisional medical statistics for personnel attached to nuclear powered submarines: FY66-69. Report Number 674, Naval Submarine Medical Center, Groton, Connecticut, 1971.
- Hoiberg, A. Causes of hospitalization for enlisted Navy men and women. Report Number 78-1, Naval Health Research Center, San Diego, California, 1978.
- Hoiberg, A. Military effectiveness of Navy men during and after Vietnam. Report Number 79-13, Naval Health Research Center, San Diego, California, 1979a.
- Hoiberg, A. Personal communication, 1979b.
- Hoiberg, A. & Ernst, J. Cancer among Navy personnel: occupational comparisons. Report Number 79-57, Naval Health Research Center, San Diego, California, 1979a.
- Hoiberg, A. & Ernst, J. Motherhood in the military: conflicting roles for Navy women. Report Number 79-31, Naval Health Research Center, San Diego, California, 1979b.
- Hoiberg, A. & Ernst, J. Cancer among Navy personnel: Incidence and mortality. Military Medicine, in press.
- Hoiberg, A., Ernst, J. & Uddin, D.E. Effects of sickle cell trait and G-6-PD deficiency on health and performance among black navy enlisted. Report Number 79-46, Naval Health Research Center, San Diego, California, 1979.
- Hoiberg, A. & Pugh, W.M. Predicting Navy effectiveness, expectations, motivation, personality, aptitude, and background variables. Personnel Psychology, 1978, 31, 841-852.
- Hunter, W.L., Biersner, R.J., Sphar, R.L. & Harvey, C.A. Aseptic bone necrosis among U.S. Navy divers: Survey of 934 nonrandomly selected personnel. Undersea Biomedical Research, 1978, 5, 25-36.
- Jackson, D.H., Oberman, A., Mitchell, R.E., and Graybiel, A. Factors contributing to the ballistocardiographic waveform in healthy middle aged adults. American Journal of Cardiology, 1967, 20, 531-540.
- Kinney, J.A.S., Luria, S.M., McKay, C.L., and Ryan, A.P. Vision of submariners. Undersea Biomedical Research, 1979, 6, (Supplement: Preventive Aspects of Submarine Medicine), S163-S173.
- MacIntyre, N.R., Mitchell, R.E., Oberman, A., and Graybiel, A. Long-term pulmonary function patterns in the aviator: The thousand aviator study. AGARD Conference Proceedings, Number 232, NATO, Neuilly sur Seine, France, 1977.
- McMahon, B. & Pugh, T.E. Epidemiology: Principles and methods. Boston: Little, Brown and Co., 1970.
- Mitchell, R.E. The thousand aviators: Aging and the blood pressure. AGARD Conference Proceedings, Number CP129, NATO, Neuilly sur Seine, France, 1972.

Oberman, A., Doll, R.W. and Graybiel, A. Interdependence among some factors associated with coronary heart disease. Report Number 887, Naval School of Aerospace Medicine, Pensacola, Florida, 1964.

Oberman, A., Lane, N.E., Harlan, W.R., Graybiel, A. and Mitchell, R.E. Trends in systolic blood pressure in the thousand aviator cohort over a twenty-four year period. Circulation, 1967, 36, 812-821.

Packard, J.M., Graettinger, J.S. and Graybiel, A. Analysis of the electrocardiograms obtained from 1000 young healthy aviators: Ten-year follow-up. Circulation, 1954, 10, 384-400.

Sawyer, R.N. and Baker, J.H. The longitudinal health survey: I. Description. Report Number 733, Naval Submarine Medical Research Laboratory, Groton, CT, 1972.

Tansey, W.A. The longitudinal health study: A multiphasic medical surveillance program for U.S. Navy submarine and diving personnel. Report Number 786, Naval Submarine Medical Research Laboratory, Groton, Ct, 1974.

Tansey, W.A., Wilson, J.M. and Schaefer, K.E. Analysis of health data from 10 years of Polaris submarine patrols. Undersea Biomedical Research, 1979, 6, (Supplement: Preventive Aspects of Submarine Medicine), S217-S246.

Tappan, D.V., Jacey, M.J., Heyder, E. and Harvey, C.A. Biochemical and hematologic profiles of 1000 submariners. Undersea Biomedical Research, 1979, 6 (Supplement: Preventive Aspects of Submarine Medicine), S191-S199.

Wilken, D.D. Significant medical experiences aboard Polaris submarines: A review of 360 patrols during the period 1963-1967. Report Number 560, Naval Submarine Medical Center, Groton, CT, 1969.

The Veterans Administration Normative Aging Study*

Raymond Bossé, Ph.D.

David J. Ekerdt, Ph.D.

Jeremiah E. Silbert, M.D.

The Normative Aging Study, VA Outpatient
Clinic, 17 Court Street, Boston, Massachusetts 02108

*Supported by the Medical Research Service of the Veterans Administration and in part by the Council for Tobacco Research, Grant #455, by the National Institute on Aging, Grants #AG00467-05 and AG 02287-01, and by the National Cancer Institute, Grant #1 R18 CA29479-01. The authors wish to acknowledge the assistance of their colleagues in preparing this chapter: Gary A. Borkan, A.J. Garvey, Robert Glynn, Leonard Poon, Charles Rose and David Sparrow.

The Veterans Administration Normative Aging Study*

Raymond Bössé, Ph.D.

David J. Ekerdt, Ph.D.

Jeremiah E. Silbert, M.D.

The Normative Aging Study, VA Outpatient

Clinic, 17 Court Street, Boston, Massachusetts 02108

The Veterans Administration, (the VA), as the largest single health care delivery agency in the United States, is involved with the care of some 25 million veterans. Of this veteran population, approximately 3 million are currently 65 years of age or older. By the year 2000, however, the population of veterans aged 65 and over will have grown to 8 million. This means that the VA could be involved in the care of 28% of all people 65 and over in the population, or 67% of all males in that age range. Aware of its responsibility for an increasingly older population of veterans, the VA has directed special attention to the study and understanding of the aging process. One step toward the fulfillment of that mission was the establishment of the Normative Aging Study (NAS).

The NAS is a longitudinal and interdisciplinary study of aging in men, established in 1963 as an intramural research project of the Veterans Administration (Bell, Rose & Damon, 1966, 1972; Rose, 1965; Rose & Bell, 1965). Historically, the NAS evolved at the Boston Veterans Administration Outpatient Clinic from an earlier study of Spanish-American War Veterans. One hundred and fifty veterans aged 80-90 were treated in a combined clinical and research program starting in 1958. It became apparent, however, that to identify predictors of good health in old age effectively it was not sufficient just to study older individuals, but that a younger group also needed to be studied as they aged. This gave rise to the idea of a longitudinal study of healthy veterans of various ages, and led to the initiation of the NAS.

A total of 6,000 men were screened for possible entry into the Study; 2,280 healthy males aged 21-81 at time of entry were eventually admitted. The NAS participants have physical examinations once every five years,

except for men over 52 years of age who return once every three years for their physical examinations. The anthropometric measurements and special senses testing are on a five-year cycle regardless of age.

The first cycle of data collection spanned the years 1963-1968, while the next two cycles were roughly between 1969-1974 and 1975-1980. The Study is currently completing the third cycle of data collection; about one-quarter of the men have been examined for the fourth time.

Description of the Population

Age

On September 1, 1979, the mean age of the active participants in the NAS was 54.6 years; 262 men (13%) were aged 65 and older, while 680 men (34%) were between the ages of 55 and 64. Because of the age structure of the panel, its value as a resource for the study of aging will increase over the next ten years and beyond as greater and greater numbers attain older ages.

Health

The NAS sample was designed to investigate aging using a gerontologic approach, which views aging as a developmental phase of life. Because there is an increase in the risk of disease with the passing years and because changes due to disease can mask or mimic actual aging processes, normal aging is most effectively observed in healthy individuals. In order to differentiate the aspects of age change from those of disease, we selected only healthy, community-dwelling men for entry into the Study.

Although no individual is free from all signs or symptoms of disordered function, volunteers were screened for acceptance into the Study according to criteria that ensured an initial state of good health and promised the expectation of continued health as far as possible. Laboratory, clinical, radiological and electrocardiographic evaluations formed the basis for such judgements. Primary causes for disqualification included the history or presence of such chronic conditions as heart disease, cancer, peptic ulcer, gout, and recurrent asthma, bronchitis and sinusitis. Also disqualifying was clinical evidence of hypertension and glucose intolerance. Acceptable conditions included childhood diseases (such as rheumatic fever or kidney infection) that had not forbidden prior military service. Other acceptable conditions included history of hepatitis, malaria, jaundice, or anemia so long as no sequelae were present and functions were intact.

Isolated abnormal findings were related to the whole complex of laboratory and clinical observations. When such findings were not compatible with serious disease or impaired function, they were not considered to be disqualifying. These included minor arthritis, obesity, a previous renal

stone or a currently elevated erythrocyte sedimentation rate. Also acceptable were behaviors and habits (such as minor psychiatric complaints, large consumption of alcohol or tobacco) that did not appear to interfere with the individual's activities of daily living. Certain disqualifying criteria were established for particular areas of interest to the NAS. For example, gross defects of limbs or disabling spinal injuries were considered disqualifying so as not to bias anthropometric investigations. The research program in special senses required the disqualification of persons with cataracts, chronic internal eye disease or impaired hearing in both ears.

Since their selection into the Study, participants have begun to differentiate into groups which have developed disease and groups which have not. Despite the need for healthy participants, men who develop disease serve an important purpose in providing comparison groups to be looked at along with healthy members of the sample. Thus, we are able to observe the effects of aging itself in the healthy men and the effects of aging overlaid with disease in the others. In addition, our longitudinal observations of laboratory, physical, social and behavioral characteristics enable us to conduct prospective epidemiologic analyses of the precursors of disease and to evaluate risk factors in our sample.

Social Characteristics

In order to have the findings apply to as wide a population as possible a concerted effort was made to recruit volunteers from a broad socioeconomic range (Rose & Bell, 1965). The effort resulted in a population with the following characteristics:

Education: 14% less than high school, 25% high school graduates, 35% beyond high school and 26% college graduates. Members of the cohort tend to be more educated than their peers in the behavioral population.

Occupation: 23% profession, 21% managers and proprietors, 8% clerical and sales, 16% craftsmen, 28% service workers (mostly policemen and firemen) 2% operative and 2% laborers.

Our population is from higher socioeconomic levels than the general local population. One advantage of our population is that it shows no decrement by age in the average educational attainment and white collar occupation. These special characteristics of our panel erase the general demographic relationship between older age and lower social class levels, a relationship which usually confounds cross-cohort comparisons in adult samples.

Ethnicity: Ethnically, our population accords fairly closely with the distribution found in the Boston area. However, less than 2% of the volunteers are Black, which is lower than the 3.7% figure for Black adult males in the Boston metropolitan area (U.S. Bureau of the Census, 1973).

Geographic Stability: Because of the plan for lifelong participation, volunteers were selected to be geographically stable in order to assure their availability for taking the recurring tests and examinations. For this reason participants were preferentially recruited from occupations or companies with low rate of turnover. The social characteristics of our eventual panel do indeed suggest geographic stability: 82% were born in the Boston area, 71% lived in a single-family dwelling, 92% were married, 84% had children of school age, and 8% planned to remain with their present employers until retirement.

Population Maintenance

Any constraints on the representativeness of the NAS population are offset by our ability to retain our sample over time. In prospective epidemiological studies, priority is given to the accurate description of within-individual change, which requires effective maintenance of the original population. There are three features of the NAS that maximize population size over time: (1) initial selection for health promises greater longevity among participants; (2) selection for geographic stability assures their availability for repeated testing; and (3) the participants, as volunteers, are highly motivated and interested.

Our success at retaining the original NAS population over time is shown in Table 1. During a period of 16 years, only 12.4% of the initial group has been lost to the Study for all reasons, an average of less than 1% per year. A detailed evaluation of attrition in the NAS (Rose, Bosse & Szretter, 1976) indicated that the loss of volunteers has not altered the compositional characteristics of the initial population. Men lost to the Study were not significantly different from those who remained in health, socioeconomic status, job satisfaction and other measures. Only the deceased, as might be expected, were significantly older than those who remained in the Study. Not all participants who move away cease their participation in the Study. Five percent of the remaining population now live outside of the New England area but continue to answer questionnaires and continue to keep scheduled examinations when they return to the Boston area for business or family visits.

Areas of Investigation

The NAS research program involves a wide range of variables in the biomedical and psychosocial domains. Measures and frequency of data collection within domains are described below. In addition to the areas of investigation conducted by NAS staff, the Study also has two associated research programs: The Dental Longitudinal Study and the Mental Performance and Aging Laboratory. Both programs are located at the VA Outpatient Clinic and historically have been closely associated with the NAS by sharing some personnel, space and equipment.

Biomedical Areas

The biomedical research program, including ophthalmology, audiology and dental medicine, is concerned with the accurate description of changes in physical or medical function over time. Related research goals include the identification and evaluation of the precursors and risk factors of age-related disease. The aim of the anthropometric research program is to describe the physique changes that accompany the aging process and to evaluate the role of body composition as a predictor of disease.

1. Medicine

The periodic medical examination includes: medical history, physical examination, chest and hand x-ray, pulmonary function tests and electrocardiogram. Over 30 standard blood and urine tests are made, including complete blood count, liver function, kidney function, glucose, lipids and protein electrophoresis. Along with the self-administered Cornell Medical Index, participants also answer a questionnaire on salt intake, medication, specific illnesses and allergies. Diseases and physical conditions uncovered during the medical examination, as well as diseases and hospitalizations occurring since the previous exam, are coded according to the adapted International Classification of Disease (USDHEW, 1968).

2. Special Senses

Systematic collection of data in audiology, vision, and taste perception has existed since the inception of the NAS. Among the measures taken as part of the ophthalmological examinations are tests of retinal fields, visual acuity, depth perception, pupil reaction, muscle action, a fundus examination, and a test for glaucoma. The audiology measures include tests for air and bone conduction thresholds, speech reception thresholds, and speech discrimination. In addition to the specific tests taken by an optometrist and an audiologist, both eyes and ears are examined by an ophthalmologist and/or otologist if need be. Also, eyes, ears, and nose are examined periodically by a physician as part of the physical examination. Measures of taste perception are collected in conjunction with the Dental Longitudinal Study.

3. Dental Study

The Dental Longitudinal Study is an associated research program of the NAS investigating oral health and disease among a subgroup of 1,200 NAS participants.

4. Anthropometry

Anthropometric measurements, like the special senses data, have been collected every five years since the beginning of the Study. A battery of 44 measurements includes height, weight, skinfolds, and a variety of circumferences, lengths and breadths which serve to delineate the physique of each participant. For many years the anthropometric measurements were taken by consultants associated with Harvard University. More recently these measures have been taken by members of the NAS staff under the supervision of its own physical anthropologist.

Psychosocial areas

Psychosocial data are collected in order to study life course changes in their own right and also to involve these data as control or mediating factors in epidemiological analyses. The social characteristics and circumstances of individuals are represented by standard demographic data, as well as by information on changing behavioral, familial, work and retirement statuses. The research programs in personality and cognitive functioning investigate developmental changes of structure, function and performance, their relation to disease and longevity, and the effects of illness on traits and abilities. Alcohol consumption and tobacco smoking are two behavioral patterns studied for their changes with age and their direct implications for health and aging.

1. Demographic data

Basic demographic data such as education, occupational history, income, marital status, family composition, type of residence, etc. have been collected routinely as part of specific batteries; others have only been updated on occasion as part of questionnaires which focus on specific topics such as smoking, drinking, or retirement. A few measures such as ethnic origin, military service and parental characteristics were collected only at the initial screening because responses are not likely to change over time.

2. Work and Retirement

Information on work and retirement was first gathered on the NAS population during the first cycle of data collection. In 1975 a comprehensive questionnaire on work and retirement was administered to the entire sample. The questionnaire had separate versions for workers and for retirees, and covered experiences, plans, and attitudes relative to work and retirement. The questionnaire was administered again three years later in 1978 with plans to repeat its administration every three years.

3. Personality and Cognitive Functioning

The NAS administered various tests of personality dimensions and cognitive functioning in the course of the first cycle of data collection. The Cattell 16 Personality Factors Questionnaire (16PF) was administered to about half of the NAS participants. The same individuals also took the General Aptitude Test Battery, the Allport-Vernon-Lindzey Study of Values, and the Strong Vocational Interest Blank. Subsequently a short form of the 16PF was administered to a subset of the original sample which took the longer version.

In 1976 a short version of the Eysenck Personality Inventory developed by Floderus (1974) and called the EPI-Q was administered by mail to the entire population. At about the same time a number of other tests were administered to subsets of the population. These included such tests as the Bradburn Affect Balance Scale, the Beck Hopelessness Scale, the Holmes-Rahe Social Readjustment Scale, and several others.

It should be noted that the Cornell Medical Index (CMI), which has been administered systematically at every cycle, has a number of items which provide indices of emotional well-being. Finally, the retirement questionnaires have included two measures of life satisfaction and happiness: the Neugarten-Havighurst-Tobin Life Satisfaction Index and the Cantril Self-Anchoring Scale.

4. Memory and Decision Making

The Mental Performance and Aging Laboratory is an associated research program of the NAS funded in part by the VA and in part by the National Institute of Health. Part of the research involved studies of normal age changes in memory and decision making in a subgroup of 250 NAS participants and a comparison sample of community adults. Cognitive aging is examined in a battery of six studies which evaluate changes in information retrieval from primary, secondary and tertiary memory. Research participants were first examined in 1971-1973 and again in 1977-1979.

5. Drinking

Data regarding drinking habits relative to coffee, tea and especially alcohol were first collected toward the end of the first cycle in 1968. Subsequently, a more detailed questionnaire was developed and administered to the entire NAS population by mail in 1973. A second administration of this instrument will occur during 1980.

6. Smoking

The development of the smoking research program at the NAS was aided by funds from the Council for Tobacco Research. Two detailed smoking behavior questionnaires were administered by mail to the entire NAS population in 1973 and 1976. Two additional basic questionnaires on smoking behavior are administered routinely. One is administered when NAS participants are given their periodic medical examinations, the other at the time of their periodic anthropometric measurements.

Biomedical Findings

Medicine

By 1979, 1,500 participants had completed three cycles of medical examinations encompassing a period of ten years. This has allowed us to initiate a series of analyses which can be considered to be prototypes of more extensive future analyses based on additional longitudinal data.

Earlier reports from the NAS analyzed cross-sectional age differences in clinical laboratory variables (Burney & Bonus, 1972a, b; Koff, Garvey, Burney & Bell, 1973) as well as a 5-year longitudinal changes in these measures (Burney & Enslein, 1972). Recently, we have completed the analysis of the 10-year changes in 31 of the 42 laboratory variables.

Medical and epidemiological studies of NAS participants have also involved behavioral variables. Using 10-year follow-up data on 1,500 participants, psychological scales from the Cornell Medical Index were used to predict subsequent onset of myocardial infarction, angina pectoris and hypertension (Garvey & Sparrow, 1979), while controlling for unknown risk factors such as age, systolic blood pressure, cholesterol, smoking, and glucose intolerance. Multiple logistic regression analyses indicated that behavioral measures of exhaustion, worry, inability to rest, and anger were related to the incidence of hypertension. The results provide support for the belief that some psychosocial factors contribute directly to the development of coronary heart disease and hypertension.

The NAS occasionally conducts pilot studies to assess the promise of new domains for our core data collection. An example was a pilot study of the influence of age on anterior pituitary function in healthy men, which compared various hormone levels of 44 young men (aged 31-44) and 42 older men (aged 64-88).

Another aspect of the NAS research program concerns collaborations with other laboratories. An example is a project investigating age-related differences in responsiveness of the sympathetic nervous system.

Mortality information would be the logical endpoint of the NAS data collection effort. As yet, it is too soon to initiate research on differential aging processes and longevity since only 95 participants (4.2%) of the original sample are known to have died (Table 1). The age of death ranged from 30 to 82, with a mean age of 51. Leading causes of death were ischemic heart disease (32 cases), cancer (26 cases) and accidents, poisoning and violence (10 cases). Although the routine autopsy of NAS decedents would be desirable, there are cost and logistic problems which have delayed such procedures.

Special Senses

A number of NAS publications have resulted from the evaluation of cross-sectional age-related differences in vision (Bell, 1972a; Bell, Wolf & Bernholtz, 1972; Wolf, 1967; Wolf & Gardiner, 1965; Wolf & Schraffa, 1964; Harrison & Wolf, 1964), one in taste perception (Kapur, Glass, Loftus, Alman & Feller, 1972) and three in hearing (Bell, 1972b; Koff, Bell & Garvey, 1973; Welch, Luteran & Bell, 1969). The NAS has longitudinal data on change in taste perception, speech perception, sound thresholds and clinical assessments of the fundus and structure of the eye. However, longitudinal analyses have not been initiated. A summary of NAS findings and the future directions for research in the area of vision have been reviewed by Fozzard, Wolf, Bell, McFarland and Podolsky (1977).

Anthropometry

One of the primary goals in collecting anthropometric data has been to describe the physique changes which accompany the aging process. Descriptions of physique changes which accompany aging have included a cross-sectional report (Damon, Seltzer, Stoudt & Bell, 1972) and a five-year longitudinal follow-up (Friedlaender, Costa, Bosse, Ellis, Rhodes & Stoudt, 1977).

Functional Age

It is a common observation that some people appear much younger or older than would be expected based on their chronological age. It has been suggested that such discrepancies reflect differences in individual rates of developmental change. The concept of biological age, apart from chronological age, has led to effective and widely used techniques in the study of child growth. Similar techniques for assessment of the aging process have been slower to appear, and have not achieved their potential.

Investigators at the NAS have approached this problem from the standpoint of functional age, with the view that different aspects of aging may progress at different rates within the same individual. Six functional age domains were proposed: biochemical, auditory, anthropometric, cognitive ability, personality, and social age. These studies

were described in series of publications which appeared in one issue of Aging and Human Development (Vol.3, 1972).

Present and planned research in functional age at the NAS is directed at analyzing the extent to which difference in longitudinal rates of change between individuals may be attributable to genetic or environmental sources.

Smoking Research Program-Biomedical

Tobacco smoking has been directly implicated in the development of numerous diseases, and as a major cause of premature illness and death. Consequently, it has become almost impossible to study any aspect of health and aging without taking into account differences between smokers and non-smokers (Bosse, Costa, Cohen & Podolsky, 1975; Bosse, Sparrow, Garvey, Costa, Weiss & Rowe, in press; Sparrow, Bosse & Weiss, 1979; Seltzer, 1974; Garvey, Bosse & Seltzer, 1974; Bosse, Garvey & Costa, in press).

Psychosocial Findings

Psychosocial Aspects of Smoking and Aging

An interesting finding of the smoking behavior data is the differential response rate of current, former, and never smokers to a smoking questionnaire (Seltzer, Bosse & Garvey, 1974). A detailed smoking history questionnaire was mailed to the 2,000 volunteer members of the NAS. Although a total response rate of 96% was achieved after six months, a comparison of response rates showed that significantly few cigarette smokers had returned their questionnaires after 30 days and 60 days from the original mailing date. By the same token, heavier smokers were the slowest to respond overall. Cigarette smokers would therefore be under-represented if the data collection had ended after the 30-day and 60 day intervals, spans which are comparable to the duration of most surveys. Our findings suggest that sample bias may be introduced in cohort studies of smoking and health. This bias may become significant if there is additional difference in the health status of responding smokers and responding nonsmokers. The findings of a Kaiser Permanente mail survey (Oakes, Friedman & Seltzer, 1973) did show a tendency for ill or disabled smokers to respond more quickly than well smokers.

The Normative Aging Study has also engaged in a number of descriptive studies of the current, former, and never smokers within the NAS population. These studies include differences in demographic, social class, family, occupational, and health behavior characteristics (Rose, 1972) and differences in anthropometric and somatotype features between cigar and pipe smokers (Seltzer, 1972).

Personality and Cognitive Functioning

In order to obtain a smaller set of measures which could be used to describe personality globally and identify personality groups, scale scores of the Cattell 16PF were cluster-analyzed (Costa & McCrae, 1976). Three broad dimensions of personality resulted: Anxiety, Extraversion, and Openness to Experience.

The Anxiety-Adjustment dimension was used in a study of self-reported health (McCrae, Barton & Costa, 1976). In another study, the psychological portion of the Cornell Medical Index (CMI) was factor-analyzed (Costa & McCrae, 1977) in an attempt to identify psychiatric symptom-dimensions among the NAS males. Two of the six factors were identified as clinical and normal forms of anxiety and both showed significant correlations with the anxiety cluster scores.

The second cluster, Extraversion-Intraversion, has been particularly useful in interpreting and extending our psychological understanding of occupational interests. This was shown in a paper which related the second cluster to the Strong Vocational Interest Blank and to the Allport-Vernon-Lindsay Study of Values (Costa, Fozzard & McCrae, 1977).

The third cluster, labelled Openness to Experience, is of particular interest because it showed age-related differences in structure (Costa & McCrae, 1976).

Memory and Decision Making

Findings of age-related changes in basic memory processes were utilized to suggest improvement for memory diagnosis batteries (Erikson, Poon & Walsh-Sweeney, 1980; Poon, 1980) and to incorporate memory remediation procedures with the aged (Poon, Fozard & Treat, 1978; Treat, Poon & Fozard, 1978; Poon, Walsh-Sweeney & Fozard, 1980).

The Retirement Process

Occupational retirement is in part an artifact of economic policy and not a necessary aspect of human aging. It has, nevertheless, become a normal feature in the lives of older men. The prospective design of the NAS allows us to observe the retirement process in an occupationally heterogeneous sample and to describe the personal changes of circumstance and identity that accompany this transition. Moreover, by following multiple cohorts through the retirement experience we are able to view the historical variability of retirement, as its forms and practices are modified by emerging social, economic and labor market conditions.

As greater numbers enter retirement, we have begun to combine medical and psychosocial data in longitudinal studies of retirement and health. These include specifying the effect of retirement on physical health

while controlling for pre-retirement states, as well as evaluating the occupation-specific "disability" among the increasing numbers of men claiming disability retirements.

Table 1

Attrition in the Normative Aging Population, 1963-1979.

	<u>N</u>	<u>% of initial population</u>
Total Initial population	2280	100.0
Attrition	282	12.4
Deceased	95	4.2
Lost interest	159	7.0
Can't locate	28	1.2
Remaining population	1998	87.6

A complete list of Normative Aging Study publications is available from the Veterans Administration Outpatient Clinic, 17 Court Street, Boston, MA 02108

References

- Anders, T.R. & Fozard, J.L. Effects of age upon retrieval from primary and secondary memory. Developmental Psychology, 9:411-415, 1973.
- Bell, B., Rose, C.L. & Damon, A. The Veterans Administration Longitudinal Study of Healthy Aging. The Gerontologist, 6(4):179-184, 1966.
- Bell, B. Retinal field shrinkage, age, pulmonary function, and biochemistry. Aging and Human Development, 3:103-109, 1972 (2).
- Bell, R. Auditory functional age. Aging and Human Development, 1972(b), 3, 183-187.
- Bell, R., Rose, C.L., & Damon, A. The Normative Aging Study: an interdisciplinary and longitudinal study of health and aging. Aging and Human Development, 3, 5-17, 1972.
- Bell, B., Wolf, E. & Bernholtz, C.D. Depth perception as a function of age. Aging and Human Development, 3:77-81, 1972.
- Bosse, R., Costa, P.T., Cohen, M. & Podolsky, S. Age, smoking inhalation, and pulmonary function. Archives of Environmental Health 30:495-498, 1975.
- Bosse, R., Rose, C.L. & Ekerdt, D.J. Adjustment to approach of retirement. XIth International Congress of Gerontology Abstracts. Scimed Publications, Akosade, Minato-Ku, Tokyo, 1978.
- Bosse, R., Garvey, A.J. & Glynn, R. Age and smoking addiction: a cross-sectional and longitudinal analysis. The Gerontologist, 18 (5, Part II): 52, 1978.
- Bosse, R., Sparrow, D., Garvey, A.J., Costa, P.T., Weiss, S.T. & Rowe, J.W. Cigarette smoking, aging, and decline in pulmonary function: A longitudinal study. Archives of Environmental Health. In press.
- Bosse, R., Garvey, A.J., & Costa, P.T. Predictors of weight change following smoking cessation. International Journal of Addictions. In press.
- Burney, S.W. & Bonus, L. Cross-sectional assessment of laboratory variables in a healthy male population. I. Decade age group differences. Aging and Human Development, 3:83-88, 1972(b).
- Burney, S.W. & Bonus, L. Cross-sectional assessment of laboratory variable in a healthy male population. II. Cigarette smoking and laboratory values. Aging and Human Development, 3:89-94, 1972(b).
- Burney, S.W. & Enslin, K. Investigation of changes in clinical laboratory tests related to aging and smoking. Aging and Human Development, 3:95-101, 1972.

- Burr, W.A. & Damon, A. Physique and the eosinophil count. Human Biology, 42:202-205, 1970.
- Cattell, R.B., Eber, H.W. & Tatsuoka, M.M. The Handbook for the Sixteen Personality Factor (16PF) Questionnaire. Champaign, IL., Institute for Personality and Ability Testing. 1970.
- Costa, P.T., Fozard, J.L., McCraw, R.R. & Bosse, R. Relations of age and personality dimensions to cognitive factors. Journal of Gerontology, 31:663-669, 1976.
- Costa, P.T. & McCrae, R.R. Age differences in personality structure: cluster analytic approach. Journal of Gerontology, 31:564-570, 1976.
- Costa, P.T. & McCrae, R.R. Psychiatric symptom dimensions in the Cornell Medical Index among normal adult males. Journal of Clinical Psychology, 33: 941-946, 1977.
- Costa, P.T., Fozard, J.L. & McCrae, R.R. Personological interpretations of factors from the Strong Vocational Blank Scales. Journal of Vocational Behavior, 10:231-243, 1977.
- Costa, P.T., McCraw, R.R. & Bosse, R. Smoking motives factor: A review and replication. International Journal of the Addictions. In press.
- Damon, A. Predicting age from body measurements and observations. Aging and Human Development, 3:169-174, 1972.
- Damon, A. & Seltzer, C.C. Anthropometry in the Normative Aging Study of veterans: Physique, age, serum cholesterol, uric acid, and personality. Aging and Human Development, 3:71-76, 1972.
- Damon, A., Seltzer, C.C., Stoudt, H.W. & Bell, B. Age and physique in healthy white veterans at Boston. Gerontologist, 27:202-208, 1972.
- Ekerdt, D.J., Rose, C.L., Bosse, R. & Costa, P.T. Longitudinal changes in preferred age of retirement. Journal of Occupational Psychology 49:161-169, 1976.
- Ekerdt, D.J. & Bosse, R. Rejection of retirement. The Gerontologist, 17(2, Part II):58, 1977.
- Ekerdt, D.J. Second thoughts: Change in preferred age for retirement among older male workers. (Doctoral dissertation, Boston University, 1979). Dissertation Abstracts International, 39, 7545A, 1979. (University Microfilms No. 79-12138)
- Ekerdt, D.J., Bosse, R. & Moge, J.M. Concurrent change in planned and preferred age for retirement. Journal of Gerontology, 35, 232-240, 1980.

- Erikson, R.C., Poon, L.W. & Walsh-Sweeney, L. Clinical memory testing of the elderly. In Poon, L.W., Fozard, J.L., Cermak, L.S., Arenberg, D. & Thompson, L.W. (Eds.), New directions in memory and aging: Proceedings of the George A. Talland Memorial Conference. Hillsdale, New Jersey: Lawrence Erlbaum Assoc., Inc., 1980.
- Floderus, B. Psycho-social factors in relations to coronary heart disease and associated risk factors. Nordisk Hygienisk Tid Skrift Supplementum 6, Stockholm, 1974.
- Fozard, J.L. & Carr, G.D. Age differences and psychological estimates of abilities and skills. Industrial Gerontology, 13:19-43, 1972.
- Fozard, J.L. & Nuttall, R.L. General Aptitude Test Battery scores for men in different age and socio-economic groups. In G. Shatto (Ed.), Employment of the middle-aged: Papers from the Industrial Gerontology Seminars. Springfield, Illinois: C.C. Thomas, Chapter 10, 117-138, 1972.
- Fozard, J.L. Facts and fiction about behavior changes with age implications for the concepts of obsolescence and vitality as applied to the engineer. In: Dublin, S.S. & Skelton, H.S. (Eds.), Maintaining Professional and Technical Competence of the Older Engineer-Engineering and Psychological Aspects, Washington, D.C. American Society for Engineering Education, Publ. 55-71, 1974.
- Fozard, J.L. & Thomas, J.C. Psychology of Aging. In Howells, J.G. (Ed.) Modern Perspectives in the Psychiatry of Old Age, New York: Brunner/Hazel, Inc., 1975.
- Fozard, J.L., Wolf, E., Bell, B., McFarland, R.A. & Podolsky, S. Visual perception and communication. In J.E. Birren & K.W. Schaie (Eds.) The Handbook of the Psychology of Aging, New York, Van Nostrand Reinhold, 1977.
- Fozard, J.L. & Popkin, S.J. Optimizing adult development: ends and means of an applied psychology of aging. American Psychologist 11:975-989, 1978.
- Friedlaender, J., Costa, P.T., Bosse, R., Ellis, E., Rhodes, J. & Stouder, H.W. Longitudinal physique changes among healthy white veterans at Boston. Human Biology, 49:541-558, 1977.
- Garvey, A.J., Bosse, R. & Seltzer, C.C. Smoking, weight change and age: A longitudinal analysis. Archives of Environmental Health, 28:327-329, 1974.
- Garvey, A.J. & Bosse, R. Age, smoking type, and other factors related to cessation of smoking. The Gerontologist, 18(5, Part II):75, 1978.

- Garvey, A.J. & Sparrow, D.W. Relationships of psychosocial factors to hypertension and coronary heart disease in adult males: A prospective analysis. The Gerontologist, 19(5, Part II), 76, 1979.
- Glynn, R., Rosner, B. & Silbert, J.E. Change in cholesterol and the development of ischemic heart disease in middle aged and older men. The Gerontologist, 19(5, Part II), 79, 1979.
- Harrison, E.W. & Wolf, E. Ocular studies with special reference to glaucoma in a group of 107 elderly males. American Journal of Ophthalmology, 57:235-241, 1964.
- Kapur, K.K., Glass, E.R., Loftus, E.R., Alman, J.E. & Feller, R.P. The Veterans Administration longitudinal study of oral health and disease. Aging and Human Development, 3:149-152, 1972.
- Koff, R.S., Garvey, A.J., Burney, S.W. & Bell, B. Absence of an age effect on sulphobromophthalein retention in healthy men. Gastroenterology, 65:300-302, 1973.
- Koff, R.S., Bell, B. & Garvey, A.J. Hearing levels in a normal veteran population. Digestion, 9:36-40, 1973.
- McCrae, R.R., Bartone, P.T. & Costa, P.T. Age, anxiety, and self-reported health. International Journal of Aging and Human Development, 7:49-59, 1976.
- McCrae, R.R., Costa, P.T. & Bosse, R. Anxiety, extraversion and smoking. British Journal of Social and Clinical Psychology, 17:269-273, 1978.
- Oakes, T.W., Friedman, G.D. & Seltzer, C.C. Mail survey response by health status of smokers, nonsmokers, and ex-smokers. American Journal of Epidemiology, 98:50-55, 1973.
- Poon, L.W. & Fozard, J.L. Speed of retrieval from long-term memory in relation to age, familiarity, and datedness of information. Journal of Gerontology, 5:711-717, 1978.
- Poon, L.W., Fozard, J.L. & Treat, N.J. From clinical and research findings to intervention programs. Experimental Aging Research, 4:235-254, 1978.
- Poon, L.W., Fozard, J.L., Paulshock, D.R. & Thomas, J.C. A questionnaire assessment of age differences in retention of recent and remote events. Experimental Aging Research, 5:401-411, 1979.
- Poon, L.W. A systems approach to the assessment and treatment of memory problems. In Ferguson, J. & Taylor, C.B. (Eds.), Comprehensive Handbook of Behavioral Medicine, Vol. I. New York: Spectrum Publications, 1980.

- Poon, L.W. & Fozard, J.L. Ager and word frequency effects in continuous recognition memory. Journal of Gerontology, 35, 77-86, 1980.
- Poon, L.W., Walsh-Sweeney, L. & Fozard, J.L. Memory skill training for the elderly: salient issues on the use of imagery mnemonics. In Poon, L.W., Fozard, J.L., Cermak, L.S., Arenberg, D. & Thomas, L.W. (Eds.), New directions in memory and aging: Proceedings of the George A. Talland Memorial Conference. Hillsdale, New Jersey: Lawrence Erlbaum Assoc., Inc. 1980.
- Rose, C.L. Representativeness of Volunteer Subjects in a Longitudinal Aging Study. Human Development, 8:152-156, 1965.
- Rose, C.L. & Bell, B. Selection of Geographically Stable Subjects in Longitudinal studies of aging. Journal of the American Geriatric Society, 13(2):143-151, 1965.
- Rose, C.L. Social correlates of smoking in a healthy male population. Aging and Human Development, 3:111-124, 1972.
- Rose, C.L. & Mogen, J.H. Aging and preference for later retirement. Aging and Human Development, 3:45-62, 1972.
- Rose, C.L., Bosse, R. & Seltzer, W.T. The relationship of scientific objectives to population selection and attrition in longitudinal studies: The case of the Normative Aging Study. The Gerontologist, 16:508-516, 1976.
- Rowe, J.W., Costa, P.T., Burney, S. & Podolsky, S. Impact of age on anterior pituitary hormones in adult males. The Gerontologist, 17(5, Part II), 112, 1977.
- Seltzer, C.C., Stoudt, H.W. & Bell, B. Reliability of relative body weight as a criterion of obesity. American Journal of Epidemiology, 92:339-350, 1970.
- Seltzer, C.C. Differences between cigar and pipe smokers in healthy white veterans. Archives of Environmental Health, 25:187-191, 1972.
- Seltzer, C.C. Effect of smoking on blood pressure. American Heart Journal, 87:558-564, 1974.
- Seltzer, C.C., Bosse, R. & Garvey, A.J. Mail survey response by smoking status. American Journal of Epidemiology, 100:453-457, 1974.
- Semine, A.A. & Damon, A. Costochondral ossification and aging in five populations. Human Biology, 47:101-116, 1975.
- Sparrow, D., Bosse, R. & Weiss, S.T. Effect of age and smoking cessation on pulmonary function. The Gerontologist, 19(5, Part II), 147.

- Sparrow, D. & Rowe, J.W. The relationship of cigarette smoking to blood pressure: A longitudinal study. Fourth World Conference on Smoking and Health, Stockholm, Sweden, June, 1979.
- Sparrow, D., Bosse, R. & Rowe, J.W. The influence of age, alcohol consumption and body build of gonadal function in men. Journal of Clinical Endocrinology and Metabolism, In press.
- Sparrow, D., Rowe, J.W. & Silbert, J.E. Cross-sectional and longitudinal changes in the erythrocyte sedimentation rate in men. Journal of Gerontology. In press.
- Sparrow, D., Silbert, J.E. & Rowe, J.W. The influence of age on peripheral lymphocyte count in men: A cross-sectional and longitudinal study. Journal of Gerontology, 35, 163-166, 1980.
- Thomas, J.C., Fozard, J.L. & Waugh, N.C. Age related differences in naming latency. American Journal of Psychology, 90(3), 499-509, 1977.
- Treat, N.J., Poon, L.W. & Fozard, J.L. Age, practice and imagery in paired associate learning. The Gerontologist, 18(5, Part II), 134.
- U.S. Bureau of the Census. Census of population: 1970. Vol. 1, Characteristics of the population. Part 23, Massachusetts. Table 24, U.S. Government Printing Office, Washington, D.C., 1973.
- U.S. Department of Health, Education and Welfare. International Classification of Diseases Adapted for Use in The United States (eighth revision). Public Health Service Publication No. 1693. U.S. Government Printing Office, Washington, D.C., 1968.
- Welsh, O.L., Luterman, D. & Bell, B. The effects of aging on responses to filtered speech. Journal of Gerontology, 24(2), 189-192, 1969.
- Wolf, E. & Schraffa, A.M. Relationships between critical flicker frequency and age in flicker perimetry. Archives of Ophthalmology, 1964, 72, 832-843.
- Wolf, E. & Gardiner, J.S. Studies on the scatter of light in the dioptric media of the eye as a basis of visual glare. Archives of Ophthalmology, 1965, 74, 338-345.
- Wolf, E. Studies of the shrinkage of the visual field with age. Highway Research Record, 1967, 157, 1-7.
- Yntema, D.B., Fozard, J.L., Poon, L.W., Zeller, P. & Freeman, B. Age and keeping track of several things at once. In Poon, L.W. & Fozard, J.L. (Eds.), Design Conference on Decision Making and Aging. VA Technical Report 76-01, 1979.

Young, J.C., Rowe, J.W., Pallotta, J.A., Sparrow, D. & Landsberg, L.
Enhanced plasma norepinephrine response to upright posture and oral
glucose administration in elderly human subjects. Metabolism, In
press.

Longitudinal Studies of Impaired Competence in the Community:
A Description of the Riverside Community Research Project (1963-1979)*

Edgar W. Butler

Susan Schuller Friedman

Department of Sociology

University of California, Riverside

*This investigation was supported by Public Health Service Research Grants
No. MH-08667 and MH-5687 from the National Institute of Mental Health;
Grant No. G-74-0095 from the National Institute of Education; and Grant
No. AG-00320 from the National Institute of Aging.

Longitudinal Studies of Impaired Competence in the Community:
A Description of the Riverside Community Research Project (1963-1979)

Edgar W. Butler
Susan Schuller Friedman
Department of Sociology
University of California, Riverside

Introduction

The Community Research Project begun in 1963 has consisted of an ongoing panel study. This research is a longitudinal study of individuals who were identified earlier in comprehensive household and agency surveys (1963-1964) as being intellectually, emotionally, developmentally or physically impaired as opposed to being nonimpaired. The research focus has been on examining the effects of family, individual and environmental characteristics on the health and social wellbeing of impaired persons over time. The effects of being labeled impaired and the influence that intervention strategies have were also a focus of the study. Particular attention has been paid to various outcomes by different types of impairment (i.e., physical impairment versus intellectual impairment) and the consequences of multiple impairments versus single impairments.

From a social system perspective, an individual is not mentally retarded or otherwise impaired unless and until a label is applied. A person may be considered normal by family, neighbors and peers and yet be labeled "retarded" by the schools or by another community agency. Still other agencies may apply different labels or may even consider the individual normal. As the social system now exists, there are "normal" individuals with characteristics like those of persons who have been labeled, referred to here as "the eligibles." Moreover, both the individual and the system change over time so that movement may occur

between categories of "normal," "eligible for labeling" and "labeled," as well as between subcategories of impairments bearing different labels.

A number of issues have been raised in response to the belief that labeling is a necessary prerequisite to intervention and thus positive. This, of course, is the original reason for "special education" classes; some individuals have needs that cannot be met in the regular classroom. By most definitions, special classes includes those for the handicapped and the superbright. Positive aspects of labeling, of course, are suggested by Alcoholics Anonymous who argue that the label must be accepted before the problem can be overcome.

A more prevalent view is that deficit labels are a form of stigma which have dire effects on those so labeled (Becker, 1963; Goffman, 1961; Kitsuse, 1962). Actually, little is known about the effects of particular labels (e.g., retarded, educationally handicapped), changes in labels (e.g., from mentally retarded to educationally handicapped) or decertification (being labeled and then delabeled). One fact remains: the social system is the primary identifying agency of impaired children. During a 1963-64 agency survey of both public and private agencies in Riverside, California, 2/3 of the labeled mentally retarded were identified and labeled by the schools. In that study from school records, it was possible to identify children eligible for labeling in addition to the labeled and normal children.

Surveys conducted in the Riverside community over the past 10 to 15 years clearly indicate that mental retardation, for example, cannot be studied in isolation. Behavior of the sort that leads to the label "mental retardation" may lead, instead, to alternative labels, e.g., "emotional disturbance" or "educational handicap." Not only minority status, but socioeconomic status and other factors such as cultural deprivation influence the labeling process. Further, there appears to be an implicit assumption that impairments occur as separate entities. However, some children have "clusters" of impairments, whereas others have only one kind. The clustering suggests complex relationships between impairments and implies a host of intervention strategies and programs geared toward maximizing favorable outcomes in subsequent school and community careers.

The study of individual impairments and interrelationship of impairments leads to an expanded focus on impairments in the school and community. A critical objective is to determine the effect that impairments and programs devised to serve populations with specific impairments have upon an individual's subsequent career in the school and community. Concerning the mentally retarded, some knowledge is available about school careers, yet there are only a few studies of the adjustment of

retardates outside the school. There are even fewer studies of the preschool history of students labeled by the schools. Followup studies of impaired children other than the mentally retarded are even more rare. Furthermore, dependent measures (e.g., achievement and "adjustment") and independent measures (IQ) have been sufficiently narrow in scope as to make differential assessment (kind of child x kind of intervention) impossible. In addition, some previous longitudinal studies were poorly designed and replete with sampling biases making results uninterpretable.

The in depth longitudinal studies underway here are designed to examine changes that persons may undergo as they proceed from one social system to another and as they are alternately labeled or not labeled. Being a longitudinal study, it is possible to determine what happens to individuals who were labeled impaired several years earlier. As yet, there is relatively little research to tell us how well or how poorly the impaired succeed in their subsequent careers in the school and community. The research that has been done gives us reason to believe, however, that in some instances the impaired manage to cope well with their school and community environments. This project explores some factors determining various levels of adjustment to school and community social systems.

A number of studies have focused on particular impairments, but no previous research project has possessed the capacity to track as many individuals with as many types of impairments as undertaken here. The principal thrust of this research has been on determining the differential needs of individuals with different kinds of impairments or combinations of impairments and examining how these needs are affected by different intervention programs and strategies or lack thereof, and how in the long run these factors working together affect levels of success in school and community careers.

The Community Research Project originated with an interview survey of a geographically and socioeconomically stratified random sample of 10% of the households in the City of Riverside, California in 1963. One adult from each household was interviewed and asked to provide information on the entire family as well as on each member of the household. The household survey file included personal and demographic data, family attitudes toward the individual, physical health, medical and behavioral history, school and family background data (N=8,750 individuals). The purpose of the original study was to provide a community epidemiology of mental retardation. In addition, in 1963-64 a survey of agencies (N=241) providing services for mentally retarded persons was conducted (for a fuller description of this study, see Lei, Butler, Towitz & McAllister, 1974). Demographic and clinical information was secured on all individuals labeled by that agency as mentally retarded. The agencies were classified into four types: public schools, public agencies other than schools, private clinical agencies and private nonclinical agencies. The

agency file contains personal and demographic data, prenatal history and impairment and labeling data from 241 community agencies for 835 persons. In order to determine the effects of labeling and intervention programs, individuals who were classified as "impaired" or "nonimpaired" in the original 1964 agency survey were traced. While it is somewhat more difficult and expensive to follow migrants than those who have remained behind, there was no practical alternative since the impaired may have been more or less migratory than the nonimpaired.

Two determinations of impairments were made in this research. First, all persons identified and labeled with a deficit by a community agency were considered to be "labeled impaired." Second, impairments were determined by the use of questionnaire data gathered in the original 1963 household survey. Four impairments are considered in this research. (1) Intellectual: learning problems, educable and trainable mental retardation, and retardation below the trainable level; (2) Developmental: primarily concerned with behavioral problems; (3) Physical: blindness and partial sight, deafness and hardness of hearing, orthopedic handicaps, speech handicaps and other related physical deficiencies; and (4) Emotional: having extreme depression, frightening, upsetting or other emotional traumas. Within the framework of this research, it is possible for a person to be assessed as impaired on any, all or none of these four impairments.

Our research reveals that there are many individuals with characteristics similar to individuals who have been labeled as retarded, but who themselves are not so labeled. Specifically, it is estimated that for every labeled mental retardate, there are at least seven persons who have not been identified by any community agency and who could be considered eligible for the label "mentally retarded" or some alternative label implying impaired intellectual competence. In our research such persons were characterized as the "eligibles" for purposes of comparisons with the agency labeled impaired and nonimpaired.

An expanded focus on impaired competence in the community was one objective of this study. The second and perhaps even more critical objective was to determine what effect impairments have upon an individual's subsequent life career. The research analyses enlightens us not only about the consequences of being labeled, but even more about the results of setting the criteria for labeling but not for being labeled. Also, it provides crucial information regarding the general lifestyles of such people, as well as information on a variety of specific outcomes which are based on data from a survey of previous research and speculative literature on impaired competence. The obtained analyses and information greatly aid in filling gaps in knowledge about the community life of individuals with impairments.

Since the original study paid particular attention to the process of identification of normal and impaired community residents, followup studies of the sample permit an examination of the effect of changing impairment status related to the aging process. The fact that respondents in the 1963 study were a random sample of households and individuals also allowed us to study changes in residential, family and ethnic environments and how they effect people as they become older. Finally, the study's longitudinal data is useful in examining changes in the characteristics and mental and physical wellbeing of persons as they chronologically move through the life cycle. Substantial sections of the followup survey instruments were similar to the 1963 schedule. Thus, individual and family characteristics such as socioeconomic status (SES), ethnicity, race, religion, marital status, sex, etc., were included as important dimensions. Similarly, data regarding work, living arrangements and family participation patterns was also collected in the followups and comparisons were made between the two time periods to determine changes that had taken place.

The research models shown in Figures 1 and 2 served as general guidelines for this research; these models assume that there are environmental, individual and family characteristics, mediated by exposure to institutional and community based intervention programs that lead to differentially successful adjustment in the community (outcomes). We assume that the characteristics of the impaired (including the type and degree of impairment), their families and the formal and informal community social systems (as measured in the 1963 baseline data) are important dimensions in differential outcomes (as ascertained by 1974, 1975 and 1977 followup interviews). Thus, these preceding factors are considered as independent dimensions that set the stage for various kinds of intervention strategies by schools and public agencies that may or may not have been utilized during the intervening years. The combination of these characteristics and interventions lead to differential outcomes which are the dependent variables of this research. In the analyses, each demographic, medical and developmental, attitudinal and structural item is interrelated with school and other agency intervention programs and linked to different outcomes.

Throughout the project, particular attention has been paid to activity and lifestyles as reflected in work (or school), social interaction and life space of the various age categories of impaired and nonimpaired. Therefore, the thrust of all analyses has been toward determining whether all or any of the four types of impairments considered -- physical, emotional, intellectual and developmental -- make a difference in subsequent adjustment.

CONTINUED

4 OF 10

Baseline Samples: 1963 and 1964

An essential ingredient for a longitudinal research project is an established baseline sample. For this research, a baseline sample from a 1963-64 research project was available. The data gathered in 1963-64 in the comprehensive household and agency surveys provided the baseline data for a series of followup studies conducted in the 1970s. Dr. Edgar W. Butler who was the field director for the 1963-64 surveys was the principal investigator for these followup studies. At the time of the 1963 baseline study, the population of Riverside was about 85,000. The basic sampling unit was the housing unit; a 10.5 percent sample of the housing units in Riverside was selected. Certain types of group living quarters (i.e., convalescent homes) and temporary quarters (i.e., motels) were excluded. Using enumeration district data, the sample was stratified geographically and socioeconomically to achieve proportional representativeness of all city areas. Five hundred blocks were selected on a probability basis. Each of these blocks was canvassed and every structure was listed. Using a sampling ratio calculated for each block, housing units to be sampled were selected at random from the list of units on the block. (For a more detailed discussion of the sample, see Mercer, 1973 and Mercer & Butler, 1967.) The original sample consisted of 3,192 occupied households in Riverside.

One year preceding the fieldwork, a pilot study was conducted in the nearby city of Pomona, California to pretest the questionnaire. The Riverside fieldwork was conducted in the Summer of 1963. It was preceded by media coverage of the project and a letter was sent to each sample household describing the survey as a study of chronic handicaps with an emphasis on mental retardation. Interviewers, all college graduates, received extensive formal training before and during the fieldwork. The average interview lasted one and one-half hours. Spanish speaking interviewers conducted all interviews in Mexican-American homes and when preferred by the respondents, conducted interviews in Spanish. One member of each household was designated as the respondent, providing information about himself/herself and all other family members. Separate questionnaires were filled out for each family member. Respondents were adult members of the household (over age 18 years) and most respondents were female. Where unrelated persons resided in the same household, the unrelated person answered questions about himself/herself. In total, 2,661 households containing 8,750 persons were interviewed representing a completion rate of 90.7 percent.

A very important dimension was added to the research by combining the original household survey data and agency data. The household survey coincidentally included a number of agency labeled retarded persons. Thus, it was possible to determine what effect labeling had upon individuals. Therefore, we compared agency labeled individuals

with eligibles (those with similar characteristics who were as yet unlabeled by official agencies) and nonimpaired persons. Such comparisons permit an evaluation of the impact labeling has upon subsequent school work and life careers.

Followup Samples: 1972-1977

In the 1970s, a series of followup studies were conducted. The research followed up persons in specified age categories who were classified as impaired based on the measures used in the 1963 household survey. Also any persons in the original 1963 survey who were labeled as impaired by any agency contacted in the 1964 agency survey were studied. In addition to reinterviewing impaired persons, a sample of nonimpaired persons was also interviewed.

The followup studies examined the adjustment of impaired and nonimpaired individuals in the community over a period of time of up to 14 years. In the followup studies, the focus was on individual, family and environmental characteristics and other factors as they related to the social and psychological wellbeing of normal and impaired persons over time. The research was concerned with both impaired who were agency labeled and household identified "eligibles." The research provided an empirical basis for evaluating the influence intervention strategies or the lack thereof have on subsequent adjustment in the community by impaired persons. It also studied the long range effect that various labels used by community agencies have upon impaired individuals.

The objectives of the longitudinal research were as follows:

1. To determine what congruence there was among different kinds of impairments and to identify patterns of movement which may take place among various types of impairments over time.
2. To determine the significance of being impaired in terms of an individual's ability to cope within the everyday formal (e.g., the school system, work) and informal social systems (e.g., kin and friendship groups) of the community.
3. To identify the effects (short and long term) of being labeled and to determine the positive and/or negative consequences of differential labeling procedures.
4. To analyze the impact personal and family characteristics and intervention strategies have upon the community adjustment of impaired and normal individuals.

5. To determine what happens to the subsequent community careers of individuals who at some earlier point in time were identified and labeled "impaired" and who may or may not have been exposed to school and/or community agency rehabilitative and other types of improvement programs in the community.

6. To determine the sociocultural and ecological factors associated with the acquisition of various labels and how such labels are acquired and removed.

7. To determine whether exposure to various community agencies and services results in favorable adjustment.

In 1972, a pilot followup study was conducted in which some 247 persons (then aged 18-24 years) still living in the city of Riverside were contacted. This study was undertaken to determine the feasibility of relocating the original sample for a larger, more comprehensive followup study. A relocation rate of 93 percent was achieved. The sample included all of the impaired (N=400) and a random sample of nonimpaired (N=200). About 93 percent of the 600 selected individuals were located using relocation techniques which proved very effective (see McAllister, Goe & Butler, 1973; McAllister, Butler & Goe, 1973). Of the relocated individuals, those still residing in Riverside in 1972 were interviewed (N=247). The interview was used to pretest the survey questionnaire for the proposed larger scale study to determine 1972 handicap status, to evaluate whether any changes in impairment status had occurred between 1963 and 1972 and to examine the extent of community adjustment as measured by a variety of indicators including educational attainment, employment status, marital stability, use of alcoholic beverages and/or drugs, etc.

In 1974, followup research was conducted in which persons aged 9-15 years in 1963 (or in the age category 20-26 at the time of the followup study) were contacted. The 1974 followup survey sample (N=986) consisted of all persons identified as impaired (N=484) and a random sample, matching in number, of nonimpaired (N=484) in the age category of 9-15 years at the time of the 1963 household survey. The systematic and efficient process for relocating individuals developed during the 1972 pretest again proved effective and nearly 90 percent of the potential respondents were relocated, 88 percent of whom were interviewed. Of those interviewed in 1974, 189 had also been interviewed in 1972.

Two age categories were studied in 1975: (1) those 0-5 years in the 1963 (NIMH funded) and (2) those 16-35 years in 1963 (NIE funded). The 0-5 year old age sample included all persons resident in the household in 1963 who were 0-5 years of age at the time of the baseline study (N=1,117). Data on 935 of the total 1,117 sampled subjects were

collected by interviewing mothers or mother surrogates. The same relocation techniques used in 1972 and 1974 were employed to locate all subjects in this age category. Relocation proved even more efficient and successful in locating younger children since their families were less mobile than others. Nearly 93 percent of the sample was located. The subject's mother or mother surrogate was the designated respondent for these interviews. The sample selected for the study of the 16-35 year old category included all impaired persons in this age category (N=526) and a matching number of randomly selected nonimpaired persons (N=526) for a total sample size of 1,052. Both 1975 surveys were successfully accomplished with more than an 80 percent rate of interview completion.

In 1977, a sample of middle aged and aged persons was recontacted. This sample was composed of two subsamples, those 36-65 years in 1963 and those aged 66 years and over in 1963. All persons in the 66 year and over age category, whether impaired (N=341) or nonimpaired (N=325) in 1963 were selected for the followup sample (N=666). The procedure for selecting the sample for those age 36-65 years was to include all persons who were impaired (physically, intellectually or with multiple impairments) in the sample (N=489). In addition, for each of six five year age categories, a matching proportion of nonimpaired persons within each five year age category were randomly selected (by computer) to be included in the sample (N=483). Thus the total subsample figures were age 36-65 years, 972 subjects and age 66 years and over, 666 subjects. The analysis of these data is currently in process. This study is funded by a grant from the National Institute of Aging.

For each of the followup studies, face to face interviews were carried out in Riverside City or within a 50 mile radius, unless the respondent did not want to be interviewed in person. Telephone interviews were obtained for most of the remaining people with a few exceptions who lived in a foreign country. For these interviews, questionnaires were sent and returned by mail. The questionnaires used in the followup surveys were a result of careful evaluation of the original 1963 household survey instrument, the 1972 pretest schedule and questions and scales used in other related studies, as well as suggestions by various staff members. The results provided ample data and a variety of measurements for the longitudinal study.

Measures

While the original research done in Riverside focused primarily on the mentally retarded, the broader, more inclusive problem of impaired competence was expanded in systematic followup research. Competence

refers here to the performance of a social role in a social system, i.e., one that fulfills the system's norms for acceptance. Impaired competence, then, designates behavior that does not reach normative expectations. Four dimensions of impairments were considered in this research: (a) intellectual - learning problems, educable and trainable mental retardation and retardable below the trainable level; (b) developmental - essentially concerned with behavioral problems; (c) physical - blindness and partial sight, deafness and hard of hearing, orthopedic handicaps, speech handicaps and other related impairments; and (d) emotional - having extreme depression, frightening, upsetting or other emotional traumas. The critical objective of this longitudinal research was to determine the effects that impairments, clusters of impairments and intervention programs devised to serve populations with impairments have upon subsequent careers in the community.

A typology of impairments derived empirically from the original household survey data is presented in Table 1. A complete description of the original sample in terms of various types of impairment is shown. Individuals in this table are separated by broad age categories in order to clarify distinctions existing in impairments by age.

Insert Table 1 About Here

From this table it may be seen that nearly 29 percent of all persons in the original sample are impaired. This is a total of over 2,500 individuals in the original sample. Of the impaired, over 26 percent (N=658) have multiple problems. There is a remarkable increase in the percentage of impaired persons from middle through old age. The principal reason for this is the substantial increase in intellectual and physical impairments among older age categories as well as an increase in multiple impairments. Part of the high level of intellectual impairment in the older age categories is attributable to cohort differences in level of education and schooling. The percentage of "nonimpaired" persons declines from almost 75 percent in the younger and middle ages to 65 percent in the 46-55 year old category and declines to approximately 46 percent for those 65 and over. These are minimal percentages since they do not include emotional impairments.

Two determinations of impairments were made in this research. First, all persons identified and given a deficit label by a community agency were considered to be "labeled impaired." Second, impairments were determined by the use of questionnaire data gathered in the original 1963

household survey. A typology of impairments was derived from the 1963 household survey data. A person could be assessed as impaired on any, all or none of these four impairments: intellectual, developmental, physical and emotional.

In general, the measure of intellectual performance was based on responses to items pertaining to academic performance in the school system. The measures of developmental impairment were based on responses to a number of behavioral items derived from the Vineland Social Maturity Scale (Doll, 1965) and the Gesell Developmental Scales (Gesell, 1948a, 1948b, 1956). These scales were revised and adapted for use in a questionnaire format and were augmented by other items developed to measure various dimensions of impairments (for the original formulation and a report from the pretest on the utility of these items in effectively identifying the impaired, see Mercer, Butler & Dingman, 1964). The items were used to form dimensions that conform rather closely to clinical practice and definition; they are age graded, reflecting the increasing number and complexity of roles and are calibrated with the pace of expected social development.

Physical impairment was measured by two sets of items: the first focused on medical history of damage to the central nervous system and the second was based on responses to a set of items measuring current physical functioning. The measure of emotional impairment was based on responses to a series of items about mentally disturbing situations. Individuals failing a sufficient number of any one of the four impairment measures so as to be functioning at the lowest 16 percent level for their age category were identified as having that particular type of impairment.

The household survey included personal and demographic data, physical health condition, medical, educational and behavioral history and family background of each individual household member. Part of the questionnaire administered in 1974, 1975 and 1977 were similar to the original 1963 survey instrument. This resolved some measurement problems connected with the use of different scales to measure "same" phenomena at two different time periods. Thus, individual and family characteristics such as socioeconomic status (SES), ethnicity, race, religion, marital status, sex, etc., were included as important dimensions. Similarly, work information, living arrangements and family participation patterns were determined in the followup studies and comparisons were made between the two time periods to determine changes that took place.

The surveys included measures of formal and informal social participation including contact with family, friends, co-workers, neighbors and voluntary associations. In addition, questions were asked in regard to interaction (the nature, frequency and perceived benefit derived from such contacts) with various community service agencies. Outcome measures

differed depending on the person's age and stage in the life cycle. For those still in school, outcome measures included educational performance and attainment, home adjustment and social adjustment other than school (i.e., confrontation with the law and other agencies). For those out of school, outcome measures included educational attainment, economic situation (i.e., income, home ownership), work patterns (i.e., occupational level, job stability, job satisfaction) and social adjustment and independence (i.e., confrontation with the law, drug and alcohol use, satisfaction with living quarters and with interpersonal relationships).

Contacts with various agencies and other community organizations designed to mediate problems incurred by impaired individuals was ascertained by determining the nature, frequency and perceived benefit derived from such agency contact. Investigation of the role of agencies in determining outcomes is critical in evaluating the effect of various intervention strategies -- an important goal of the research. Hence, dependent variables included measures of community related adjustment which were broader in scope than those employed in previous investigations with a focus on the longitudinal aspects of adjustment in community settings. Interviews with a significant subsample of persons with a variety of impairments were done to assess the extent of impaired competence (and changes in impaired competence status over time), and the degree to which persons are successfully coping in the community. Outcomes, defined by a variety of individual performances in the community, constituted the principal dependent variables of the study. The ways in which various intervention strategies may have affected a person's outcomes also were assessed.

Of special interest in the 1977 study of the aged was longevity and death since during the interim 14 years many of the aged were expected to have died. In our followup, longevity and date and cause of death were determined. Utilizing data obtained from the 1963 interview schedule and information obtained from the "certificate of death," individual, family and environmental characteristics and their relationship to differential death risks were ascertained.

Summary

The Community Research Project has been conducting a series of panel studies since 1963. A critical objective of these studies has been to determine the effect that impairments, clusters of impairments and intervention programs devised to serve populations with impairments have upon subsequent careers in the community. The longitudinal, in depth studies carried out so far in this project were designed to examine

changes that persons may undergo as they go through the life cycle and as they are alternately labeled or not labeled. Since this is a longitudinal research project, it is possible to determine what happens to individuals in the community who 10 or more years earlier were labeled as impaired or were eligible for such labeling.

There is a complex interrelationship among different kinds of impairments between individuals with impairments and agencies who give labels in the community. There is a need to determine what effect these factors have upon the subsequent life careers of individuals. The principal thrust of this research has been to determine differential adjustments of persons with different kinds of impairments or combinations of impairments; seeing how these adjustments are affected by intervention programs and strategies or lack thereof; and how in the long run these factors working together produce life careers with various degrees of success.

Our research suggests definite social etiological factors in intellectual impairments and behavioral retardation, but this is less clear for physical disabilities. Further, to date the research has demonstrated that there are varying ecological distributions of impairments and social characteristics associated with different kinds of impairments and hence the need for services varies by ecological area. The baseline 1963 and subsequent 1972, 1974, 1975 and 1977 followup data allow evaluation of the relative contribution social characteristics and exposure to action programs have upon subsequent community adjustment of impaired individuals. These results have a direct impact upon policy decisions as regards evaluation of the effectiveness of action programs.

To date, much has been learned about individuals in terms of being "normal," "eligible" or "labeled impaired"; and as a result of previous research, a great deal about epidemiology of such cases is known. However, virtually all of this knowledge concerns individuals at one point in time. Proceeding longitudinally and in depth makes possible a determination of changes that persons may undergo as they proceed from one "social system" to another and as they are alternately labeled, delabeled and relabeled or not labeled. Being a panel study, it makes possible a determination of what happens to persons in later life who when younger were considered to be impaired. In addition to these considerations, the research analyses make possible better understanding of the social characteristics and long range outcome of persons whom we have classified as "eligibles." These eligibles constitute an important part of the Riverside sample.

In summary, a number of studies have focused on particular impairment and while some have followed up certain isolated types, no previous research project has possessed the capacity to track as many individuals

with as many types of impairments as has been attempted here. The principal thrust of this research has been on determining the differential needs of individuals with different kinds of handicaps or combinations of handicaps and examining how these needs are affected by different intervention programs and strategies or lack thereof, and how in the long run these factors working together affect levels of success in occupational community life careers.

The research that has been done gives us reason to believe that in some instances, the impaired manage to cope quite well with community environments. These preliminary findings are presented here only for illustration. During the next several years, more sophisticated and inclusive data analyses will be carried out to achieve a better understanding of impaired competence in an urban community.

References

- *Butler, E.W., Lei, T.J. & McAllister, R.J. Childhood impairments and subsequent social adjustment, American Journal of Mental Deficiency, 83: 223-232, 1978.
- *Butler, E.W., Lei, T.J. & McAllister, R.J. Impaired competence in an urban community: An ecological analysis, Urban Affairs Quarterly, 13: 421-442, 1978.
- Doll, E.A. Vineland Social Maturity Scale, Condensed, Rev. ed. Manual of Direction. Minneapolis: American Guidance Service, Inc., 1965.
- Gesell, A.L. The First Five Years of Life. New York: Harper and Bros., 1948a.
- Gesell, A.L. Studies in Child Development. New York: Harper and Bros., 1948b.
- Gesell, A.L. Youth: The Years From Ten to Sixteen. New York: Harper and Bros., 1956.
- *Lei, T.J., Butler, E.W., Rowitz, L. & McAllister, R.J. Agency labeled mentally retarded persons in a metropolitan area: An ecological study, American Journal of Mental Deficiency, 79: 22-31, 1974.
- *McAllister, R.J., Butler, E.W. & Goe, S. Tracking respondents in longitudinal surveys: Some preliminary considerations, Public Opinion Quarterly, 37: 413-416, 1973.
- *McAllister, R.J., Butler, E.W. & Lei, T.J. Patterns of social interaction among families of behaviorally retarded children, Journal of Marriage and the Family, February, 1973: 93-100.
- *McAllister, R.J., Butler, E.W. & Goe, S. Evolution of a strategy for the retrieval of cases in longitudinal survey research, Sociology and Social Research, 58: 37-47, 1973.
- *Mercer, J.R. & Butler, E.W. Disengagement of the aged population and response differentials in survey research, Social Forces, 46: 89-96, 1967.
- *Mercer, J.R., Butler, E.W. & Dingman, H.F. The relationship between social developmental performance and mental ability, American Journal of Mental Deficiency, (a) 69: 195-205, 1964.
- *Mercer, J.R. Labeling the Mentally Retarded. Berkeley and Los Angeles: University of California Press, 1973.

*Project publications: a bibliography of additional project publications is available on request.

Table 1

Distribution of Respondents in (1963) Household Survey,
by Age Category and Impairment

Impairment	Age													
	a		a		a		b		c		c			
	0-1		2-5		6-15		16-45		46-55		56-65		66+	
	%	N	%	N	%	N	%	N	%	N	%	N	%	N
None (normal)	75.0	(270)	78.6	(596)	63.9	(1199)	77.8	(2790)	74.8	(687)	65.1	(382)	45.7	(304)
Physical	9.4	(34)	3.7	(28)	8.4	(157)	1.0	(35)	2.2	(20)	4.1	(24)	12.3	(82)
Emotional	*	*	6.9	(52)	5.6	(105)	*	*	*	*	*	*	*	*
Intellectual	*	*	*	*	3.0	(56)	5.0	(181)	8.4	(77)	15.8	(93)	22.1	(147)
Developmental	15.6	(56)	8.1	(61)	8.7	(163)	10.2	(365)	8.8	(81)	4.4	(26)	3.2	(21)
Two Impairments	0.0	(-)	2.4	(18)	7.6	(142)	5.6	(201)	5.1	(47)	9.0	(53)	14.6	(97)
Three Impairments	0.0	(-)	0.3	(2)	2.5	(46)	0.4	(14)	0.8	(7)	1.5	(9)	2.3	(15)
Four Impairments	0.0	(-)	0.0	(-)	0.4	(7)	-	-	-	-	-	-	-	-
-340-	100.0	(360)	100.0	(757)	100.0	(1875)	100.0	(3586)	100.1**	(919)	99.9**	(587)	100.2**	(666)

*As a result of the design of the 1963 questionnaire, there is no "emotional" component for the 0-1 and 16+ population and no "intellectual" component for the 0-5 population.

**Rounding error.

a Longitudinal survey completed: NIMH 08667.

b Longitudinal survey completed: NIE G-74-0095.

c Longitudinal survey in process of being completed: NIA AG 00320.

Figure 1
Research Model 1

I. Individual, Family And Environmental Characteristics) II. Intervening Variables) III. Differential Health Status

1. Individual and Family Characteristics
 - a. Age
 - b. Sex
 - c. SES
 - d. Ethnicity/Race/Religion
 - e. Marital Status
 - f. Individual Medical History
 - g. Family Medical History
 - h. Impairments: Physical, Intellectual, Developmental and/or Emotional
2. History of Loss of Significant Other
3. Work
 - a. Income/Resources
 - b. Financial Security
 - c. Work History
 - d. Retirement
4. Environmental Characteristics
 - a. Age Peer Residential Density
 - b. Environment, Including Air Pollution
 - c. Residential Mobility Patterns
5. Lifestyle; Formal and Informal Social Participation Patterns
6. Drug, Alcohol, Tobacco Use

1. Adjustment to Bereavement or Loss of Significant Others
2. Source of Medical and Social Support
3. Adjustment to School/Work/Retirement
4. Psychological Adjustment
 - a. Stress
 - b. Depression
 - c. Morale/Life Satisfaction
 - d. Self Esteem
 - e. Anomia
 - f. Retreatism
 - g. Self Identification of Age
 - h. Future Orientation
5. Use of Medical and Paramedical Services
6. Individual Self Perception of Health

1. Well-Healthy
2. Chronic Illness
3. Acute Illness
4. Multiple Illnesses
5. Impairments
 - a. Physical
 - b. Intellectual
 - c. Developmental
 - d. Emotional

-34-

Figure 2

Research Model II

I. Individual, Family And Environmental Characteristics	II. Intervening Variables	III. Differential Outcome
<ol style="list-style-type: none"> 1. Differential Health Status <ol style="list-style-type: none"> a. Well-Healthy b. Chronic Illness c. Acute Illness d. Multiple Illnesses e. Impairments: Physical, Intellectual, Developmental and/or Emotional 2. Individual and Family Characteristics <ol style="list-style-type: none"> a. Age b. Sex c. SES d. Ethnicity/Race/Religion e. Marital Status 3. Environmental Characteristics <ol style="list-style-type: none"> a. Age Peer Residential Density b. Neighborhood and Housing Characteristics 4. Drug/Alcohol/Tobacco Use 	<ol style="list-style-type: none"> 1. Quantity/Quality/And Use Pattern - Community and Institutional Services 2. Quantity/Quality/and Use Pattern - Medical and Paramedical Services 3. Family Health Support System 4. Loss of Significant Others 5. Informal/Formal Support System 6. Change in Environment and/or Living Arrangement 7. Change in Work Pattern and/or in Financial Security 8. Change in Drug/Alcohol/Tobacco Use 9. Change in Personal Habits and Lifestyle 10. Patterns of Illness Behavior 	<ol style="list-style-type: none"> 1. Medical and Health Status <ol style="list-style-type: none"> a. Well-Healthy b. Recovery-Well c. Impaired d. Disabled e. Institutionalized f. Death 2. Functioning Level: Ambulation, Hearing, Seeing, etc. 3. Individual Self Perception of Health 4. Social and Psychological and Adjustment: Social Role Adjustment, Patterns of Social Participation, Mental Wellbeing, etc.

The Nebraska Annual Social Indicators Surveys

David Richard Johnson

Lynn K. White

Department of Sociology

University of Nebraska-Lincoln

Lincoln, Nebraska 68588

The Nebraska Annual Social Indicators Surveys

David Richard Johnson

Lynn K. White

Department of Sociology

University of Nebraska-Lincoln

Lincoln, Nebraska 68588

The Nebraska Annual Social Indicators Survey (NASIS) is an omnibus survey of the quality of life in the State of Nebraska. Each year a representative sample of Nebraskans is interviewed by telephone about the environment, housing, transportation, health, crime, occupation, education, family life and other matters that reflect the quality of life they experience. The data are used by a variety of state and local governments for policy and program planning efforts and by faculty and students at the University of Nebraska and other colleges for research.

The survey is designed to have both a cross-sectional and a panel component: each odd year the initial sample, first interviewed in 1977, will be reinterviewed; each even year a new cross-sectional sample will be drawn. This design maximizes the ability of the survey to monitor change in quality of life both in terms of statewide trends and changes in individuals.

Background

The annual survey is conducted by the Bureau of Sociological Research, the research arm of the Department of Sociology at the University of Nebraska-Lincoln. Since its inception in the middle 1960s, the Bureau has supported and administered the research facilities of the Department. Support for these research capacities came in part from contract research for nonprofit organizations, largely state agencies, but also from

scholars within the university who had funded research projects. While continuation of contract research was essential for the financial integrity of the Bureau, much of the research itself was largely irrelevant to the research interest of the faculty, or it was gathered in a manner which limited its usefulness.

The Nebraska Annual Social Indicators Surveys were designed to make the research of the Bureau more productive in terms of scholarly research and also to correct some of the inefficiency with which data about Nebraskans were being collected by a variety of public agencies. One carefully designed and conducted interview survey a year could meet many of the data collection needs of state agencies, each of which in the past had fielded their own separate surveys. Such a combined survey has several advantages. First, the cost of meeting each agency's information needs will normally be less than the cost of fielding several separate surveys. The costs of designing and drawing the sample, developing the instruments, training interviewers, collecting basic demographic data on the respondents and data processing and basic computer programming are shared. Thus, higher quality data with a larger sample are possible at a lower cost. Second, by combining data collection, the concerns of the various agencies can be related, often for the first time, as information needed by each user is gathered from the same set of respondents. Third, by repeating the survey annually and including a panel component, the data produced are of use to scholars and policy analysts concerned with trends in quality of life and in demographic and social factors. Finally, the economical nature of the combined survey makes it possible to set aside part of the interview schedule for questions related to the sociological research interests of the department's faculty.

Study Design

The Nebraska Annual Social Indicators Surveys are designed to provide data on a random sample of 1800 adults living in households in the State of Nebraska. The sampling and field procedures are based on techniques used extensively by the major national survey research centers. Probability procedures are used to permit accurate estimation of sampling errors and to minimize problems associated with sampling bias.

The survey is designed to provide contemporary data on single years and to provide a picture of changes in quality of life in the state. To this end, the surveys have two components, a trend and a panel segment. The NASIS study design calls for using the 1977 respondents as the basis for the panel. These respondents will be recontacted every two years (1979, 1981 and so on) until they leave the population through death or migration, refuse to participate any longer, or cannot be located after a residential move. The panel will be used both to describe conditions

in the current year and to describe changes over time. Supplementary procedures will be used to keep the panel representative of the population from year to year and to prevent the dating and aging of the panel. Each even year a cross-sectional sample will be independently selected and will serve as a trend sample. Persons interviewed in these trend surveys will be used to supplement the panel study: all individuals in an even year who have just turned 18, have just been released from institutions or have just moved to the state will automatically become permanent members of the panel. In addition, quota procedures will be used to select respondents from the trend studies who can replace persons leaving the panel because of nonresponse.

The following sections review briefly the basic design features including sampling methods and item selection. Readers interested in more detail should refer to Johnson (1977) or White (1978, 1979) for full descriptions.

Sampling

A careful review of the most recent literature on sampling strategies demonstrated that telephone interviewing is much less expensive than face-to-face interviewing with very little sacrifice in data quality (Rogers, 1976). Because adequate comparisons of the reliability, validity and representativeness of telephone and face-to-face interviewing were not available for regions with large rural and small town populations, however, the initial study design called for utilizing both methods. In the first wave of the study (1977), one-third of the sample was selected using standard area probability techniques and interviewed in person; the other two-thirds were selected using random digit dialing techniques and interviewed by telephone (Hauke & Cox, 1974). Detailed comparisons of sample representativeness and item responses in both survey strategies demonstrated few significant differences and equal data quality and representativeness. Thus all further surveys in this series utilize random digit dialing for sample selection and telephone interviewing techniques. This allows considerable cost savings while maintaining high data quality.

The samples for NASIS are drawn from a population consisting of noninstitutionalized persons in households with telephones residing in the State of Nebraska during the survey period. Persons under 18 years of age, persons in custodial institutions and on military reservations and transient visitors to the state were excluded from the sampling universe. Selection of respondents from households with telephones is accomplished by means of a set of selection tables randomizing the person to be interviewed in households with more than one eligible respondent. As only one person per household is selected, when respondents are reporting on characteristics of their household, then the sample is representative of all households with telephones in the state. Since

there are differences in selection probabilities for persons in different sized households, a weighting procedure is used in the computerized data files to compensate for these differences and produce an alternative data set representative of all persons in the population.

The normal random digit dialing procedures used in the telephone sampling will not be discussed in detail here as they have been adequately described elsewhere (Hauck & Cox, 1974). However, one major modification of normal procedures was incorporated to deal with the special problems of sampling small town and rural respondents. Since most of the three-digit telephone exchanges in small town and rural areas use only a small fraction of the possible four-digit numbers, an enormous number of calls would have to be made before an assigned number was reached. Because exchanges that have only a small number of listings assign the active numbers to a restricted range of all possible four-digit numbers, we are able to greatly reduce the number of calls needed to get a valid number without reducing representativeness. For example, a community with only 100 telephones may assign them all in the range from 1000-1250. Ranges are ascertained from telephone directory listings and calls are not made to randomly selected numbers which fall outside of the range for the community. Without this procedure, the cost of random digit dialing in small town and rural areas would be prohibitive.

Item Selection

The data gathered by the survey vary from year to year. There is, however, a core of quality of life items and demographic indicators that appears on each survey. These items are selected in consultation with an Advisory Committee of state agency representatives to meet two criteria: (1) provide information of most use for policy formation at the state and local level; and (2) provide maximum comparability with national social indicator data banks, especially the General Social Survey (NORC, 1978).

In addition to this set of core items are the questions added by participating agencies. Some agencies purchase questions on only one survey and these questions appear for one year only. Other agencies buy the same items year after year in order to assess change. Each year room is set aside on the instrument to include a small set of questions from the Study Director (one of the perquisites of that position), some questions of general sociological interest and some items from members of the Department of Sociology.

Interviewing

All interviewing is done on the telephone by experienced, professional interviewers. (Most calls are made on a University WATS line.) The

interviewing is done from the offices of the Bureau of Sociological Research in Lincoln and takes approximately two months. This procedure makes possible a high degree of quality control as interviewers are centrally located and can be closely supervised.

Re-Interview of the 1977 Respondents

In 1977, 1,867 persons were interviewed. Strenuous efforts were made to reach each of these individuals for inclusion in the 1979 survey. These efforts included the following procedures:

1. One year after the survey, each respondent was sent a brief review of the results of the survey.
2. Four weeks prior to the beginning of the 1979 study a letter was sent to each respondent. This letter notified them that an interviewer would be calling them, described the reasons for the interview and enclosed a collage of newspaper articles reporting on the 1977 survey. These letters were sent first class with a "forwarding address requested" stamp.
3. Respondents who had moved or changed telephone numbers in the interim were traced. These efforts were quite successful and, in spite of the fact that a significant number of our respondents had refused to give us their names or addresses, all but 17 of the original 1,867 were traced.
4. Each respondent's home was called until a conclusive response was obtained: completed interview, refusal, knowledge of respondent's death, illness or prolonged absence. In some cases, well over 300 calls were made to the same number.
5. A letter was sent to all respondents who refused to be re-interviewed, asking that they reconsider. After two weeks they were called again. This procedure resulted in completed interviews in the case of 15 percent of the initial refusals. It was concluded that any further efforts to reach the continued refusals could be construed as harassment, and no further efforts were made.
6. Finally, personal letters were sent to allay individual respondents' particular anxieties, personal interviews were arranged with a few individuals uneasy about talking on the telephone and a few respondents were allowed to complete their schedules by mail.
7. Respondents who had moved out of the state were no longer eligible for inclusion in the survey. Nevertheless, these people were traced and a short schedule was administered which covered their reasons for leaving the state and their migration history.

The results of this procedure are displayed in Table 1. Of the original 1,867 respondents from the 1977 wave, successful re-interviews were obtained in 1,515 or 81 percent of the cases. There was a significantly greater likelihood of re-interview in the case of respondents originally interviewed in person: 85 percent of the personally interviewed respondents successfully completed the re-interview whereas only 79 percent of those interviewed by telephone in 1977 completed the re-interview.

Insert Table 1 About Here

The noninterviews are broken down into refusals, deaths, migration and other reasons (illness, prolonged absence). The two most important reasons for noninterview were refusals and migration. Neither was unexpected. Experience in other surveys suggests that while refusals may be significant in the second wave (sometimes as high as 20%), they should drop to five percent in subsequent interviews. Migration, however, should continue to be a cause of panel loss in future years.

Updating the Panel

In order to keep the panel representative of the population of the state rather than of nonmobile Nebraskans in 1977, additional respondents had to be added to the panel. This included all of those people who had become adults (18) since 1977 as well as those who migrated to the state since 1977. The most efficient way to identify these individuals was through the 1978 NASIS file. While the list of 1978 respondents represented 1978 rather than 1979, they provided the best and most recent list of migrants and new adults. All of the individuals who were identified as 18 years of age in 1978 were added to the panel as well as all migrants since Spring 1977.

Raising the Sample to 1800

The many purposes of NASIS required that each year we not only renew the panel but maintain it with at least 1800 respondents. This was done by first making adjustments for refusals. Since deaths and migration are both "natural" losses reflecting changes in the population to which we want to generalize, these losses were not included in the adjustments. Stratified sampling from the 1978 survey respondents was used to make adjustments for refusals. A three-dimensional matrix of 1978 respondents was produced to which respondents were classified by sex, residence (farm, rural nonfarm and urban), and age (in 10-year intervals). The number of respondents to be drawn from each cell was determined by an analysis of the characteristics of refusals at the end of the first seven weeks of interviewing.

In anticipation of this procedure, the original 1977 respondents had been divided into two random groups on the basis of odd or even I.D. numbers. The odd-numbered respondents were concentrated during the first six weeks of interviewing so that the refusals at this midpoint could be considered a representative sample of total refusals. The actual selection of replacements for the refusals was done by systematic sampling of the 1978 respondents with a random start within each cell. Through this procedure a total of 394 individuals were selected resulting in 303 completed interviews.

The total sample for the 1979 NASIS was 1,882 persons, representing a combination of three separate procedures drawing respondents from the independent 1977 and 1978 cross-sections. These 1,882 individuals represent the new panel which will be eligible for re-interview in 1981. Of course, only for the original 1,515 will panel data be available back to 1977; however, the inclusion of the additional 367 respondents from 1978 will ensure that the panel remains contemporary, representative and sufficiently large to permit detailed analysis.

Survey Response Rate

A combination of the respondents reinterviewed from the 1977 and 1978 surveys yielded a total of 1,882 usable interviews. Of these, six were only partially completed because the respondent was unable or unwilling to continue until the end of the interview schedule. These individuals were retained in the sample as some information on reluctant respondents is more valuable than no information at all.

Table 1 includes a breakdown of the eligible sample by type of response to persons refusing to be interviewed and persons moving out of the state. Table 2 presents a breakdown by sex, age and residence for the two major categories of nonresponse. This analysis shows that the attrition of the 1979 panel results from a disproportionate migration of younger people and urban residents and from a higher refusal rate among the rural and older respondents.

Insert Table 2 About Here

Results of the Survey and Future Prospects

On the criterion of demand, the survey is successful. While a prospectus describing the survey is sent each year to agency heads, the survey has never needed to be "sold" to people. During each of the last

two years, participants who wanted to be included in the survey were turned down because of lack of space. Again, many of the participants in 1979 indicate desire to participate again in 1980 or 1981. There appears to be every reason to expect the survey to continue for another decade.

The surveys are being productive of data for scholarly research. This research is distributed in two general ways. First, each year the faculty puts out approximately 10-12 reports aimed at the general public. These are distributed to influential people in the state and are given wide attention in the media. Secondly, faculty and graduate students use the data for their own research. In the last two years NASIS data have formed the basis for papers read at scholarly meetings and published papers. As the panel data become available, this number should increase. In addition, the data have been used in both undergraduate and graduate courses, reaching well over 2,000 students per years.

Hauck, Matthew & Cox, Michael. Locating a sample by random digit dialing, Public Opinion Quarterly, 1974, 28.

Johnson, David R. Designs, procedures, instruments, and forms for the 1977 NASIS. Report No. 5 to the 1977 Nebraska Annual Social Indicators Survey, Bureau of Sociological Research, Department of Sociology, University of Nebraska, Lincoln, Neb., 68588.

National Opinion Research Center. General social surveys, 1972-78: Cumulative codebook. Chicago: University of Chicago Press, 1978.

Rogers, Theresa F. Interviews by telephone and in person: Quality of responses and field performance, Public Opinion Quarterly, 1976, 40, 51-65.

White, Lynn K. Designs, procedures, instruments, and forms for the 1978 NASIS. Report No. 3 from the 1978 Nebraska Annual Social Indicators survey, Bureau of Sociological Research, Department of Sociology, University of Nebraska, Lincoln, Neb., 68588

White, Lynn K. Designs, procedures, instruments, and forms for the 1979 NASIS. Report No. 1 of the 1979 Nebraska Annual Social Indicators Survey, Bureau of Sociological Research, Department of Sociology, University of Nebraska, Lincoln, Neb., 68588.

Table 1
Response Rate to Second Wave of Panel Study

	First Wave 1977		Additions From 1978 Cross-Section		Total	
	N	Percent	N	Percent	N	Percent
Successful reinterview	1515	81.1	367	75.5	1882	80.0
Refusal	173	9.3	63	13.0	236	10.0
Left Nebraska	107	5.7	28	5.8	135	5.7
Deceased	36	1.9	3	0.6	39	1.7
Untraceable	13	0.7	4	0.8	17	0.7
Other	23	1.2	21	4.3	44	1.9
TOTALS	1867	99.9	486	100.0	2353	100.0

Table 2
Second Wave Non-Response Rates by Sex, Age and Residence

	Percent Refusing	Percent Leaving State
Total	10.0	5.7
Sex:		
Male	9.8	6.6
Female	10.2	5.2
Age:		
18 - 19	7.8	7.8
20 - 24	8.2	14.6
25 - 29	7.1	9.0
30 - 34	4.9	9.2
35 - 39	7.6	8.1
40 - 44	5.8	8.3
45 - 49	12.9	1.8
50 - 54	13.4	1.3
55 - 59	11.0	0.6
60 - 64	7.1	0.6
65 - 69	14.9	1.3
70 - 74	11.5	0.8
75+	18.8	1.8
Residence:		
Farm	12.5	2.7
Rural Non-Farm	10.9	3.2
Town or City	9.1	7.2

The Woodlawn Mental Health
Longitudinal Community Epidemiological Project

Sheppard G. Kellam, M.D.*

Margaret E. Ensminger, Ph.D.**

Jeannette Branch, M.A.***

C. Hendricks Brown, M.A. ****

John P. Fleming, Ph.D.*****

from the
Social Psychiatry Study Center
Department of Psychiatry, University of Chicago

*Professor of Psychiatry and Director of the Social Psychiatry Study Center

**Assistant Professor of Sociology, Illinois Institute of Technology, and Sociologist, Social Psychiatry Study Center, University of Chicago

***Field Work Associate Professor in Psychiatry and Senior Research Analyst, Social Psychiatry Study Center, University of Chicago

****Chief Statistician, Social Psychiatry Study Center, University of Chicago

*****Postdoctoral Clinical Research Training Fellow in Adolescence at the Social Psychiatry Study Center, University of Chicago

Acknowledgements

The authors wish to acknowledge the crucial contributions of the Woodlawn community, its families and children and the community board members who over the last 18 years have provided support and guidance for this research and service enterprise. The faculty and staffs of the Woodlawn public and Catholic elementary schools and those of the Chicago Public High Schools made crucial contributions. Over the years Dr. Curtis Melnick, former Associate Superintendent of the Chicago Board of Education, was very important to this project.

We wish to acknowledge the contribution of our co-authors in the studies presented here and our colleagues who are currently involved in the project: George Bohrnstedt, collaborator and consultant in methodology; Pamela Hecht, sociologist and chief of data management; David Stevenson, postdoctoral fellow; Rebecca Adams, sociologist; and Mark Schmidt, medical student. We would also like to recognize the work of Center staff: Gary Coleman, David Creagan, Mark Jaeger, Cora Jenkins and Elizabeth Keller.

These studies have been supported by the following grants: State of Illinois Department of Mental Health Grant Numbers 17-224 and 17-322; P.H.S. Grant Number MH-15760 and Research Scientist Development Award (Kellam) Grant Number 1-K01-MH-47596; the Maurice Falk Medical Fund; and National Institute on Drug Abuse grants DA-00787 and DA-02591. John Fleming is supported by an NIMH postdoctoral training grant (T32 MH-14668) issued to the University of Chicago and the Institute for Psychosomatic and Psychiatric Research and Training at Michael Reese Hospital and Medical Center.

The Woodlawn Mental Health
Longitudinal Community Epidemiological Project

Sheppard G. Kellam, M.D.

Margaret E. Ensminger, Ph.D.

Jeannette Branch, M.A.

C. Hendricks Brown, M.A.

John P. Fleming, Ph.D.

Social Psychiatry Study Center

Department of Psychiatry, University of Chicago

The Woodlawn Psychiatric Community Epidemiological Project is a prospective longitudinal study of children and their families in a Black, urban, poor community on Chicago's South Side. Starting in 1963, four successive cohorts consisting of all first grade children in the Woodlawn community were assessed. All the children in their classrooms were studied three times in the first grade and again in the third grade. For the first and third cohort, family data were also collected through an extensive interview with the children's mothers or mother surrogates. Finally, children and their families in the third cohort were followed up 10 years after the initial assessment when the children were 16 or 17 years old. These followup data include information on family structure, interaction, psychological atmosphere and other social characteristics, measures of psychological and psychiatric status, measures of drug use, delinquency and social role performance of the teenager.

An important characteristic of the Woodlawn Project is its focus on a well defined population using a prospective life span developmental orientation and methods. Such an approach helps solve the difficulties inherent in retrospective accounts of development and socialization. The Woodlawn data are epidemiological in that they were collected from a total community population rather than from a pathological population. Epidemiological data are of special value because they allow for the comparison of healthy versus ill and adapting versus maladapting individuals within a total population. Early studies were also replicated in conservative cohorts in this community, thereby offering the opportunity for testing and validation of hypotheses.

Social Adaptational Status and Psychological Well Being:
A Two Dimensional View of Mental Health

Before discussing the research programs pursued in the Woodlawn studies, it is important to examine first the conceptual frameworks which have guided the project since its inception. In the Woodlawn studies, we have considered it important to distinguish between two broad classes of outcomes -- social role performance and psychological status. While the two may be empirically related (and in the Woodlawn research we found this to be the case), they are nevertheless conceptually quite different and the relationships we have found are by no means simple. Many more first grade children were thought by their teachers to be maladapting than were symptomatic in the view of observing clinicians. The short and long term courses of social role performance and psychological well being were strikingly different and their interrelationships very specific.

In considering this two dimensional view and in particular the role performance aspect, the life course social field concept may help (Kellam, Branch, Agrawal & Ensminger, 1975; Kellam & Ensminger, 1980). It is based on the theoretical development of Havighurst (1952), Erikson (1959, 1963), Neugarten (1968) and others over the last several decades.

Every individual in society passes through stages of life which at certain times seem more clear and at others more vague (Neugarten, 1979). Each stage involves that individual in specific social fields in which there are persons who define social tasks and judge the adequacy of the individual's performance in that field. Such persons, whom we have termed "natural raters," are similar to Lippitt's "socialization agents" (1968). The parents at home, the teacher in the classroom, the foreman at work are examples of such persons in specific social fields. The process involved is highly interactional and we have named it social adaptation. The natural raters rate the

adequacy of each individual's performance sometimes very formally as teachers do with grades, sometimes informally as in parental attitudes about how well the child is behaving. We conceive of this interactional process as the basic interface between the individual and society.

Social adaptational status (SAS) -- the adequacy of performance as rated by the natural rater in a specific social field -- is a societal judgment of the individual's performance and is therefore external to the individual. It may be thought of as similar to Parsons' (1964) concept of role performance, i.e., the adequacy with which an individual meets the expectations of social roles. The concept of social adaptational status further specifies and elaborates this notion.

In contradistinction to SAS, there is the question of how the individual is feeling inside; that is, his or her psychological well being (PWB). By PWB we mean the thought processes, affective status, self esteem and other aspects of the psychological status of the individual. These two components -- SAS and PWB -- represent the two major dimensions of mental health. One represents mental health from the viewpoint of society; the other represents mental health from the viewpoint of the individual. This two dimensional concept of mental health builds on a variety of concepts from social and behavioral sciences. Charles Horton Cooley (1922/1962) and George Herbert Mead (1934) were both concerned with the individual's perception of how others see him or her and its effect on his/her sense of self.

From an epidemiological standpoint, the social fields represent the locations within the community where the individuals at a particular stage of life can be found. Equally important, within each social field can also be found the natural rater. In defining social tasks and rating the individuals within the social fields, the natural raters are the sources of social adaptational measures for the epidemiological investigator. For child psychiatric epidemiology, the family and the classroom are the two major social fields in which young children can be found, and parents in the home and teachers in the classroom are the natural raters. Very soon the peer group becomes a major locus also.

Quasi SAS measures can be taken from other indications of success or failure in a social field. For example, we have considered achievement test scores, IQ scores and grades in school as indirect measures of SAS. They are classroom tasks the child is expected to perform. Reliability of SAS measures is often examined by the test-retest method, without assessing interrater agreement, because there often is only one teacher in a classroom or only one foreman at the job. In many social fields such natural raters carry sole authority and another person's opinion, no matter how expert, often does not carry

the authority of the natural rater. Chance and idiosyncrasy, therefore, may play a large role in SAS. The teacher's expectations and tolerance, the fit of the students with each other and with the teacher, as well as the skills of the child may also play a role in the child's SAS. What makes this evaluation even more important in childhood is that the child has much less choice in teacher or classroom or parents compared to the choice of natural raters and social fields later in adolescence and adulthood.

Measures of SAS can be obtained by first systematically asking the natural raters in a specific social field what social tasks are expected. Then scales can be made and the natural raters asked in standardized fashion to rate the children in their social field regarding each child's performance on each task. This procedure was followed in Woodlawn with the teachers in the classrooms and the mothers at home.

In contrast to SAS, PWB is an internal state and measures can only be inferred from self reports of individuals, particularly as the child grows older and becomes more competent to report; from reports of significant others such as parents, teachers or peers; and/or observations of clinicians, who use reports of others as well as direct observation of the children.

Mental health clinicians and research investigators have most often viewed these two perspectives, SAS and PWB, as opposite sides of the same coin. Psychological status is usually considered primary, with social maladaptation the measure of illness and its severity. Educators, including classroom teachers of young children, have tended to do the opposite, viewing the social adaptation of the child as primary and the psychological status or presence of psychological symptoms as a measure of severity. They too assume that these two perspectives are opposite sides of the same coin, but a different coin than that seen by the mental health professional.

By separating these two dimensions, we are able to investigate their interrelation over time and to discover the long term outcomes of each. Such investigation has been at the center of attention in the Woodlawn studies and the results suggest that SAS in first grade in particular has very important long term predictive and possibly developmental significance at least as far as adolescence (Kellam et al., 1975; Kellam, Adams, Brown & Ensminger, in press-a; Kellam, Branch, Brown & Russell, in press-b). These results tend to be similar to those found by Robins (1966), for example, in her studies of the long term importance of early aggressiveness to adult clinic and criminal outcome. Watt (1978) also has recently reported the long term importance of aggressiveness in young males as an antecedent of schizophrenia in adulthood. Loney (1980) has shown evidence that early aggressiveness

is the key prognostic component of the hyperactivity syndrome. While there are modest correlations in longitudinal studies between early PWB and later PWB or SAS, the evidence from our research and other studies suggests that the psychological dimension does not have the same importance for children to long term outcomes as does SAS. The circumstances which influence these long term relationships of SAS to PWB and the relationships of each to other outcomes requires much more prospective longitudinal research.

Community Epidemiology

The Woodlawn studies centered around the concept of community and the role of community epidemiological studies in our understanding of the incidence and prevalence as well as the etiology of psychiatric and social dysfunction. Community epidemiological studies act as an important complement to demographic studies by holding constant the general characteristics of the community, while studying variation in the independent, mediating and dependent variables (Kellam & Ensminger, 1980).

The concept of community focuses our attention on the familial, educational, political, religious and economic institutions which, along with the physical environment, act to enhance or inhibit the psychological or social development of the individual. The effects of variations in the families, in the schools and their classrooms and in peer groups within a particular kind of community can be examined while holding constant the more general characteristics of the community. Such a strategy requires replication in similar and different kinds of communities to determine relationships which are community specific and those which transcend community.

Especially important to those interested in children and adolescents is the concept of the "community of limited liability" (Janowitz, 1967; Suttles, 1972). This concept recognizes that the local community may be of varying importance for those in different stages of the life cycle or different social situations. For example, the community may have more importance for the very old and the young. The young are limited in their geographic mobility because their main activities of play and school usually take place within the local neighborhood. Services designed for the young, such as educational, recreational and medical facilities, are usually provided in local settings.

The Woodlawn Study and Its Population

Woodlawn is an urban, poor community on the South Side of Chicago. Between 1955 and 1966 it changed from 40 percent to almost totally Black and became substantially overcrowded. Median income and employment rates

also dropped. It ranked among the four most impoverished Chicago neighborhoods by 1966 (DeVise, 1967). However, there was an is heterogeneity within Woodlawn, some sections having higher median income and more home ownership than others.

Between 1964 and 1969 we made assessments of the mental health of all the first graders in Woodlawn at several points in each school year. Further assessments were made on samples of these children in their third grade years. We also conducted interviews in the Spring 1965 and Spring 1967 with the mothers (or mother surrogates) of the children who were in first grade in those two years. These assessments were coupled with service and evaluation programs which were partly directed at systematically developing community wide prevention and early intervention programs (Kellam et al., 1975). Developing sanction from the community is an important step in epidemiological studies. In all of our service and research ventures we were supported by a community board composed of leaders from the community's larger citizen organizations. The support and assistance of the board has been essential throughout the course of this work. We have discussed the initial negotiations with the community and the formation of the community board elsewhere (Kellam & Branch, 1971; Kellam, Branch, Agrawal & Grabill, 1972; Kellam et al., 1975).

The repeated collection of data in multiple different cohorts is important in that it allows us to strengthen our conclusions by separating cohort specific results from those relationships which hold across cohorts. In the first stage of a 10 year followup, we located and reinterviewed 939 (75%) of the mothers or mother surrogates of the 1241 families from the 1966-1967 study. The mothers' refusal rate was 5.9 percent. We could not locate 11.7 percent of the families; 6.5 percent of the families had moved out of the Chicago area; and 4 or 0.3 percent of the children from the study population had died between time 1 and time 2. Following the reinterviews with the 939 mothers and with their permission, 705 teenagers (75%) participated in the reassessments. 14.5 percent refused to participate; 6.7 percent could not be found; 1.5 percent had moved out of Chicago; and 2.2 percent had run away from home or were in an institution.

We have studied the differences among the three followup populations: (1) those children and families seen in first grade but not followed up; (2) those mothers and children seen at first grade and the mothers interviewed at followup; and (3) those who were seen in first grade and both mother and teenager were reassessed at followup. We found few significant differences, based on the first grade family, social adaptational and psychological well being data (Kellam, Ensminger & Simon, 1980; Kellam et al., in press-a). The mothers whom we could not reinterview were somewhat younger, more mobile before and during the child's first grade year and their children were somewhat more

likely to have been in parochial schools in first grade. We found no differences between the two groups of mothers in the 1966-1967 psychological well being, early family income, welfare status or the adults living in the households. It was somewhat harder to trace parochial school families because these schools lacked the centralized, computerized records that the Chicago public school system maintains.

We found significant differences in only three of the 21 first grade measures of the children's social adaptation. Those measures which we examined included IQ and school readiness test scores, along with a set of teachers' ratings of social adaptational status. Only maturity as assessed by teachers early in first grade and the grade point averages for both semesters of first grade were different. All three differences were small and the other 18 SAS measures did not provide any indications of important trends which the three significant differences might be signalling. For further studies and description of how we followed up the population, see Agrawal, Kellam, Klein and Turner (1978).

The Woodlawn Data

We chose the classroom and the home for the measurement of SAS during the first grade because of importance of these social fields to the first grade child. The teacher and mother served as the natural raters in the respective social fields. In both social fields, adaptation was assessed in three major areas: (1) social contact or shyness; (2) authority acceptance or aggression; (3) maturation, cognitive achievement and concentration, collectively referred to as learning problems. Ratings by teachers were made three times in first grade and at the end of third grade. The mothers were interviewed in two of the cohorts at the end of first grade.

Our measures of first grade PWB were based on clinical observations (twice in first grade) and a symptom checklist of 38 items completed by the mother of the child. The items in this checklist were grouped into 13 symptom categories and during analyses were often collapsed into single global construct (Kellam, Ensminger & Turner, 1977).

In the 10 year followup the now 16 or 17 year old teenagers provided us with self reports of psychiatric symptoms, self esteem and the teenagers' perceptions of their own social adaptational status (SAS) in school, family, peers and with the opposite sex. We reinterviewed the mothers along the same dimensions. We tried to use as many as possible of the questions we had used with the mothers when the children were first graders. In addition, information regarding drug, alcohol and cigarette use and delinquency was obtained from both mothers and teenagers. We also collected school records, including achievement and attendance.

Extensive data on the families of these Woodlawn children are also available. Measurements of family structure, including household and family composition and marital history are available both during the study child's first grade year and when the child was 16 or 17. Information collected on the family included such things as patterns of interaction within the family, rule setting and punishment, employment and income, social involvement of the family, stressful life events, psychological well being of the parents, parents' satisfaction with different areas of their life and parents' values.

Research Areas of the Woodlawn Studies

Family Structure

Our previous research into the Woodlawn families has indicated the importance of family structure for the social adaptation of children, though not family structure as it is usually conceived of in terms of the presence or absence of the father. Detailed taxonomic descriptions of family compositions indicated a wide diversity in family groupings in Woodlawn.

Teenage Motherhood

As an extension of our studies on family structure, we have recently completed two studies investigating teenage motherhood and its relation to family composition and to psychological distress in mothers.

Welfare

One social structural indicator which has received research attention in the Woodlawn studies is welfare reciprocity (Ensminger, 1978, 1979).

Drug Use

In a series of three different papers (Kellam et al., 1980, in press-a, in press-b), we have examined patterns of drug use among Woodlawn teenagers and have attempted to identify antecedents of drug use among the measures of SAS and PWB collected in first grade.

Antecedents of Teenage Psychiatric Symptoms

Past and current investigations of teenage psychiatric symptoms (Kellam, Brown & Fleming, in press-c; Kellam, Simon & Ensminger, in press-d) have employed a general measure of psychiatric symptoms and five factor analytically derived scales labeled "anxiety," "depression,"

"bizarreness," "paranoia" and "obsessive-compulsive." This work has been focused primarily on determining the relationship between early SAS, PWB, evolving family characteristics and teenage symptoms.

Delinquency

We are currently undertaking a longitudinal research program aimed at the delineation of social adaptational, psychological and family characteristics which differentiate those adolescents who become delinquent from those who do not. Social control theorists have emphasized the importance of existing social bonds in discriminating between delinquents and nondelinquents (Hirschi, 1969). The Woodlawn data enable us to trace the development of social bonds and measure the influences on delinquency of family structure and atmosphere, role performance evaluations in the home and school and PWB including self esteem.

Early Intervention

During the course of the first grade school year we provided an intervention program involving half of the 12 neighborhood schools. Six of the schools served as treatment schools and the other six were used as matched controls. The intervention is described in detail by Kellam et al. (1975) and consisted of weekly classroom visits by a mental health professional in which the students were asked to discuss their feelings and whatever problems they might be experiencing in becoming students. This was coupled with staff training and crisis response programs.

Who Comes for Treatment

As part of the 10 year followup of the Woodlawn population, a multilevel treatment program was designed which included such services as psychotherapy, remedial reading and mathematics classes, vocational counseling, group discussion sessions and a variety of other activities directed at a wide range of adolescent psychosocial needs. The characteristics of those adolescents who chose to enter a treatment program have not received such investigation.

Conclusions

Our primary goal in the Woodlawn studies has been to examine the interrelations between social adaptation and psychological well being in the context of specific social fields over a significant span of the child's and family's life courses. The broad nature of the Woodlawn data has allowed us to pursue this goal by simultaneously assessing sets of independent and dependent variables. Analyses are now being done which are increasingly moving from prediction to explanation.

Much more analytic work remains to be done on the Woodlawn data and replication in similar and different populations is critical in refining hypotheses and testing them.

Kohn (1976) argued for multivariate approaches to social, behavioral and health sciences, suggesting that the absence of research on sets of outcomes has inhibited the development of theories which explain specific outcomes and understanding of a set of outcomes. Our research has been generally consistent with Kohn's suggestion. We hope that this multidimensional approach will provide a more complete understanding of the antecedents and evolving courses leading to specific outcomes as well as providing a clearer understanding of teenage outcomes generally.

References

- Agrawal, K.C., Kellam, S.G., Klein, Z.E. & Turner, J. The Woodlawn mental health studies: Tracking children and families for long-term follow-up. American Journal of Public Health, 1978, 68, 139-142.
- Brown, C.H., Adams, R.G. & Kellam, S.G. A longitudinal study of teenage motherhood and symptoms of distress. To be published in R.G. Simmons (Ed.), Research in community mental health (Vol. II). Greenwich, Conn.: JAI Press, 1980.
- Cooley, C.H. Human nature and the social order. New York: Schocken Books, 1964.
- DeVise, P. Chicago's widening color gap. Interuniversity Social Research Committee, Report No. 2, Chicago Regional Hospital Study, 1967.
- Ensminger, M.E. Welfare: Social and psychological differences between recipients and non-recipients. Unpublished doctoral dissertation, University of Chicago, 1978.
- Ensminger, M.D. Welfare status and feelings of psychological distress. Paper presented at a meeting of the American Sociological Association, Boston, 1979.
- Erikson, E.H. Identity and the life cycle, selected papers. In A.S. Klein (Ed.), Psychological issues. New York: International Universities Press, 1959.
- Erikson, E.H. Childhood and society (Rev. ed.). New York: Norton, 1963.
- Gold, M. Delinquent behavior in an American city. Belmont, Calif.: Brooks Cole, 1970.
- Havighurst, R. Developmental tasks and education (2nd ed.). New York: Longmans Green, 1952.
- Hirschi, T. Causes of delinquency. Berkeley: University of California Press, 1969.
- Janowitz, M. The community press in an urban setting: The social elements of urbanism (2nd ed.). Chicago: University of Chicago Press, 1967.
- Johnston, L. & O'Malley, P.M. A cross-cohort comparison of the drugs-delinquency connection. A paper presented at the symposium on cohort studies held in Uppsala, Sweden on August 11-12, 1978 under sponsorship of the Research Committee for Social Deviance and Social Control of the International Sociological Association, the American Society of Criminology, and the Scandinavian Research Council for Criminology.
- Kellam, S.G., Adams, R.G., Brown, C.H. & Ensminger, M.E. The long-term evolution of the family structure of teenage and older mothers. Accepted for publication in Journal of Marriage and the Family, in press-a.
- Kellam, S.G. & Branch, J. An approach to community mental health: Analysis of basic problems. Seminars in Psychiatry, 1971, 3, 207-225.
- Kellam, S.G., Branch, J., Agrawal, K.C. & Ensminger, M.E. Mental health and going to school: The Woodlawn program of assessment, early intervention and evaluation. Chicago: University of Chicago Press, 1975.
- Kellam, S.G., Branch, J., Agrawal, K.C. & Grabill, M.E. Strategies in urban mental health. In S.E. Golan & C. Eisdorfer (Eds.), Handbook of community mental health, pp. 711-727. New York: Appleton-Century-Crofts, 1972.
- Kellam, S.G., Branch, J.D., Brown, C.H. & Russell, G. Why teenagers come for treatment: A ten-year prospective study of Woodlawn. Accepted for publication in Journal of the American Academy of Child Psychiatry, in press-b.
- Kellam, S.G., Brown, C.H. & Fleming, J.P. Longitudinal, community epidemiological studies of drug use: Early aggressiveness, shyness, and learning problems. To be published in L.N. Robins (Ed.), Studying drug use and abuse. New York: Neale Watson Academic Publishers, Inc., in press-c.
- Kellam, S.G. & Ensminger, M.E. Theory and method in child psychiatric epidemiology. In F. Earls (Ed.), Studying children epidemiologically. New York: Neale Watson Academic Publishers, 1980.
- Kellam, S.G., Ensminger, M.E. & Brown, C.H. Epidemiological research into the antecedents in early childhood of psychiatric symptoms and drug use in mid-adolescence. In Newsletter, Society for Research in Child Development, Inc., Winter, 1979, 5-6.
- Kellam, S.G., Ensminger, M.E. & Simon, M.B. Mental health in first grade and teenage drug, alcohol, and cigarette use. Drug and Alcohol Dependence, 1980, 5, 273-304.

- Kellam, S.G., Ensminger, M.E. & Turner, R.J. Family structure and the mental health of children: Concurrent and longitudinal community-wide studies. Archives of General Psychiatry, 1977, 34, 1012-1022.
- Kellam, S.G., Simon, M.S. & Ensminger, M.E. Antecedents in first grade of teenage drug use and psychological well-being: A ten-year community-wide prospective study. To be published in D. Ricks & D. Dohrenwend (Eds.), Origins of psychopathology: Research and public policy, in press-d. Also in Prevention of drug abuse, Hearings before the Select Committee on Narcotics Abuse and Control, House of Representatives, Ninety-fifth Congress, Second Session, April 18, 20, 25, May 16, 25, 1978. Washington, D.C.: U.S. Government Printing Office, SCNAC-95-2-4.
- Kohn, M.L. Looking back -- A 25-year review and appraisal of social problems research. Social Problems, 1976, 24, 94-112.
- Lippitt, R. Improving the socialization process. In J.A. Clausen (Ed.), Socialization and society. Boston: Little, Brown, 1968.
- Loney, J. Hyperkinesis comes of age. What do we know and where should we go? American Journal of Orthopsychiatry, 1980, 50, 23-42.
- Mead, G.H. Mind, self, and society. Chicago: University of Chicago Press, 1934.
- Neugarten, B.L. The awareness of middle age. In B.L. Neugarten (Ed.), Middle age and aging: A reader in social psychology. Chicago: University of Chicago Press, 1968.
- Neugarten, B.L. Time, age, and the life cycle. American Journal of Psychiatry, 1979, 136, 887-894.
- Parsons, T. Social structure and personality. New York: Free Press, 1964.
- Robins, L.N. Deviant children grown up: A sociological and psychiatric study of sociopathic personality. Baltimore: Williams and Wilkins, 1966.
- Suttles, G.D. The social construction of communities. Chicago: University of Chicago Press, 1972.
- Watt, N.F. Patterns of childhood social development in adult schizophrenics. Archives of General Psychiatry, 1978, 35, 160-170.

The Midtown Manhattan Longitudinal Study
vs. "The Mental Paradise Lost" Doctrine:
A Controversy Joined*

Leo Srole, Ph.D.
 Anita Kassen Fischer, M.A.
 Faculty of Medicine, Center for Geriatrics and Gerontology
 Columbia University, New York

*Reprinted from Archives of General Psychiatry, 1980, 37, 209-221.
 Copyright 1980. American Medical Association.

The Midtown Manhattan Longitudinal Study

vs. "The Mental Paradise Lost" Doctrine:

A Controversy Joined

Leo Srole, Ph.D

Anita Kassen Fischer, M.A.

Faculty of Medicine, Center for Geriatrics and Gerontology

Columbia University, New York

The Midtown Manhattan Mental Health Study, launched in 1952 and fielded in 1954, has evolved from a single point-of-time investigation of a general population, with a cross-sectional design, into a two-stage follow-up exploration, with a longitudinal, panel design. For purposes of labeling convenience, we here refer to the cross-sectional survey of 1954 as Midtown I, the follow-up field research of 1974 as Midtown II, and the analytic delineation of temporal changes observable between those two stages as the longitudinal Midtown III operation.

With Midtown III, we can now approach the issue of whether urbanization, and the big city in particular is the prime villain behind the presumed trends of deteriorating mental health in recent centuries; Manhattan, having long been tagged "crack-up city" (Lawrence, 1953), might offer the ultimate test case for that view. Midtown III is the first socio-epidemiological analysis of a large, community-based population to uncover a previously undetected change in adult well-being between two successive generations born since the turn of the present century.

Concepts and Methods

First, by way of fixing locus, Midtown is our name for a well-delineated, high-density "gold coast and slum" residential area, one that in 1954 sheltered some 175,000 inhabitants, 99% of them white. In demographic terms, these Midtowners were a close approximation of the 1.25 million, otherwise extremely

heterogeneous, non-Puerto Rican whites then occupying the island of Manhattan. The former, therefore, could be generalized to the latter as parent universe, and probably to similar populations in counterpart hub areas of America's largest cities. Under scrutiny, in short, was a cross-section slice of the most metropolitan segment of the nation's great communal diversity.

Our 1954 field operation was an enterprise in sociopsychiatric epidemiology, with the descriptive and etiological targets fixed on Midtown and its mother borough. From the long perspectives of anthropology and sociology, we viewed that community generically; that is, as presenting visible precipitates of natural experiments of human nature, but concealing the sociocultural and experiential sources of health differences embedded deeply in its cross-cutting subgroups. We accordingly sought to pinpoint descriptively and dissect analytically those sources.

Beyond such basic research goals, Midtown I's social policy sights were set on providing an epidemiological surveyor's map of the community for (A) guidance in planning the expansion and redeployment of professional helping services that were prompted by World War II experiences with military-age men, and for (B) moving preventive psychiatry forward from a dead-center position then near ground-zero. (The degree to which these goals were subsequently approached is discussed by Srole et al., 1978.)

The Midtown I researchers primarily focused on a randomly culled, area-type probability sample of 1,660 adults, who were spread across the entire prime-of-life span between the ages of 20 and 59 years. (These particular boundary lines of age coverage were chosen to facilitate demographic calibration with the like-age groups in the Midtown and white Manhattan populations, involving unpublished data that were especially tabulated for us by the United States Census Bureau (Srole et al., 1962, 1978).)

Funded by the recently activated National Institute of Mental Health, the Midtown I designers were faced with adapting a new and relatively untried post-World War II sample survey technology. Moreover, in confronting a large sample of people from a segment of perhaps the most complex metropolis on earth, and totally preoccupied with the "here and now," the Midtown I designers did not foresee or prepare for a possible restudy of our intended target subjects at some distant date. Several years later, however, while wrestling with refractory analytical problems of a cross-sectional study (Midtown I), one of us (L.S.) began to monitor the highly productive, prevention-relevant follow-up investigation of cardiovascular disease in a sample of Framingham, Massachusetts men (Gordon & Kannel, 1970).

As a result, he was drawn into a commitment ultimately to convert the original Midtown data set, although not specifically framed for the purpose, into the baseline of a two-stage restudy of a large subsample of the same people. More than a decade was to pass before this could begin to come about, in part because we had not systematically built into Midtown I the connecting informational tie-lines that, after a substantial time interval, could help us pick up the trails to the current addresses of an "on-the-move" population element. Given skepticism that our Midtown I sample could, years later, be tracked and corralled in adequate numbers, it was first indispensable to establish follow-up feasibility. With the full support of the National Institute of Mental Health, in 1970 we started a systematic "bureau of missing persons" search to locate the nomads among our Midtown I respondents.

Of course, the fundamental principle of probability sampling for a new

of Study interest would probably not have deviated markedly from those yielded by the 695 panelists.

This abbreviated account of how, and with what results, our Midtown II reinterviewees were painstakingly reassembled for potentially large scientific and history-illuminating purposes has been offered to meet challenges to that panel on the ground that it cannot be representative of the entire Midtown I sample of 1954. Our answer is that the panel is not unrepresentative of its appropriate population universe, that is, the Midtown I sample members surviving to 1974, now advanced to the range of 40 to 79 years of age.

General Mental Health Classification Process

Before proceeding to our panel data, a final level of procedural scaffolding must be sketched to anticipate relevant questions as to the substantive comparability of the key Midtown I and II variable, namely general mental health. The 1954 interview systematically reviewed the presence or absence of an expert-picked sample of some 120 manifestations of mental disturbance. These principally drawn from such established symptom collocations as the Minnesota Multiphasic Personality Inventory, the Cornell Medical Index (CMI), and the World War II military screening Neuropsychiatric Adjunct (NSA). The symptom questions of the CMI and NSA (Stouffer, Lumsdaine, Lumsdaine, et al., 1949), in particular, have been established as reliable and, moreover, as valid by the measure of their power to discriminate between diagnostically heterogeneous psychiatric patient (criterion) and nonpatient (comparison) groups.

Thomas Rennie, the Midtown I study director and senior psychiatrist (until his untimely death in 1956), chose 83 of those items for their clinical relevance and their reference to the period of the interview, and the remainder, which referred to the preadult period, only to lend temporal perspective to the currently visible manifestations of mental health status. Detailed presentations of the inclusive Midtown I symptom questions and the procedures for structuring them are given in Mental Health in the Metropolis (Srole, et al., 1962, 1978).

From Rennie's classification schema of 1954, we can now move to their replication in 1974, in which both of us were centrally involved. For purposes of assessing general mental health in 1974, we repeated with identical wording the 83 "current" symptom items used in 1954, excluding as redundant those that referred retrospectively to the respondent's preadult period.

To accurately discern mental health changes since the 1954 baseline, we might have considered having the Midtown panel's 1974 general mental health status classified by the same psychiatrists and judgmental processes as 20 years earlier. Unfortunately, the Study's psychiatrists of 1954 were no longer available. And, even if they had been present to repeat their classification function of 1954, the likely evolution in their professional perspectives would probably have undermined confidence in the objective reproducibility of their subjective judgmental processes of 29 years earlier.

This dilemma presented us with Midtown III's most formidable procedural impasse, which was fortunately resolved by the availability of a surrogate "classifier" not the scene 20 years ago: The computer. As first step,

investigation is that every person in its population universe must be given an equal chance of being drawn, a principle rigorously applied in Midtown I with, to us, uncanny results. For Midtown II, Richard Remington, our epidemiology consultant (now dean of the University of Michigan School of Public Health), suggested the following adaptation of that principle for follow-up studies: Manage your search efforts so as to give all missing members of the stage I sample the same chance of being located for possible restudy in stage II.

Accordingly, we applied the same sequence of search modalities to every one of our Midtown I "address unknown" respondents (or surviving kin) without exception. At completion of this all-inclusive detective sweep (Fischer & Biel, in press), we had verified the current whereabouts of 1,124 members (67.7%) of our baseline sample, 858 of them alive and 266 certified as deceased. (Mortality, its causes, and its baseline predictors were target variables of Midtown II and III (Singer, Garfinkel, Cohen, & Srole, 1976).)

Of course, the hazard hanging over all longitudinal sample studies is the certainty of sizeable attrition, i.e., loss to follow-up, with the magnitude of losses, other things being more or less the same, tending to be a function of (A) the extent of the tracking information secured in the baseline interviews, (B) the number of years intervening, and (C) the degree of residential turnover in the sample's community base. In all three respects, the attritional risks for the Midtown I sample had been maximal. It might be mentioned that our total unlocated loss of 32.3% figures to an average annual loss of 1.6%, which compares favorably with the rate for most other longitudinal studies focused on career-launched adults. An example is the recent follow-up investigation of 3,224 military personnel exposed to an atmospheric nuclear test in 1957. Despite the uncommon advantages of baseline military records, by 1979 the unlocated loss was 40%, or an average annual loss of 1.8% per year (U.S. Center of Disease Control, 1979).

Located Subsample

We ended our search with 858 baseline respondents located alive, plus 266 certified as dead and a residual 536 as terminally unlocatable. By actuarial estimate, approximately 100 of the latter were deceased. Drawing on our baseline data bank, we could also estimate another 100 to have been never-married women, mainly aged 20 to 39 years, who predictably soon married and abandoned their former identifying surname and home precincts, for search purposes becoming anonymous "needles" in our megapolitan "haystack".

The 858 located survivors, then, comprised the Midtown I subsample finally accessible for approach to be reinterviewed after a lapse of two decades. At minimum, we can in principle generalize from them as their own 20-year controls. Beyond that, we emphatically did not expect to generalize from them to the entire Midtown sample of 1954. However, we can, potentially at least, generalize from them to their living population universe of 1,294 survivors, located and unlocated, in 1974 at age 40 to 79 years, and probably to other similarly constituted populations of like history as well. Practically, however, we could so generalize only if the 858 survivors of known addresses are an approximately representative subsample, in demographic and other respects recorded in our Midtown I information bank, of all 1,294 Midtown I survivors.

We can determine representativeness, to a large degree, by assessing the fit between the characteristics in 1954 of (A) the 858 visible survivors and (B) those of the 1,394 comprising the 1,294 survivors, unlocated and located, plus the 100 estimated but unidentified deceased (probably elderly, lower-status men in the main), who are numerically too few to significantly diminish the appropriateness of the 1,394 aggregate as a reference population universe. A comparison here of the located and the unlocated, i.e., of subsample A and subsample B, would be relevant if we were interested in the determinants of locatability. However, for purposes of generalization, the prerequisite is an answer to the question of the part-whole relationship; more specifically, how representative is subsample A of the inclusive A/B population aggregate that is of primary interest here? A secondary, technical issue is: How successful were our efforts to rigorously implement the Remington principle of equal-chance search for all the missing members of the original sample?

Results of comparing the 858 and 1,394 on the hundreds of characteristics reported in the baseline interviews cannot be documented here, but can be briefly summarized as follows: With relatively few exceptions, the accessible subsample entity did not deviate from its stated parent universe by much more than a small known margin of chance variability. Thus, the subsample of accessible survivors passed the initial test necessary to qualify as representative of that parent universe (and probably of counterpart universes of like provenance and age elsewhere), demonstrating that Midtown II was feasible.

With their qualifying test met, those 858 survivors were still a target sample on a list, whom we next had to approach individually for reinterview, without any predefined exclusions, no matter where their homes now happened to be. For this climactic, cliff-hanger purpose, we retained and cosupervised a cadre of experienced health survey interviewers on the staff of the University of Chicago's National Opinion Research Center (NORC). At closure of field work, they had conducted a total of 695 interviews with respondents scattered throughout New York City (62%), its surrounding metropolitan region (22%), 28 states beyond (14%), and while vacationing, with 14 of 18 respondents located the continental United States (2%), two as distant as Morocco and Greece.

With the exception of a few who were beyond interview reach, either geographically or by reason of a verified serious physical or mental condition, the nonparticipants had declined several written and phoned (or house call) requests for a home interview appointment communicated by NORC staff members, and as final try in every case, by the Study Director in person. Such extraordinary efforts succeeded in converting 96 (11%) of the located survivors from resisters or refusers into reinterviewees.

The 695 persons interviewed in both 1954 and 1974 comprised 81% of our accessible target sample, compared with our Midtown I 87% participation frequency. The Midtown II 81% rate can be regarded as a satisfactory achievement in light of the fact that, in recent decades, sample surveys have experienced sharp, and bias-threatening, increases in nonparticipation, with the increase largest by far in big cities and among the elderly, where Midtown II on both counts was most vulnerable. (It is revealing that among our 858 located survivors, the nonparticipation rate varied inversely with residential distance from New York City). On the basis of the 163 nonparticipants' known heterogeneous (although hardly random) baseline characteristics, representing a relatively small proportion of the 858 total, we infer as follows: Had all of the nonparticipants been reinterviewed, the findings on the main variables

with multiple regression methods applied to the 1954 Midtown corpus of adult symptom information, the computer could be programmed with an equation that parsimoniously reproduced the psychiatrists' six-grade judgmental ratings of 1954 general mental health, with an accuracy suggested by a Pearson Correlation of .83. (This compares favorably with a correlation of .75 between the 1,660 independent judgmental classifications of the two Study psychiatrists with each other in 1954. Moreover, that the psychiatrists' ratings could be predicted by multiple regression methods with such a high degree of accuracy suggests that they had been made with very considerable consistency.) The .83 coefficient assures us that the computer-generated ratings of 1954 were an acceptable surrogate for the psychiatrists' judgmental classifications of that year (Singer, Cohen, Garfinkel, & Srole, 1976).

Since we asked the identically worded battery of current symptom questions in 1974, we applied the identical regression methods to these later symptoms to produce follow-up mental health ratings. The computer and multiple regression statistics have given us a standardized, stable yardstick, one that allows us to measure and compare the general mental health status of our Midtown panel members with consistency over a considerable expanse of the adult life cycle. On the technical level, this may be not the least of Midtown III's contributions.

In this connection, parenthetically, to convert a synchronic, cross-sectional survey into a diachronic follow-up investigation, especially if conversion was not preplanned, is to multiply the technical and substantive complications exponentially. Thus, the Midtown longitudinal effort, working at baseline on relatively unexplored frontiers, with new instruments, was at especially high risk of stumbling into one or more quicksand points of empirical "no return." We are still somewhat disbelieving that none of these possible mishaps has as yet surfaced, and that all of the succession of procedural "pieces" have so far fallen into place in a manner more assuring than we could have expected on the basis of Murphy's Law ("If anything can go wrong, it will.") To successors, we pass on the suggestion that such an outcome does not emerge unless the baseline stage provides the follow-up with a robust foundation.

Summary, Explanatory Theory, and the Future

The Midtown Study confronted the Mental Paradise Lost protagonists with their arch-villain, New York City, and on that most strategic testing ground favorable to their own position, has put two of the school's three major articles of faith (Srole, 1978) under systematic empirical scrutiny.

(1) They have held that mental health in the general population has been on the downgrade over time. To this keystone of their thinking, Midtown III has responded in two parts: (A) After exposure to 20 years of living mainly in or around the City, the Midtown panelists show no significant net change in mental health composition among any of their four decade-of-birth groups. (This holds irrespective of their post-1954 residential distance from New York.) (B) On the other hand, by the temporal measure of two generation units, mental health in the later generation, with age controlled, is not worse than it had been in the earlier generation, but significantly better.

(2) Those ideologues in many cases have entertained a view of women as the

constitutionally and psychologically fragile gender, implying that, compared to men, their mental health over time has been vulnerable to erosion. Again Midtown III has countered with two generalizations: (A) After 20 years of living, the Midtown panel women of all age groups have changed no more in mental health makeup than did their male peers. (B) On the time scale of generation units, the Midtown panel's intergeneration improvement in mental health is found to be exclusively concentrated in its female ranks.

Earlier, we cautiously defined the specific population universe to which the Midtown panel's findings can be generalized, and reiterate that caution here. However, we would add that there are grounds for the hypothesis that Midtown's differential intergender progression in well-being between successive generations (A) may be underrepresented in rural and other small communities, and conversely, (B) may be principally concentrated in the big cities (Fass, 1977). If the Midtown-specific generalizations are in due course confirmed in other populations, it is predictable that the Mental Paradise Lost doctrine, spun from strands out of a long tradition of nostalgic legend and folklore, will take its place with the Ptolemaic cosmology and other intuitive prescientific formulations that evolved quasiscientific supports to sustain their popular acceptance for centuries.

Changes over Generations

The longitudinal reach of the Midtown Study enabled us to sort out the discrete changes between successive generations in the subjective well-being of their members. It is our final charge to offer a theory that might plausibly articulate the interconnections between (A) the Midtown collectivity units called generations, (B) the overarching macrosocial shifts that occurred in the separate periods of history they passed through, and (C) the sex-specific patterns of changes in well-being.

Of course, the concept of generations is one of the most multidimensional constructs in the entire repertoire of the human sciences. That it refers to a terrain of potentially bewildering proportions is documented by the huge literatures of fiction, autobiography, and biography, where it has been the scene of endless contention.

Yet our reported Midtown data insist that we address this question: How are we to explain the differential intergender shifts in well-being between the Midtown A/B generation, born in the two decades before 1915, and the C/D generation appearing during the two decades following? Because this question did not surface for us until the Midtown III stage of the study, we had not sought answers in the preceding stages. Nevertheless, in a retrospective, gender-oblivious way, Midtown II did systematically explore macrosocial and purely individual episodes, major and minor, that had marked the life course of our panelists, most episodes dating since the Midtown I interviews of 1954, some predating that occasion by ten years (e.g., World War II) or as far back as two decades (e.g., the Great Depression).

Stratification of Experience

These relatively recent life-history intrusions on a panel-wide basis remain to be analyzed for their gender-differential connections, and will be reported in a sequel article. In the meantime, toward considering the question just posed, two influences have forced us to look for empirical leads to new hypotheses reaching farther back into time than we have probed heretofore. The first influence has been the crystallizing impact of Mannheim's (1940) thought. He refers to the "phenomena of the 'stratification of experience'" in a generation group, arising from exposure to the chronological ordering of imprinting macrosocial configurations. He then emphasizes the critical distinction between the experiences of preadulthood that "happen to make those all important 'first impressions' . . . and [those] which follow to form the second, third and other 'strata' . . . Early impressions tend to coalesce into a natural view of the world. All later experiences then tend to receive their meaning from this original set . . . The continuous shift in objective conditions has its counterpart in a continuous shift in the oncoming new generations which are first to incorporate the changes in their behavioral systems" (Mannheim, 1936).

We can elaborate these important insights in terms of five propositions: (1) Given the macrosocial and microsocial changes between one historical period and the next, we can designate the behavioral milieus of the separate periods as subcultures in flux. (2) The several age-ordered generations present in a given period of time, A, can envisaged metaphorically as dominoes superimposed in a stratified ziggurat pattern, the youngest cohort at the bottom and the oldest at the top. (3) Dominant carriers of the subculture prevailing in period A are usually the cohorts in midlife who are its models. (4) The changes imminent and emergent during period A that foreshadow the subculture of period B will leave varying imprints on the hierarchy of cohorts present during period A. In specific terms of values, goals, life-styles and life-course pathways, the impacts of period-A changes will be greatest on the youth cohorts (adolescents and young adults, as in the Vietnam War years), with progressively diminishing effects on the successively higher (i.e., older) cohorts. (5) The midlife models of subculture A and the youth-stage, forthcoming models of subculture B are in relationship of contention that can be characterized as the "muted battle of the generations," across what the French call *fosse*, meaning "ditch" or "moat."

The point of key relevance here is the paramount importance of the subculture milieu during the formative stage of a generation's development. Although that stage was by no means untapped in the Midtown I and II interviews, we lacked the prescience to try to conceptualize and systematically recapture it retrospectively, if indeed that would have been possible at those times in terms relevant for present purposes.

The second influence to push the inquiry back into time, anecdotal in nature, was the coincidence that one of us (L.S.) happens to be a contemporary of the Midtown panel's cohort B, timing him to be a schoolboy hearsay-witness of World War I, and a participant in its aftermath. In the midst of those social upheavals, he was himself caught up in the tradition-shattering changes that were sweeping through the arena of intergender relationship, of course without any insight as to their historical significance or foresight as to their future long-range consequences. Those experiences remained vivid in

memory, but lacked an adequate historical perspective until the surprising Midtown cohort data were in hand, sending us to documentary works on that period and the decades immediately before and after. This literature fortunately has been highly relevant for a credible theoretical response to our question about the possible main sources of the Midtown panel's inter-generation-cum-intergender differences in well-being, compelling us here to share its illuminations and implications for a historically grounded social psychiatry.

Women in the Victorian Era and After

To face our question adequately, we must turn back to the last decades of the 19th century, which was the temporal habitat of the parents of Midtown's cohorts A and B. The distinguished historian J.H. Plumb needs only two sentences to summarize the predicament of women in that culminating period of the Victorian era: "Except in the highest ranks of the aristocracy, women . . . were as securely locked in the prison of their households as any convict . . . condemned to a daily treadmill of toil . . . Husbands could be and were tyrants" (Plumb, 1973).

Kinder, Kuche, Kirche fixed the specific confines of a woman's life, with children one-after-the-other pressing the limits of her gestation cycles and emotional reserves. Women were the victims of a second and far more penetrating contradiction: women then were placed as cult symbols of sexless purity on a public pedestal, but privately were locked into a chastity belt, so to speak, of tight moral proscription of sexual expression and impulse; all of this, while they were also forced to give silent lip service to the unwritten double standard tacitly enabling men to give their sexuality free rein in or out of the marital bed.

A more pathogenic, culturally patterned design for intergender relationships can hardly be imagined, a point confirmed by Professor Plumb's reference to the period's "agonizing preoccupation with sin, combined with a Jehovah-like inflation of the figure of the father . . . Fearful repression of sex was followed, as might be expected, by life-destroying neuroses . . . The whole Victorian scene is littered with broken minds, broken lives and broken careers" (Plumb, 1973). Students of hysteria (Hinside, & Campbell, 1970; Freedman, Kaplan & Sadlock, 1975) during the late 19th century emphasized that it was "a disease of the female sex almost exclusively . . . [that] its prevalence reached almost epidemic proportions during the reign of Queen Victoria," and that far from being wholly a middle or upper class phenomenon, "it occurred frequently [among the poor], and chiefly among the more respectable poor who, 'resisting the normal effects of passion, fall victim to the abnormal'" (Veith, 1965). Sigmund Freud himself was a product of this period of history, turning in the 1890's to a preoccupation with hysterics and their underlying, repressed sexual conflicts that were one of the prime forces impelling his subsequent breakthrough into the depths of the human psyche and and psychoanalytic theory.

It was this Victorian generation, concludes Plumb, that carried "enormous burdens of anxiety that within this last [20th] century . . . have lifted off the shoulders of men and women, particularly in the highly industrialized West, to a degree that they can scarcely appreciate" (Plumb, 1973). It is

therefore plausible to infer, with Plumb, that during this Victorian period, mental disturbance in the general population had been at higher levels of frequency and more severe levels of intensity, above all among its submerged women, than were to ever prevail thereafter.

Parenthetically, balancing all the pluses and the many remaining minuses of the present period, we hold the considered opinion that both the average man average woman beyond the age of 40 years are objectively and subjectively healthier than ever before in the history of this Republic. We were able to put this view to a partial test in 1974, by asking our Midtown panelists the following question: "Your parents grew up to adulthood in a very different period from the young people of today. In your opinion, which generation grew up in a better period to enjoy life: Your parents' generation, your generation, or today's younger generation?" In choosing among these alternatives, only 22% of our panel members replied "my parents' generation," fully twice as many (43%) declared "my own generation," and another 35% chose "today's younger generation" as growing up in the best period. In other words, almost 80% of the panel members perceive progress in the quality of the social worlds that have emerged at least since the generation of their grandparents. It is of particular interest that whereas significantly more panel men than women prefer parents' generation (28% to 19%), valuing "my own generation" most were 33% of the men, and 50.3% of the women. In any case, members of the last Victorian generation were contemporaries and counterparts of the parents who brought up the Midtown panel's own cohorts A and B, offspring appearing, as we have already seen, with disparities on all three indicators of subjective well-being that were manifestly unfavorable to their women.

Young women now, writes Fass, "defined equality not as political rights or economic opportunities, but as something more subtle: freedom . . . the right to self-expression, self-determination and personal satisfaction." In short, they intended to steer the life course primarily by their own lights and not, as in the past, exclusively by those of their men. With this shift, notes Fass, there followed "a democratization of family relations between husband and wife, parents and children, and more latitude for emotional expression for each member of the family."

A related consequence was "an expanded view of children's welfare," including "a trend toward more equal involvement with, and affection for, each child, regardless of sex." In light of the fact that the Victorian family had previously regarded the preadult son as "a man still growing," and the daughter in effect as "ever a little girl," this new parental "promotion" of the latter toward equality with her brothers probably represented a definite eugenic gain, especially in her own eyes, of socially valued status and personal self-esteem.

This, then, was the context of the generation that emerged during and after World War I, which was first socialized in the "derigidified" family and age-peer milieus of the 1920s, and appeared decades later in the Midtown Study panel as adult cohorts C and D. We find an exemplar of the women of this generation in Rosalyn Yalow, in 1921 born of immigrant parents in a working class New York City environment that could prompt her, at age 8 years, to announce her awareness that she could become not only a wife and mother, but also a career scientist. She, of course, achieved all three goals, and more, becoming a researched in the almost Frauen-frei field of physics, and at age 56, a Nobel Laureate in Medicine (Srole, 1978).

This is the generation that in the last two decades has emerged as a "new breed of women," who in growing numbers have been cutting through the barriers to university graduate schools, into male-dominated professional, semiprofessional, and other occupational strongholds, to launch the revived feminist movement and to enlarge their presence and voice in political forums at all levels (Kirkpatrick, 1976).

All of the above post-World War I milieu developments can be considered sociologically isomorphic with our observed cohort C and D women's quantum jumps in subjective well-being, representing advances over that of their A and B "preliberated" predecessors at like age, and probably even more over that of their late Victorian grandmothers who suffered the fate of a culturally contrived "iron maiden" confinement.

Midtown Results and the Future

It is a credible hypothesis that isomorphisms on such a three-generation time scale are not mere chance coincidences of the play of history, but reflect a cause-and-effect connection between the partial emancipation of women from their 19th century status of sexist servitude, and their 20th century advances in subjective well-being. Out "control subjects" in this "experiment of human nature" on the psychological consequences of social change are the Midtown panel men who had no comparable sexist "chains to lose," and showed no comparable progress on our three measures of subjective well-being. Thus, we are led to suggest the following law of sociopsychology: Improvements in a group's social position and role in a society's objective system of status allocations are conducive to improvements in that group's subjective well-being and other dimensions of health. Implicit here is the notion of relative self-fulfillment as a eugenic, socially generated counterpart of the sociopathogenesis-denoting concept of relative deprivation.

Unless unexpected technical artifacts of both sex-specific and generation-specific kinds should turn up to suggest otherwise, the magnitude of our cohort C and D women's gains in subjective well-being must be seriously assessed for their potential social policy implications. Our figures, as they now stand, support estimates, that may be conservative, of a cut of 50% to 60% in the cohort C and D female mental impairment rates below the levels of their A and B counterparts, and a roughly corresponding increase in their wellness frequencies. Consistent with these tentative estimates, one of us (L.S.) has elsewhere mounted strong circumstantial evidence that parallel long-term trends of liberation from the demeaning deprivations of poverty have had similar eugenic mental health effects (Srole, et al., 1962, 1978).

To suggest the further implications of such separate kinds of evidence, we hold that a substantial part of the psychopathology at large in the population is precipitated by long refractory, discriminatory dysfunctions that are foisted on specific, power-weak community subgroups, damaging their members and subverting the most basic Judeo-Christian canons of a democratic society. Since these social pathologies are legally and politically correctable, it can hardly be denied that in the calculus of both humanist and cost/benefit values the surest primary preventive medicine lies in a general policy of making accessible larger dosages of social equality to groups where it is in less than health-sustaining supply. (Most immediately, this emphatically

includes those unyielding points of female second-class citizenship that remain to be corrected through the instrument of the pending, unconscionably stalled, Equal Right Amendment to the American Constitution.)

However, if our findings have challenged the Mental Paradise Lost doctrine, we are not implying an expectation that the broad policy just recommended can alone bring about a Mental Paradise Regained in our enormously complex society, although it can help the latter advance much farther from its "dystopia" antecedents in the 19th century. Even so, there will still remain a large reservoir of (A) discordances between individual psychic capacities and the marketplace of demanding adult roles, and (B) circumscribed intrafamily pathologies that are destructive for children. Both of these will, of course, require the interventions of secondary preventive psychiatry, of established and newer modalities, toward fostering improved intra-psychic and interpersonal functioning. Also relevant here is the recent warning of Erik Erikson (New York Times, 1979) that "just as sexual repression characterized the Victorian era, there is a real danger that a new kind of repression may become a mark of adult life, namely repression of the urge to have children."

Nowhere, we believe, are the massive behavioral impacts and imprints of objectively changing social conditions more clearly manifest than in the procession of successive generations, each marching, to its "own music." Prompted by our compelling longitudinal findings, we have reached into the historians' archives going back more than a century to the decade that began about 1870. Looking toward the future, we project two further sets of follow-up interviews: Midtown IV, focusing on our cohorts C and D, to intensively test and elaborate the theory offered above, and alternative hypotheses, toward explaining the gender's differential intergeneration advances in subjective well-being.

Midtown V will follow up the 1,160 adult children of our four cohorts, their births spanning the years 1920 to 1960. As preparation, in the 1974 interviews with their parents, we systematically secured eight items of recent life-situation information (including psychiatric episodes) about each of those children. With such new baseline information already stored in our growing data bank, we will be able to cover four generations since the Midtown panel's earliest births in 1895.

Envisaged thereby is furtherance of an epidemiological nucleus for the crystallization of the specialty multidisciplinary field of sociopsychiatric history. On its agenda would be research questions such as the following: Will the newer adult generations of the 1980s change in levels of well-being, relative to the baselines of their parents, in a linear or cyclical, pendulum, fashion? When is a predominantly affluent society based on eugenic freedom at risk of slipping into a pathogenic, self-destructive system of narcissistic license, an outcome not without historical precedents? At least speculatively, what might be the mechanism inherent in a specific historical macro-social universe, for the differentially potentiating the gene pool of the preadult generation engulfed in it?

The answers to such questions might suggest prognoses as to the well-being of democracy itself in the century to be turned just 20 years ahead.

Table I

Midtown Manhattan
Follow-Up Panel (N=695)
General Mental Health Distributions
in 1954 and 1974

Symptom Formation Grades	GMH I	GMH II
1. Well	21.9%	25.0%
2. Mild	42.6%	42.0%
3. Moderate	21.3%	21.1%
4. Marked	9.3%	7.5%
5. Severe	4.3%	3.5%
6. Incapacitated	0.6%	0.9%
4-6 Impaired	14.3%	11.9%

References

- Bellow, S. quoted in L. Srole, & A. Fischer (Eds.), Mental health in the Metropolis. New York: New York University Press, 1978.
- Fass, P.S. The damned and the beautiful: American youth in the 1920s. Oxford, England: Oxford University Press, 1977.
- Freedman, A.M., Kaplan, H.I., & Sadock, B.J. Comprehensive textbook of psychiatry, III. Baltimore: Williams & Wilkins Co., 1975.
- Fromm, E. Escape from freedom. New York: Holt Rinehart & Winston, Inc., 1941.
- Fromm, E. The anatomy of destructiveness. New York: Holt Rinehart & Winston Inc., 1973.
- Goldhamer, H., & Marshall, A. Psychosis and civilization. Glencoe, Ill.: Free Press, 1943.
- Gordon, T. & Kannel, W.B. The Framingham, Massachusetts Study, twenty years later. In I.J. Kessler & M.I. Levin, (Eds.), The community as an epidemiological laboratory. Baltimore: John Hopkins University Press, 1970.
- Hayden, C.B. On the distribution of insanity in the United States. Southern Literary Messenger, 1844, 10, 178-181.
- Hesse, S.J. Women working: Historical trends. In K.W. Feinstein (Ed.), Working women and families. Beverly Hills, Calif.: Sage Publications, 1979.
- Hinsie, E., & Campbell, R.J. Psychiatric dictionary. New York: Oxford University Press, 1970.
- Kiev, A. The role of expectancy in behavioral change. In J.P. Brady, & H.K. Brodie (Eds.), Controversy in psychiatry. Philadelphia: W.B. Saunders Co., 1978.
- Kirkpatrick, J. The new presidential elite: Men and women in national politics. New York: Russell Sage Foundation, 1976.
- Leighton, A.H., Leighton, D.C., & Danley, R.A. Validity in mental health surveys. Canadian Psychiatry Association Journal, 1966, 11, 1670178.
- Luborsky, L. Clinicians' judgments of mental health: A proposed scale. Archives of General Psychiatry, 1962, 7, 407-417.
- Mannheim, K. Ideology and utopia: An Introduction to the sociology of knowledge. L. Wirth, & E.A. Shils (Trans.). New York: Harcourt Brace Jovanovich, Inc., 1936.
- Mannheim, K. Man and society in an age of reconstruction. New York: Harcourt Brace Jananovich, Inc., 1940.
- Myrdal, G. An American dilemma. New York: Harper & Row Publishers, Inc., 1940.
- Plumb, J.H. In the light of history. Boston: Houghton Mifflin Co., 1973.
- Sartwell, P.E. Problems of identification of cases of chronic disease. Milbank Memorial Fund Quarterly, 1953, 31, 17.
- Singer, E., Cohen, S., Garfinkel, R., & Srole, L. Replicating psychiatric ratings through multiple regression analysis: The Midtown Manhattan Restudy. Journal of Health and Social Behavior, 1976, 17, 376-387.
- Singer, E., Garfinkel, R., Cohen, S., & Srole, L. Mortality and mental health: Evidence from the Midtown Manhattan Restudy. Social Science and Medicine, 1976, 10, 517-525.
- Srole, L. Social integration and certain corollaries. American Sociological Review, 1956, 21, 709-716.
- Srole, L. Measurement and classification in sociopsychiatric epidemiology: The Midtown Manhattan Study (1954) and Midtown Manhattan Restudy (1974). Journal of Health and Social Behavior, 1975, 16, 347-364.
- Srole, L. The city vs. town and country: New evidence on an ancient bias. In L. Srole, & A. Fischer (Eds.), Mental health in the metropolis. New York: New York University Press, 1978.
- Srole, L., Langner, T.S., Michael, S.T., Kirkpatrick, J., & Rennie, T. Mental health in the metropolis: The Midtown Manhattan Study. New York: McGraw-Hill Book Co., 1962.
- Srole, L., Langner, T.S., Michael, S.T., Kirkpatrick, J., & Rennie, T. Mental health in the metropolis. Revised and enlarged edition. New York: New York University Press, 1978.
- Stouffer, S., Lumsdaine, A.A., Lumsdaine, M.H., et al. The American soldier. Vol. 2 & 4. Princeton, N.J.: Princeton University Press, 1949.
- U.S. Center for Disease Control. Leukemia among persons present at an atmospheric nuclear test. Morbidity and Mortality Weekly Report, 1979, 28, 361-362.
- Veith, I. Hysteria: The history of a disease. Chicago: University of Chicago Press, 1965.
- Witson, C.L., & Hunt, W.A. A rationale for psychiatric selection. American Psychologist, 1955, 10, 199-204.

City Children:
The Family Research Project

Thomas S. Langner
Professor of Clinical Psychiatric Epidemiology (in Public Health),
Columbia University

City Children:
The Family Research Project

Thomas S. Langner
Professor of Clinical Psychiatric Epidemiology (in Public Health),
Columbia University

Aims of the Study

The general objective of the study is to relate environmental stress to behavioral disorder in children and adolescents over time, in order to develop etiological weights and contributions. The child's total environment was not assessed, especially the physical environment (i.e., toxic factors, such as air pollution and lead poisoning). An emphasis was placed upon the family and broad social environment, and some attention was given to physical illness, accidents, and handicaps. A brief family history of emotional problems and their treatment was collected, but no attempt was made to obtain materials as detailed as those in genetic high risk studies, such as accurate parental diagnoses. Only global measures of parental disturbance were made, and their educational, marital, and police history was obtained. Given this bias toward assessing the social environment, we cannot speak to the relative strength of air pollution, or diet, as opposed to social and familial factors. Between the broad groups of demographic and biosocial variables (age, sex, race and socioeconomic status), marital-parental variables (qualities of the marriage) and parental practices, we have much to say about relative contributions to behavior variance of different types, both concurrent and long-term (5 to 7 years).

A second aim of the study was to classify the environments to which urban children are exposed in some replicable fashion. A third goal was to classify types of predominantly untreated children in a fashion more reliable and more predictive of future impairment and psychopathology. Early identification and classification was seen as a first step toward primary prevention. Changes in children and in families over time were to be measured, and child-change was to be related to family and social change. Of somewhat lower priority, but

of great importance for planning services, was the question of the prevalence and incidence rates of general impairment in the population of each sex at various ages, the ethnic and social class distribution, and the proportion of children with severe impairment who were referred and treated for their problems. The determination of the treated proportions for various types of disorder, and the comparison of change in treated versus untreated children over time, were never primary goals of the study, but clinical audiences focused on these questions, often to the exclusion of questions of etiology. Special questions were also a focus of the study, such as whether excessively stressful environments in childhood and adolescence might account for the heavy predominance of mental disorders in minorities and the poor found in most field studies of mental disorder in adults.

The Study Methods

Cross-sectional Sample: Time I. The sample area was defined as the area between Houston Street and 125th Streets on the East and West sides of Manhattan, New York City. This area included all major ethnic and income groups of Manhattan. The sample was broken down into health area, blocks within each health area, and dwelling units in each block. The total number of dwelling units in each block was translated into clusters, where, on the average, each cluster consisted of eight dwelling units. Thus, the sample was one of systematic cluster sampling, stratified by health area. A team of enumerators contacted each household that fell into the sample to determine if it contained a child between the ages of six and 18 and thus was eligible for the study. When an eligible household fell into the sample, an appointment was made by the enumerator for a trained interviewer to come and interview the mother of the household. The sample was completed with the relatively low refusal rate of 15.6 percent. The breakdown by the major ethnic groups of the final cross-sectional sample of 1,034 families was 56 percent White, 29 percent Spanish-speaking, 14 percent Black, and 15 percent Other.

Welfare AFDC Sample: Time I. The Welfare AFDC households were randomly selected from Welfare rolls which covered the same area of Manhattan as that of the cross-sectional sample. Eligibility was once again determined by the presence of a child between six and 18 years old in the family. The sample was stratified by ethnic background, to approximate equal thirds of the major ethnic groups (27% White, 36% Spanish-speaking, 37% Black) in the final sample of 1,000 AFDC families. The refusal rate for this sample was a low 8 percent.

In both the cross-section and Welfare samples each age group except the oldest comprised nearly one-thirteenth of the sample, and males and females were fairly evenly distributed across the 13 age

age groups.

Follow-up Study: Time II. Approximately five years later, the two samples were followed up. The first wave of interviewing was conducted in 1966 and 1967, the second wave in 1971 and 1972. The follow-up was conducted in such a manner as to ensure constancy of the ethnic proportions at both time points. This rule set a lower bound to the follow-up, since many of the Spanish-speaking families had moved out of the city and could not be located. In the cross-sectional sample, a total of 732 families or 71 percent of the original sample was reinterviewed. In the Welfare sample, a total of 661 families or 66 percent of the original sample constituted the follow-up sample. The follow-up (or Time II) samples did not differ significantly from the Time I samples on any of the central measures, namely, age and sex of the study child, demographic characteristics, marital and parenting dimensions, and child behavior dimensions. The follow-up samples thus represented unbiased subsamples of the original samples.

Procedure

The Mother's Interview. In both Time I and Time II, mothers were interviewed for an average of two and one-fourth hours. The questionnaire was, for the most part, a structured instrument, with some verbatim response material included. The questionnaire was designed to solicit information on the development and current behavior of the study child, aspects of parental character and the marital relationship, child-rearing practices, and a broad range of demographic variables. The questionnaires for the Welfare and cross-sectional samples were identical except for an additional section in the Welfare questionnaire on Welfare-related attitudes and behavior. The entire questionnaire was translated into Spanish for the large number of Puerto Rican respondents, and interviewers and respondents were matched by ethnic background in the interest of establishing maximum rapport.

Psychiatrist's Rating of Questionnaire. At Time I, a computer summary of the questionnaire information dealing only with the child's behavior (654 items) was used by two of the three project psychiatrists to rate each child on a five-point total impairment and five equaled severe impairment. The distribution of ratings from each rater was transformed into a distribution of standard scores. The reliability for the average of two psychiatrists for the total impairment rating (TIR) was .84. Validity information for this rating was summarized earlier (Langner, Gersten, Greene, Eisenberg, Herson, and McCarthy, 1974).

Factor Analysis of Questionnaire Data. The total pool of 654 child items in the mother's questionnaire was reduced by first eliminating those items with low frequencies and age or sex contingency. A second group of about 200 items were collapsed into a set of composite scores (number of fears, illnesses, etc.). Independently of their psychiatric

evaluation, the 287 items which remained were then factor analyzed (orthogonal varimax rotation) forming 18 dimensions using a total of 222 items. This analysis was conducted on the cross-sectional sample pool. The Welfare sample data were recoded and scored on each item. This was done to allow direct comparisons on the factors between samples. (For a detailed description of the child factors see Eisenberg, Langner, and Gersten, 1975, and Langner et al., 1980). The internal consistency reliabilities for Time I ranged from .72 to .94 (cross-section) and .57 to .92 (Welfare). At Time II they were .56 to .93 (cross-section); .57 to .93 (Welfare).

Familial Factors. Questionnaire items describing the marital relationship the character and personality of the mother and father, and the mother's physical and emotional health were factor analyzed at Time I. A total of 91 items comprised 8 parental factors. Questionnaire items describing the parent-child relationship, child-rearing practices etc. were also factored. Five-parent-child factors included 81 questionnaire items. This analysis was conducted on the cross-section item pool. The Welfare sample data were coded and scored on each item. Standard Z scores were computed for the Time I and Time II sample data on each of the 13 familial factors. The internal consistency reliability coefficients for the familial factors at Time I ranged from .66 to .88 (cross-section) and .42 to .85 (Welfare); at Time II; .58 to .85 (cross-section); .36 to .87 (Welfare).

Record Data

School records were obtained for 88 percent of the children in the follow-up sample. The continuous school record was obtained and the coding system applied to each year of the child's school career, noting its temporal relationship to the Time I and Time II interviews.

Information on officially recorded delinquent/criminal behavior was collected via a search of the records of Family Court and the Police Department (both youth and adult records). Information from agency records was transformed into a coding form. The coding-measurement indicated the data of occurrence for each event, relevant aspects of its content, and the offense charge(s) applied. For the purposes of the following analysis a police event (on a given date) was counted as a single charged offense, although various types of charges might have been involved.

Selected Findings of the Study

The findings of the Family Research Project (F.R.P.) are voluminous, partly due to the fact that the study has been supported for ten

years by a variety of sponsors and partly due to the large number of variables collected from many different sources, such as the mother, the child, school, police and probation records, and social agencies. These data were collected by different types of people, such as interviewers, psychiatrists, and record coders. Only a small selection of the findings are given here.

Prevalence: Children were rated on five-point scales of impairment by at least two of 3 psychiatrists, from summaries of their mother-reported symptomatology, without information on any environmental or biosocial variables except age and sex.

In the cross-section 13.5 percent (based on the average rating distribution percentages of 3 psychiatrists) were rated as severely impaired or incapacitated (4 or 5 on the Total Impairment scale). This group was considered to be in need of immediate intervention.

The prevalence of Child Types was assessed by the construction of child profiles based on hierarchical cluster analysis, (a Euclidean distance model) using eighteen Time I factored dimensions.

In the cross-section sample (from which most of the data in this report were taken) about two-thirds of the children were classified in healthier categories (lower impairment, fewer symptoms generally). These profile types were 16 percent sociable, 12 percent competitive-independent, and 34 percent dependent. About one-third were classified in impaired groups; 16 percent moderate backward-isolates, 4 percent severe backward-isolates, 12 percent aggressive, 0.7 percent organic, 2.8 percent Self-and-Other Destructive, and 1.5 percent Delusional. The children in these groups were rated between 15 and 57 percent impaired (4+), 14 to 29 percent had been arrested one or more times, and 34 to 54 percent had poor school or police outcomes. (Figures are conservative, due to placing of rare groups into larger categories).

Predictors of Child Behavior and Outcomes

There are few studies of large random samples of children using multi-dimensional measures of mental disorder and multidimensional predictors of disorder which can be ranked by their power. These predictors are leads for prevention. In rank order of power, they are; 1) Being Spanish-speaking; 2) having Punitive Parents; 3) having Cold Parents; 4) being Black; 5) having an Excitable-Rejecting Mother; 6) having an Emotionally Ill Mother; 7) a Large Number of Children in the Family; 8) High Rent; 9) having a Traditional-Restrictive Mother; 10) a Large Number of Addresses (Moving); 11) Number of Natural Parents the child lives with (mostly loss of father).

Changes in Parenting and Child Behavior, Overall Predictive Power, and Policy Implications

Changes in parents over the average 5 year period showed changes in children. If parents became warmer or less rejecting, children showed less Conflict with Parents and less Anxiety, Fighting, Mentation Problems, Isolation, and Delinquency. Changes in Parental Coldness and Excitability were tied to changes in aggressive behavior in all settings. Changes in Parental Punitiveness were only related to changes in intrafamilial aggression, but not outside the home. Any one predictor's unique contribution (what it alone contributes) to variation in child behavior is not great, explaining a maximum of 10 percent of a particular behavior (multiple correlation of .78). This still leaves out much of the environment that has not been measured.

Changing the child's environment might change as much as half of his behavior, if all factors were ameliorated. Minimal changes would be outlawing of child-beating, reducing coldness by encouraging parents to touch, hug and kiss children in a massive educational and conditioning program, making parental screaming and lability socially undesirable, using birth control to cut down unwanted children, and reducing the bulldozing of neighborhoods by rehabilitation, thus limiting uprooting. Support for marriage via full employment, counseling, and changes in the Welfare system can reduce broken homes. Reduction of the high risk of Black and Puerto-Rican children overlaps the other preventive techniques, but discrimination and accompanying low self-esteem are additional primary targets.

Age and Sex Effects

In general, age showed strong effect, and sex weak effect in predicting behavior. Delinquency and Undemandingness increased with age, while Weak Group Membership and Dependency decreased. Anxiety, Training Problems, Sex Curiosity, and Fighting, among other behaviors, showed sizeable decreases with advancing age (prevalence). A large number of behaviors showed on unique age contribution.

Boys exhibited more Conflict with Parents, Noncompulsivity (mild antisocial behavior) and Delinquency, while girls showed more Regressive Anxiety. None of these unique sex contributions to behavior variance exceeded 2.7 percent, while most were under 1 percent (considered a meaningful association, as opposed to a smaller figure which is statistically significant). Age contributions range from 0 percent to 15.6 percent for Delinquency (increasing) and 14.3 percent for Regressive Anxiety (decreasing).

Family Factors vs. Social (Demographic) Factors in Children's Behavior

It is important for many reasons, not the least of which is prevention, to know the relative importance of family life variables as opposed to broader social (demographic) variables in the development of child behavior and psychopathology. The family variables are those dealing with aspects of the parents' relationship and their character (eight factors) and those covering the relationship between parent and child, or parenting factors. The demographic or social variables (such as ethnic background, number of addresses, rent, number of natural parents the child lives with, and the number of children in the household, all made unique contributions of 1 percent or more to child behavior of various types. Family variables were more strongly associated with child behaviors involving interpersonal relationship (such as Conflict with Parents and Weak Group Membership) aggressive behavior (Fighting) and neurotic behavior (Regressive Anxiety). Demographic or social factors made little contribution in these areas.

The demographic variables made greater contributions than family variables in the areas of developmental or organic pathology (Isolation, Mentation Problems and Training Difficulties). This was also reflected in the areas of impairment which were ratings of particular role or biological functioning in specific behavior settings (family, school, peer). The ratings of Impairment of Development were both related somewhat more to social than family variables. Impairment ratings of Self-Confidence, and Impairment in relations with Father, Mother, Siblings and Peers, were all more strongly associated with family than with demographic variables. The contrast of the interpersonal as opposed to the organic-developmental seems to hold, whether more objective scores or psychiatric judgements are used as the criterion variables.

Stability of Child Behavior

Stability coefficients were calculated for the 732 children who remained in the cross-sectional sample at Time 2, after five years. The focus was on dimensions strongly correlated with the psychiatric impairment rating at Time 1. Six dimensions of behavior (those most related to impairment), had correlations ranging from .36 to .54; Mentation Problems, Conflict with Parents, Regressive Anxiety, Fighting, Delinquency and Isolation. On three of these which could be called aggressive behavior (Conflict with Parents, Fighting and Delinquency) there was either the same level or a mean increase in pathology between Time 1 and 2. When these same behavior dimensions were examined by age groups (rather than cohorts) at a point in time (Time 1), the results were in disagreement with the true longitudinal analysis half the time.

Suggestions for intervention are that antisocial behavior

2

becomes established or stable only at or after age 10. Treatment for antisocial behavior, if effective at all, would be best started at, or after, preadolescence. Treatment for neurotic disturbances would involve great instability and much spontaneous remission before the beginning of adolescence. Intervention before middle adolescence seems to be inappropriate, since much of the behavior is outgrown up to that time. Conflict with Parents and Mentation have stability patterns which suggest intervention before age 6 would be effective for intrafamilial aggression or for memory, speech, concentration problems and intellectual functioning.

Family Types

To create Family Types, a total of thirteen factor scores (8 parental and 5 parent-child factors) were standardized along with some demographic factors (ethnic background, Welfare status, child living with both, one or no natural parents, child always in mother's care, number of children in the household, monthly rent, number of addresses in New York City). The same hierarchical cluster analysis technique (McKeon, 1967) was used to create the types. The types are named for elevations of 0.5 standard deviations or more from the mean. A new set of types quite similar to the previous types was created using only those families found in the Time 2 follow-up (732) who also fit into a Family Change Type (total 704). The proportions of the new Time 1 types in this group were (in order of increasing environmental pathology) the White Affluent Cohesive type (28.8%), White Affluent Discordant (15.9%), Single Non-Biological Parent and Child (7.0%), Spanish Low Status Traditional Cohesive (24.3%), Spanish Welfare Isolated (5.1%) and Low Middle Status Ethnic Mix Discordant (18.9%). It should be noted that there are clearly contrasting types of White and Hispanic families. Black families were concentrated in the Mixed Ethnic type, where they comprised 68 percent of the group. While this type showed the greatest family pathology, there was no corresponding predominantly Black family type found in a subsequent analysis of the Welfare sample. In that sample, traditional Spanish families showed moderate pathology and somewhat more than a non-Spanish group with relatively educated well-adjusted mothers (both Black and White). Ethnic background, especially being Black, did not enter as much into the Welfare typology as it did the representative sample.

Family Types as Predictors of Behavior

On the basis of the concurrent association at Time 1 between the classification of family types and rates of serious impairment or a classification of child types, hypotheses were deduced regarding the predictive value of the family typology for future pathology. To

9

effectively test the construct and predictive validity of the typology, it was necessary to test it using outcome measures which were not drawn from the same data source from which the family types were developed, i.e. the mother's questionnaire. The hypothesis then stated that the risk for future outcomes be differentiated by family type. In other words, the relative risk for one or more future pathologies would escalate as the demographic and/or familial stresses within the family increase. For the six Time 1 family types represented in the follow-up sample, the percentage of children who evidenced one of the following four adverse outcomes after the Time 1 interview was determined: held over, discontinued, or expelled from school as ascertained from school records or had a charged offense in the Police-recorded information. The most advantageous family, White Affluent Cohesive, evidenced the lowest risk for at least one of these adverse outcomes (8.9%) followed by White Affluent Discordant (13.4%), Single Non-Biological Parent and Child (30.6%), Spanish Low-Status Traditional Cohesive (33.3%), Low-Middle Status Discordant (42.9%) and Spanish Welfare Isolated (58.3%). Children living in a family which was considered stressful (Low-Middle Status Discordant) had nearly five times greater risk for evidencing at least one of these adverse outcomes than did children in the family type with greatest advantages and fewest stresses. In fact, four to six out of every 10 children living in respectively Low-Middle Status Discordant or Spanish Welfare Isolated families will show at least one adverse outcome in the future.

The use of multiple adverse outcomes (2 or more, 3 or more) produced similar risks varying greatly across Family Types, and in approximately the same rank order. The risk ratios became greater, however. Thus, 20 percent of the children in the most stressful family type had 3 or more adverse outcomes, while less than 1 percent of children in the least stressful Family Type did. Many other single behavioral outcomes were examined with respect to the Time 1 Family Type of the child. These findings are reported in (1) (37) and (19).

Family Change Types

The question now at issue was to ascertain if the risk of future pathology, as measured by the outcome variables used above, varied as a function of the consistency in quality of the environment. A hierarchical profile analysis was performed on the standardized regressed change scores of demographic, parental and parent-child variables to develop patterns of familial change. These types or patterns were then cross-classified with the initial family patterns in order to arrive at a set of environments which exhibited variance or invariance over the course of the follow-up period. Within each family type, three basic patterns were developed: Stable, Worse or

Better. Depending on the nature of the original family type and in order to compensate for ceiling or floor effects, Stable was combined with either the Better or Worse change pattern. When the variable of at least one adverse outcome was examined, the findings for five of the six family types were in accord with expectation. Namely, a favorable environment which changed for the worse, had a higher percentage of children who exhibited one adverse outcome than if that environment was stable or changed for the better. Likewise, an unfavorable environment which changed for the better had a lower percentage of children at risk than if constancy or worsening was noted over time.

Child Change Types

Profiles were developed of the regressed change scores of the dimensions of behavioral disturbance in children and adolescents. Four major change profiles were abstracted: 1) an improvement profile; 2) a profile of stability; 3) a profile of worsening on most dimensions; and 4) a profile of increased aggression. To exemplify the results, it was found that among non-affluent families who were discordant at Time 1 and then evidenced a worsening in their familial conditions, 41.4 percent of their children worsened over time. In contrast, 24.4 percent of the children displayed a worsening profile who were initially in the same type of family at Time 1, but the family remained stable, and only 3.3 percent of the children worsened whose identical initial environment improved over time. In other words, while the children were exposed to the same stressful environment at Time 1, a change for the worse in the environment resulted in a twelvefold increase in risk for future disturbance than did an improvement. Stability in such an environment, however, carried a substantial risk for deterioration in behavior with nearly one-fourth of the children worsening.

Relative Power of Child Behavior vs. Child Environment To Predict Later Behavior

The question was asked "Does early child behavior or early environment best predict later child behavior?" The power of the early Family Types far exceeded that of the Child Types (early behavior) to predict arrests, school problems, and mother-reported outcomes five years later. The question of whether child change or family change was better in accounting for later child behavior outcomes from record data was also addressed.

Prediction of Child Behavior Change, and Outcomes of Change Types

As described before from the raw data on behavior change between the Time 1 and Time 2 measurements, base-free change scores were computed for each of 16 Child Behavior Factors. The 16 regressed change scores were then used as input to a hierarchical clustering program, to obtain an empirical typology of change. Four major change types were noted, termed More Aggressive (change in the direction of increased fighting and conflict with parents), General Improving (general decreases in anxiety and interpersonal conflict), Stable Independent (no appreciable change), and General Worsening (increases in mentation problems, anxiety, aggression, and isolation). Four other change types with small numbers of cases represented various types of psychotic deterioration (increases in delusions, hallucination, self-destructiveness, repetitive motor behavior), and were collectively termed Other Worsening. A discriminant function analysis was performed to determine the prognostic measures in the Time 1 data that predicted various types of change. The discriminant analysis was generally successful, with 45 percent - 50 percent successful classification for the major categories, and an overall classification accuracy of 40 percent. There were two major discriminant functions, accounting for 53 percent and 31 percent of the explained variance; these functions were loaded by variables including child's age, parents' socioeconomic status, child-rearing practices, and marital relationship. The More Aggressive change type was characteristic for younger children with relatively punitive parents of higher socioeconomic status; the General Improving type was characteristic of older children of Spanish-speaking parents who were relatively nonpunitive; and the Stable Independent type was characteristic for older children of higher-SES parents (without regard to ethnic group) who had a relatively happy marriage. The General Worsening type was related to some extent to the following combination of attributes: Younger child, lower-SES parents with unhappy marriage, and AFDC Welfare status. Analysis of Time 2 outcome measures showed the General Worsening type to have significantly worse outcomes and the Stable Independent type to have significantly better outcomes. Analysis of treatment status showed the expected pattern, with children in the More Aggressive, General Worsening, and Other Worsening types more frequently referred for psychological treatment by Time 2. Interestingly, it was noted in the Time 1 data that the Stable Independent children were more likely to have been initially referred for treatment, although they had the least need for it; probably this is attributable to the higher socioeconomic status of the parents of these children.

Effect of Severely Disturbed Parents on their Children

It is generally believed, and it has been widely found, that children of mentally disturbed parents tend to have emotional disturbance themselves. However, many such studies deal specifically with schizophrenia,

which has shown a definite genetic connection. High risk studies with cross-adoption techniques have established this connection, but they have, by and large, neglected to measure the parental behavior toward the child carefully, and multidimensionally.

An examination of behavior measures of children of 22 cross-section and 36 Welfare families with one or both parents who had been hospitalized before or during the study period found no difference between their mean scores and those of the general child population. This was also true for the children of 142 cross-section and 193 Welfare out-patient parents.

A correlation was found, however, between mother's mental disorder (.33) as measured by her self-report of symptoms, and the child's total impairment rating at Time 1. The relationship between mental hospitalization and/or psychological or psychiatric treatment, on one hand, and on the other hand, the mother's symptom score, was .32. This suggested that the mother's current level of disturbance was directly (though moderately) related to her child's behavior and impairment, while her psychiatric history was not. Since the "children" in the study were now mostly in their early twenties, signs of adult psychoses could be expected to show themselves, especially in children of parents who were in mental hospitals, and thus conferred high genetic risk on their children.

It was logical to check the parental behavior of present or former patients, to see if they treated their children more harshly, or distantly. On the contrary, there was little difference in the means of parental behaviors between treated and untreated parents. Patients don't necessarily make "bad" parents.

A final analysis, using both the cross-section and Welfare samples separately, showed an interaction effect of parenting behavior on children, depending upon parent patient status. This analysis used a number of parental practice variables, and a summary statement here does some injustice to the results. In general, however, it is striking that the relationship of parenting variables (coldness, punitiveness, lability, restrictiveness and supportive-directing) to child Total Impairment (Time 1 and Time 2) averages .20 for non-patient parents, .40 among out-patient parents, and between .60 and .70 for in-patient parents (past or present treatment for all patients). Conjectures as to why poor parenting among patients is more damaging, given equal levels of mean severity, can range from fluctuations in range of parenting practices, guilt engendered in the child by the hospitalization, the effects of separation from the parent, to the loss of parental authority due to the stigma of treatment, etc. Further work is called for.

Children of Broken Homes

A 3-way analysis of variance was conducted of the effects of broken homes (as indexed by a father figure variable coded "has natural father, has no father, has surrogate father") with ethnic background, and socioeconomic status as factors, with 18 child behavior dimensions as criterion variables, at both Time 1 and Time 2 (16 variables). Main effects for the father figure variable were found for Delinquency, Noncompulsivity (one end of this variable was compulsivity, the other end involved petty stealing, slovenliness, and failure to come home on time, among other behaviors) and Undemandingness. These three dimensions belonged to a larger domain labeled Delinquency, Super-Factor III which involved older children, peer rather than parent orientation, and general anti-social behavior. There was also a main effect at Time 1 for Self-Destructive Tendencies, but it was associated with the presence of a father surrogate only, not with father absence. This suggests depressive tendencies not due to object loss, but rather to clashes with a surrogate. Evidence of this was found in previous studies of adults, but only among low-SES males (Langner & Michael, 1963).

There were no significant main effects of ethnic background, but several interactions between father-figure and race. For example, with Total Impairment as the criterion variable, children living with natural fathers were less impaired than those without fathers or with surrogates. However, among Hispanic children, those with surrogate fathers were more impaired than those living with mothers alone.

This confirms the strength of father absence and surrogate fathers as conditions for the development of antisocial behavior and some closely related behavior. It also suggests depressions is associated with father replacement (surrogate), but not with father loss or presence of natural father.

Treated Prevalence, True Prevalence and Incidence

Only about half the severely impaired children in the sample, (or 6.7 percent of the whole sample) were ever referred to any type of mental health service. Only one in five got "long-term" treatment of 6 months or longer. Of all children referred for help, regardless of impairment level, only 35 percent got long-term treatment, which mirrors the national figure. Unfortunately, treated children with comparable levels of anxiety-depression, organic-developmental problems, and delinquency showed no improvement in 16 behavior dimensions over those untreated. Reflecting a British study, (Shepherd et al., 1971) the Manhattan children with long-term treatment showed worse average current psychopathology on many of the behaviors than the short-term treatment group. Perhaps up to five visits help the family to mobilize itself.

About 4500 children per year aged 6 to 18 may become new seriously impaired "cases" in New York City, (an incidence rate of 2.8/1000), using these sample estimates.

The effects of spontaneous change and intervention (by treatment) were assessed. The measure of spontaneous change was based on the stability or change of the child's family. Families were classified as being either better, stable, or worse environments for the child (between Time 1 and Time 2). The intervention variable was simply whether or not a child received professional treatment between Time 1 and Time 2. All cases in which treatment began prior to Time 1 were excluded from the analysis so as not to contaminate the intervention variable. Interestingly, the analyses performed clearly implicate spontaneous effects over treatment effects. In other words, the spontaneous changes in the child's environment (i.e., the child's family) were much more closely related to the changes in the overall impairment (TIR) of the child at Time 2, than were the effects of professional treatment. The correlation between family change and change in TIR was 0.19 (p .001), while for intervention the correlation was nowhere near significance (r = -0.2). The data were broken down even further to determine if family change interacted with treatment effects. In this analysis, it was found that those children who were never in treatment still showed the relationship between family change and change in TIR (r = 0.21, p .001). However, the children in treatment between Time 1 and Time 2 showed a much smaller, and statistically insignificant effect (r = 0.14, n.s.). This was interpreted as treatment (and particularly long-term treatment) acting as a drag on the improvement of the child. Simply stated, spontaneous change is more powerful without treatment than with it.

There is a possibility that Time 1 TIR is not a sufficient control for the severity of the child's original disorder. If so, slippage in measurement could have allowed for more severely impaired children in the treated (and long-term treated) group. However, controls for initial impairment and type of disorder (above) produced similar results.

It should be stressed that these results involve any and all kinds of treatment, whether voluntary or involuntary, of good or poor quality. Many children are referred by the courts and schools, with treatment the alternative to punishment or imprisonment. The closed settings are often the equivalent of jails, with minimal treatment facilities or staff. It may seem unfair to represent the average results of treatment and referral as practiced in a metropolis, when some practitioners are making great efforts to improve therapeutic techniques. These effective services are obviously rarer than we would wish, and are somewhat more available to wealthier families.

Our impression is one of good consonance and reasonable fit between diagnosis and psychometric type. For example, Oppositional, Mistrustful, and Impulse-ridden Diagnoses were all represented well beyond change in the Aggressive psychometric type, and Compulsive and Inhibited children (GAP) were quite similar to the Dependent and Sociable Child Types. In all, over half of the children in the most populated diagnostic categories were accounted for (i.e., fell into one of the two closest child types) by the psychometric typology.

Outcomes of Welfare Family Types, Welfare Child Types, and Welfare Findings

Five types of Welfare families were constructed, using hierarchical cluster analysis. This clustering was done within the Welfare sample of 1000 ADC (Aid to Dependent Children) families, and not matched to cross-section types (as was done in the case of Welfare child types). The tables of these types give only a superficial description of the demographic and intrafamilial variables which went to make up the profiles. A pathological rank (1 being the most disturbed family), based upon the means of the variables in the profile (8 parental factors, 5 parent-child factors, and 7 demographic variables) was assigned to each type. These were rank 1. Discordant homes, emotionally ill mother, 2. Intact atypical Welfare, 3. Spanish-speaking, traditional, 4.5 Educated, well-adjusted mothers, not Spanish, and 4.5 Disrupted homes, mother surrogate present. Only an impression of the range of associations and outcomes can be given. The proportion who got off Welfare after Time 1 varied between Family Types from 17.5 percent to 40 percent. The proportion whose study child had a police record ranged from 26.7 percent to 51.4 percent. Parents' police records ranged from 13.3 percent to 51.4 percent. The proportion of children held over in school after Time 1 ranged between 17 percent and 28 percent. School dropouts ranged from 8.7 percent to 19 percent. Thus, Welfare Family Types discriminated well on many outcomes.

Job acquisition was the primary reason given for getting off ADC, and involved 9 percent of the follow-up sample of 664 cases at Time 2. Remarriage, children growing up, and mother's further education only accounted for about 1 percent each as reasons for getting off Welfare completely. Only 4 percent gave other reasons.

In the random sample, 12 percent of the families were on ADC. These cross-section ADC children were disproportionately classified into the Delinquent-Aggressive and Severe Backward Isolate profile types. No Self-Destructive and very few Sociable and Competitive-Independent types were found in the ADC group within the cross-section sample. This result conflicts with later profile matching between the cross-section and ADC samples.

16

The Measurement of Antisocial Behavior and its Implications for Theory Construction

Analysis was conducted with the aim of refining the measurement of delinquency and clarifying the constructs tapped by different operational measures. At issue were the convergent and divergent aspects of survey reported and officially recorded delinquency. Overall, Survey-Reported Delinquent Behavior and Official Delinquency are positively but moderately related ($r = .36$; $.41$ for mother and self-report). There is a high degree of concordance at the low end of survey reported delinquent behavior (over 90 percent of survey "non-delinquents" are similarly identified by the police). There is, however, a large proportion of those with moderate and high survey delinquency who are not identified by the police (85 to 67%). Analysis using information preceding police identification indicates that future official delinquents can be significantly discriminated from those with only survey reported delinquent behavior. The function which discriminates these groups is heavily determined by Income Level, Parental Criminal Record, Sex, and the Mother's Attitude to Responsibility and strongly suggest a distinct course of development prior to police identification.

As an additional test a model was applied to the data to determine the extent to which these two measures can be used as alternative indicators in a causal model. The results further demonstrated that they do not provide equivalent operationalizations of the same construct. Indeed, the use of official record information to represent delinquent behavior will grossly distort the importance of certain factors while ignoring others. On the other hand, delinquent behavior is not very telling with regard to delinquent status.

The Effect of Social Factors on Development Patterns of Behavioral Disturbance

The extent to which environmental factors critically shape or modify patterns of stability in behavioral traits was examined. Specifically, the effect of societal recognition (i.e., arrest, referral to treatment, and interruption in school career) and of the environmental context on pattern of stability in delinquency, anxiety, and mentation problems was examined. In general, it was found that the extent to which a disturbance is normative (e.g., delinquency in the Welfare vs. cross-sectional samples) reinforces the stability of problematic behavior patterns and postpones the consolidation of conforming behavior. In addition, across the types of behaviors, societal recognition prior to the age of consolidation increased the stability of pathology. The results strongly support the importance of the environment in modifying developmental patterns.

17

Sex Differences in Antecedents of Delinquency

Three measures of antisocial behavior were constructed using, respectively, police-recorded information, mother-reported information on "chargeable" behaviors, and delinquency in a congruence sample for which both sources indicated non/moderate/highly delinquent behavior. Time 1 mother-reported and school information account for a significant portion of the variance in later antisocial behavior as measured by each of the criterion variables, but the prediction is improved when greater validity is achieved in measuring the criterion. In the congruence sample, for which a dual criterion was used, 59 percent of the variance in male delinquency and 63 percent in female delinquency were accounted for. For females, a relatively uniform etiology was identified across the criterion measures indicating the unique importance of parental rejection, conflict with parents, early fighting, delusions-hallucinations and self-destructive behavior. The pattern of antecedents indicates early severe emotional pathology. For males, two relatively distinct patterns appear to be involved. One emphasizes the importance of antisocial parents, early mentation problems and unanxious non-compulsive behavior. The second identifies the importance of residential mobility, conflict with parents, early fighting, self-destructive tendencies and early behavioral problems at school. The former pattern suggests the impact of socialized (or subcultural) aggression, while the latter identifies early social-emotional adjustment problems as having major etiological importance. (25, 37)

Summary

In a summary of methods and findings of one study, there is hardly room to discuss the agreements and disagreements with other studies. However, there has been a tendency for our findings to corroborate previous findings or to expand upon them in more detail. It is not possible to list even all the major studies, both longitudinal and single surveys of child and adolescent behavior, but some are given here for comparison. The type of study or data source is listed, and examples of studies are given.

LONGITUDINAL STUDIES: Macfarlane, Allen, and Honzik (1954); Kellam, Branch, Agrawal, and Ensminger (1975); Richman, Stevenson and Graham (1975); Robins, L.N., (1966); Graham and Rutter (1973); Jessor and Jessor (1977).

MOTHER OR PARENT INTERVIEW: Lapouse and Monk (1958); Miller, Hampe, Barrett, and Noble (1971); National Health Survey (1971); Shepherd, Oppenheim, and Mitchell (1971); Tuddenham, Brooks, and Milkovich (1974); Achenbach and Edelbrock (1979).

PARENT, TEACHER AND/OR CHILD INTERVIEW: Rutter, Tizard and Whitmore (1970); Rutter, Cox, Tupling, Berger and Yule (1975).

TEACHER INTERVIEW: Werry and Quay (1971).

INTERVIEW AND EXAM: Richardson and Higgins (1964).

The FRP publications have extensive references to the findings of other studies, so we will not discuss them in detail here. A few examples of differences in our findings will suffice. There does not seem to be much of a synergistic effect of various factors affecting child behavior. Most of what we find is linear and additive, not interactive. This may not seem important, but it questions the basic assumption that the sequence or specificity of events and processes impinging on the child is critical to the prediction of his or her behavior.

The FRP generally found that the particular behavior or criterion must be specified, and that generalizations about predictors of behavior can be very misleading unless one talks about isolation, antisocial behavior, depression, school failure, or anxiety, etc. Three large groups of behaviors were distinguished; Anxiety-Depression, Organic-Developmental, and Delinquency. The social, family, and physical health variables associated with, or forecasting, these behaviors varied widely.

Multiple ways of looking at children and classifying them seemed a great advantage, given the parlous state of child classification at this time. Diagnoses, profiles, dimensions and symptoms all helped in understanding the interrelationships of dependent variables. Where their predictors overlapped, there was some confirmation that a meaningful association really existed. Similarly, multiple classifications of the environment, into factors (family and social, marital and parental) and into family profile types, helped to clarify the relationships. Data reduction was crucial.

The splitting of parental rejection into factors of coldness and excitability brought great dividends. There is a great contrast with some other studies in associated outcomes. The stability of anxiety and other similar dimensions suggests that antisocial behavior is not the only stable phenomenon, and that some researchers have been misled by the fact that the age range of their samples was truncated, which made some behaviors appear unstable.

The quantification of contributions to behavior is crucial to an understanding of the data. There is a great amount of redundancy in the environmental factors believed to influence child behavior. Multiple regression and discriminant function analysis helped to sort out what was contributing, and what was freeloading. The use of large item pools (654 behavior items, several hundred parenting and marital items) allowed for factoring within domains. This increased the

accounting strength of the study considerably. It also allowed for generally high reliability.

The use of multiple sources and multiple methods, as well as multiple traits or behavior dimensions, was in part a boon, since we could compare sources. When there was agreement, it helped, but when sources disagreed, we were often unable to understand why. As others have found, many behaviors seem to be setting-specific, so that informants at school, for example, do not agree with parents (at home) because the same behavior is actually not exhibited by the child in those two settings.

The original motivation for the study had been to circumvent the possibility that adults in the Midtown study (Langner & Michael, 1963) had reported greater childhood stresses because of their adult mental disorders. This circularity was circumvented by studying children during the time they were growing up. Our mother and child reports allowed us to rank stresses for children, and these ranks can be compared with the strength of retrospective reports of childhood stresses made by Midtown adults (in relation to their adult mental health ratings).

The Midtown adults reported on their childhood experiences, which can be estimated to have occurred some 30-40 years previous to the report, since the mean age of respondents was 40.

There is reasonable agreement between the strength of retrospective and prospective predictors on four of the six factors which are comparable. This gives us some confidence that adults are not distorting their childhood estimates of parental behavior, poverty or illness. The parallels between the studies are only approximate, and the scales used in the FRP are often quite long and highly reliable, whereas the Midtown scales are short.

The strongest predictors which were not tapped, or were not analyzed, in the Midtown data, but were in the FRP, are Being Spanish-speaking (rank 1); Being Black (rank 3); (almost no minority members were in the Midtown sample), Parents Traditional-Restrictive (rank 5). Number of children (7); High rent (8); and Number of addresses (10) are listed under socioeconomic factors, but they are really unique in their contributions to behaviors, even though they are correlated with socioeconomic status. Strictly speaking, they should not be compared directly with childhood economic deprivation (for example, there are quite a few rich families with numerous children).

The advantages of prospective studies are obvious, but it seems that the "distortions" of some retrospective studies may have been exaggerated.

Project Publications and Reports

Books

Langner, T.S., Gersten, J.C., Eisenberg, J.G., Greene, E.L. and Herson, J.H. Children Under Stress: Family and Social Factors in the Behavior of Urban Children and Adolescents. Columbia University Press, (in press).

Articles

Langner, T.S., Greene, E.L., Herson, J.H., James, J.D., Goff, J.A., Rostowski, J. and Zykorie, D. Psychiatric impairment in Welfare and non-Welfare children, Welfare in Review, Vol. 7, No. 2, 1969, pp. 10-21.

Langner, T.S., Herson, J.H., Greene, E.L., Jameson, J.D. and Goff, J.A. Children of the City: Affluence, poverty and mental health, in Psychological Factors in Poverty. Vernon L. Allen (Ed.), Chicago: Markham Publishing Co., 1970, pp. 185-209.

Greene, E.L., Langner, T.S., Herson, J.H., Jameson, J.D., Eisenberg, J.G., McCarthy, E.D. Some methods of evaluating behavioral variations in children six to eighteen, Journal of the American Academy of Child Psychiatry, Vol. XII, No. 3, July, 1975, pp. 531-553.

Gersten, J.C., Langner, T.S., Eisenberg, J.G., Orzeck, L. Child behavior and life events: Undesirable change or change per se? In Stressful Life Events: Their Nature and Effects, B.S. Dohrenwend

Langner, T.S., Gersten, J.C., Greene, E.L., Eisenberg, J.G., Herson J.H., McCarthy, E.D. Treatment of psychological disorders among urban children. Journal of Consulting and Clinical Psychology, 1974, 42 (2), pp. 170-179.

Langner, T.S., Gersten, J.C., Eisenberg, J.G. Approaches to Measurement and definition in the Epidemiology of Behavior Disorders: Ethnic Background and Child Behavior, International Journal of Health Services, 1974, 4(3), 483-501.

Gersten, J.C., Langner, T.S., Eisenberg, J.G. and Simcha-Fagan, O. Spontaneous recovery and incidence of psychological disorder in urban children and adolescents. Psychiatry Digest, August, 1975, pp. 35-43.

Eisenberg, J.G., Langner, T.S., and Gersten, J.C. Differences in the behavior of Welfare and non-Welfare children in relation to parental characteristics. Journal of Community Psychology, Monograph Supplement, October, 1975, (#48) 33 pp.

Gersten, J.C., Langner, T.S., Eisenberg, J.G., Simcha-Fagan, O. and McCarthy, E.D. Stability and change in types of behavioral disturbance of children and adolescents, Journal of Abnormal Child Psychology, 1976, 4 (2), pp. 111-127.

McCarthy, E.D., Langner, T.S., Gersten, J.C., Eisenberg, J.G., and Orzeck, L. The effects of television on children and adolescents: Violence and behavior disorders. Journal of Communication, 1975 Vol. 25, No. 4, pp. 71-85.

Langner, T.S., Gersten, J.C., McCarthy, E.D., Eisenberg, J.C., Green, E.L., Herson, J.H. and Jameson, J.D. A Screening Inventory for Assessing Psychiatric Impairment in Children Six to Eighteen. Journal of Consulting and Clinical Psychology, 1976, 44, (2), pp. 286-296.

Langner, T.S. Life stress, behavior and intervening variables. Proceedings of the A.A.A.S. 1975, pp. 83-91.

Gersten, J.C., Friis, R., and Langner, T.S. Life dissatisfactions, job dissatisfaction and illness of married men over time. American Journal of Epidemiology, 1976, 103, (3), pp. 333-341.

Eisenberg, J.G., Gersten, J.C., Langner, T.S., McCarthy, E.D., Fagan, S.O. A behavioral classification of Welfare children from survey data. American Journal of Orthopsychiatry, 1976, 46(3), pp. 447-463.

Gersten, J.C. Measures of family functioning. In G.D. Grave & I.B. Pless (Eds.), Chronic childhood illness, assessment of outcome. Washington, D.C.: Fogarty International Center Series on the Teaching of Preventive Medicine; DHEW Publication No. (NIH), 76-876, 1976, pp. 193-202.

Gersten, J.C., Langner, T.S., Eisenberg, J.G. and Simcha-Fagan, O. An evaluation of the etiologic role of stressful life-change events in psychological disorders. Journal of Health and Social Behavior, September, 1977, Vol. 18, #3, 228-244.

Langner, T.S., Gersten, J.C. and Eisenberg, J.G. The epidemiology of mental disorder in children: Implications for community psychiatry. In New Trends of Psychiatry in the Community, G. Servan (Ed.), Cambridge, Mass.: Ballinger Publishing Co., 1977, Chapter 2, pp. 69-109.

Langner, T.S., Gersten, J.C., Wills, T.A., and Simcha-Fagan, O. The relative roles of early environment and early behavior as predictors of late child behavior. In: Origins of Psychopathology: Problems in Research and Public Policies. David Ricks and Barbara Dohrenwend

Langner, T.S., Gersten, J.C. and Simcha-Fagan, O. Environmental change and behavioral change in children: Does early damage irreversibly affect later functioning? (Invited paper given at the World Psychiatric Association Section Committee on Epidemiology and Community Psychiatric Triennial Meeting, St. Louis, October 1978).

Langner, T.S., McCarthy, E.D., Gersten, J.C., Simcha-Fagan, O., Eisenberg, J.G. Factors in children's behavior and mental health over time: The Family Research Project, in Research in Community and Mental Health - An annual Compilation of Research, January 1979, R.G. Simmons (Ed.), Part II, 127-181.

Langner, Thomas, S., "Predictors of Child Behavior and Their Implications for Social Policy," In R.G. Simmons (21).

McCarthy, Elizabeth, D. "Family Structure: Its Relation to Social Class and Child Behavior." In R.G. Simmons (21).

Gersten, Joanne, C. "The Role of Historical Change in the Development of Types of Behavioral Disturbance," In R.G. Simmons (21).

Simcha-Fagan, Ora. "The Prediction of Delinquent Behavior Over Time: Sex Specific Patterns Related to Official and Survey-Reported Delinquent Behavior." In R. G. Simmons (21).

Eisenberg, Jeanne, G. "The Welfare Children: An Overview of Longitudinal Findings." In R.G. Simmons (21).

Wills, T.A. and Langner, T.S. Psychosocial stress. In I.L. Kutash & L.B. Schlesinger (Eds.), Pressure Point: Perspectives on Stress and Anxiety, San Francisco, Jossey-Bass, in press.

Simcha-Fagan, O., Gersten, J.C., Langner, T.S. and Eisenberg, J.G. Antecedents and correlates of illicit drug use in a sample of urban youth. To be revised for the Journal of Consulting and Clinical Psychology.

Life Events, Role Conditions and Adaptation:

A Panel Study of Social Stress

Elizabeth G. Menaghan

Department of Sociology

The Ohio State University

Leonard I. Pearlin

Laboratory of Socio-Environmental Studies

National Institutes of Health

Morton A. Lieberman

Committee on Human Development

The University of Chicago

Joseph T. Mullan

Department of Family Relations and Human Development

The Ohio State University

**Life Events, Role Conditions and Adaptation:
A Panel Study of Social Stress**

Elizabeth G. Menaghan

Department of Sociology

The Ohio State University

Leonard I. Pearlin

Laboratory of Socio-Environmental Studies

National Institutes of Health

Morton A. Lieberman

Committee on Human Development

The University of Chicago

Joseph T. Mullan

Department of Family Relations and Human Development

The Ohio State University

Introduction

The research project discussed here addresses a core issue in adult development and aging: the effects of life events and life conditions on psychological functioning. Previous analyses, our own as well as others, have documented the association between social stresses and various dimensions of emotional distress and psychological functioning. Our goal

has been to move beyond such demonstrations toward an understanding of how these associations come about, and an explanation of the wide variation in outcomes among people who share similar social statuses, experience similar events, or face similar role problems.

In attempting to achieve this goal, we have focused on the life problems, social contexts, and coping responses of ordinary people engaged in the mainstreams of social life. The body of data we have developed has several important advantages for the understanding of social stress and adaptation. First, the study population is large, representative of urban, Chicago-area adults, and widely varied in age, social class, and event experience. Second, the data are comprehensive, including information about conditions and events in each major adult role: occupational, economic, marital, and parental. Finally, we have longitudinal as well as cross-sectional data. As many researchers have pointed out, it is often impossible with cross-sectional data to disentangle prior conditions from subsequent effects. In contrast, the panel design we have used gives us an unambiguous baseline from which to assess stability and change over time.

Two waves of interviews have not been completed, and analysis is continuing on those data. It is possible that a third interviewing wave will be completed in 1982. This report reviews the research design of the project thus far, and details the measurement of major concepts.

Research Design

Introduction

In 1972, scheduled interviews were conducted with 2,300 people representative of the adult population of the Census-defined urbanized area of Chicago. These interviews had three main foci: the assessment of a wide range of problems and hardships people experience as workers, breadwinners, husbands and wives, and as parents; the identification of resources and responses they bring to bear in coping with these life strains; and the enumeration of symptoms indicative of emotional stress and psychological disturbance.

The follow-up survey in 1976, like the first, was also broadly concerned with the problems and challenges that converge on the lives of people and their psychological effects. To permit an examination of change, we included identical measures of role-related difficulties and dissatisfactions, coping strategies, and overall psychological functioning in both the Time One and Time Two survey. In addition, however, the scope of the 1976-77 survey was enlarged to include information about the life-cycle transitions through which people had passed and the crises they had confronted in the four-year period following the first interview. At each time point, information was also obtained about demographic

characteristics and patterns of social interaction. All interviews were conducted in the respondents' homes by trained interviewers, with the average interview lasting an hour and forty minutes.

Characteristics of the Study Population

A cluster technique was used to draw the 1972 sample, each cluster consisting of four households per block. Since the total sample of 2,300 was to be made up of clusters of four households, the interviewing was done in 575 blocks, one-fourth the size of the sample. The 1970 Census (U.S. Bureau of the Census, 1972) reports that there are 2,137,185 households in the Chicago Urbanized Area; this total was divided by the total number of blocks in which households were to be chosen (575), and the results, 3,719, was the skip factor for the selection of households. After each 3,716th household was selected, three additional households from the same block were chosen by dividing the total number of households on the block by four and using the result as the factor for counting from the initially selected address. In anticipation of refusals--30% of those contacted--and to make allowance for households where contact could not be established within three callbacks, substitute addresses in each block were also prelisted. The sex of the person to be interviewed in each household was predesignated in order to have as equal a number of males and females as possible. Finally, only those between 18 and 65 were included, producing a sample still actively engaged in occupational life.

At the time of the 1972 interview, subjects were asked if they would be willing to be reinterviewed in the future. Eighty-eight percent, approximately 2,000 of the original sample, agreed to this. In 1976-77, a sub-sample of over 1,100 people was drawn whose social characteristics differ only slightly from the original. Specifically, the sub-sample contains 5% fewer nonwhites than the 1972 sample, 4% fewer nonmarried, 6% fewer whose education is less than high school, and 3% fewer unskilled workers. There is also some difference in the economic makeup of the sub-sample. In 1972, 29% of our sample had annual family incomes of \$8,000 or less, but only 21% of the sub-sample are drawn from that category, with a corresponding increase at the higher levels. Finally, there was a slight loss of the youngest group of respondents; whereas 15% of the original sample should presently be in their twenties, 11% of the sub-sample are actually in that age range. The current sub-sample, therefore, is slightly older than it would be if it were perfectly representative of the original sample, somewhat more highly placed in socioeconomic status, and a bit more likely to be married; and it contains proportionately fewer nonwhites. To a small extent, it mirrors less accurately the population characteristics of the Chicago Urbanized Area, but its discrepancies should not distort the kinds of relationships we shall be examining.

Measures of Psychological Functioning

Our measures of psychological functioning included two types: those capturing the person's own view of his or her well-being and potency, and those assessing the frequency of symptoms of distress and breakdown. In the former category are a seven-item scale of personal mastery developed by Pearlin and described in detail in Pearlin and Schooler (1978), and a ten-item scale of individual self-esteem developed by Rosenberg (1965). In the latter are twelve-item scales of depression and anxiety (Derogatis, Lipman, Rickles, Uhlenhuth, & Covi, 1974), assessments of physical health problems, and scales of felt distress in each of the person's major life roles.

Conditions in Major Life Roles

We also developed several specific scales in each role area to tap the quality of the person's life experience. In occupational life, we assessed:

- (1) the noxiousness of the physical work environment (e.g., noise, dirt, dust, and danger);
- (2) the degree of overload and pressure in meeting job demands;
- (3) the amount of extrinsic rewards the job holds (e.g., opportunities for salary increases and promotion);
- (4) the extent of depersonalization in the work environment.

In parental life, we assessed:

- (1) the frequency with which children violated parental standards for behavior and performance in the larger environment;
- (2) the degree of threat to parental aspirations and values embodied in their children's behavior;
- (3) the extent to which parents felt their own parental self-image to be violated or disregarded.

In economic life, we measured the degree of struggle to obtain and maintain an adequate family life-style (e.g., adequate food, clothing and medical care and sufficient space and furnishings).

In their marital relationships, we assessed people's:

- (1) sense of restriction of their own identity;
- (2) absence of desired affection and expressivity;

- (3) denial of equity or reciprocity in daily interaction;

Thus, in each area we identified concrete conditions likely to bring distress and frustration to the respondents.

Coping Strategies

In each major adult role, we have asked about three kinds of coping strategies: those aimed at preventing or eliminating the problem directly, those functioning to deflect or buffer the stressful impact of problems by modifying the meaning of the problem (e.g., through cognitive reinterpretation, focusing on counter-balancing virtues of the situation, or finding some worse possibility with which to compare it), and those aimed at either controlling or expressing the unpleasant feelings that may be aroused by the problem.

Strategies to modify the situation directly were assessed in each role area. They include:

- (1) attempts at negotiation and compromise in marriage;
- (2) discipline and scolding of children;
- (3) careful budgeting and purchasing in economic life;
- (4) group action and discussions with supervisors and co-workers in occupational life.

Strategies to modify the meaning of difficult conditions are primarily of two types:

- (1) optimistic comparisons of one's situation relative to the past and relative to one's peers;
- (2) selective inattention to unpleasant aspects and close attention to positive features of the situation.

Each of these was measured in all four role areas.

Strategies to manage or control the distress aroused by role problems were:

- (1) a conscious suppression of feelings and withdrawal from interaction (measured in the marital role);
- (2) a restriction of expectations and denial of importance of a role (measured in the occupational role);

- (3) a restriction of responsibility for conditions in a role (measured in the parental role).

Demographic Indicators

At each time point, information was obtained about the respondent's age, sex, race, employment status and marital history; the age, sex, and current residence of any children; the current or most recent occupation of the respondent; the highest level of education completed; and the total current family income. Married respondents were also asked about the employment status, occupation, and educational attainment of their spouses, and all respondents were asked about the occupation of the family breadwinner of their family of origin.

The Occurrence of Life Events

The Time Two interview ascertained whether any of a series of possible transitions or life events had occurred in the time since the first interview.

In occupational life, the following events were queried about: labor force entry or reentry, labor force exit (through retirement, job loss, or a decision to become a homemaker), promotion, demotion, temporary job disruption, and change in employer.

In marriage, the focus was on entry through first marriage or remarriage, and exit through separation, divorce, or widowhood.

Possible changes in the parental role included children's births, their starting school, becoming teenagers, and leaving home. For each, we distinguished events occurring to oldest, middle, and youngest children. Other possible events connected with family life were the serious deterioration or death of one's parents or parents-in-law; serious illness of self, spouse, or child; death of a sibling, friend, or other relative; and change in residence.

If an event had occurred, respondents were asked to appraise its impact. Specifically, the study asked: (1) whether there was anything about the transition that bothered them, and if so, how bothered they were; (2) whether they were pleased or excited about the transition, and if so, how much; (3) whether the transition changed their lives in some ways, and if so, how much; and (4) whether the transition changed the way they felt about themselves, and if so, in what ways (more or less: responsible, in control of their lives, distant from other people, dependent, disappointed in life, wise, capable, and self-respecting).

Measures of Social Interaction

At each time point, a variety of questions was asked about the availability of possible social supports and the frequency of contact with them. Length of time in the same neighborhood, attendance at religious services, and membership in clubs and organizations were also ascertained.

The Time Two interview enlarged upon this basic set of information in two ways. First, there was more detailed questioning about the intimacy and dependability of the social network. (Brown [1978] discusses these measures in greater detail.) Second, a standard sequence of help-seeking questions was asked whenever an event was appraised as troubling or when difficult conditions in a particular role were felt to be distressing. Respondents who said they had talked to someone about the problem were asked to whom they had spoken, what the persons approached did, and whether they felt any better afterward. Helpers were categorized as family members (one's spouse, one's own parents, or one's children), informal helpers (neighbors, friends, co-workers, or other relatives), and formal helpers (doctors, nurses, teachers, clergy, police, lawyers, politicians, or therapists). The kinds of help received included: talking (listened, asked questions, suggested someone else to talk to), insight (showed a new way to look at things), and action (took some action, suggested what action to take, or took respondent to see someone who could take action).

Respondents who did not seek any help were asked about possible reasons for refraining from helpseeking: a sense that they could handle the problems by themselves; the feeling that no one would be willing, able, or available to help; and/or a belief that seeking help would be too costly, in time, effort, or amount of self-disclosure needed. Thus, both general descriptions of social interaction and problem-specific reports of discrete help-seeking episodes were obtained.

Completed Analyses

Four major sets of analyses have been completed. Those exploring:

- (1) relationships between social position, social stress, and adaptation;
- (2) the effectiveness of coping strategies in the management of social stress, examined cross-sectionally;
- (3) the use of social resources in managing life events and social stress; and
- (4) relationships between life events, role conditions, and psychological adaptation.

Analyses in Progress

Much of our most recent work has focused on the issue of life events and their effects. Like many life-span investigators, we believe that transitional events may help to provide a causative model for explaining the patterns and regularities in adult social and psychological functioning. In the course of our work thus far, however, it has become increasingly clear to us that life events cannot be understood apart from a detailed examination of the prior and subsequent conditions people face in their major life roles. When people confront difficult conditions in their roles, they seem to be more vulnerable to events touching on these roles. Just as important, life events themselves often importantly shape the ongoing role conditions which people subsequently endure. In fact, it seems likely that life events involving no major alterations in enduring role conditions will have few consequences for psychological adaptation.

At the same time, attention to prior and subsequent role conditions may help to explain some of the variation in outcomes among people experiencing a given event. For example, many of the major events in later adult life have been described in terms of role loss or role exit: retirement from occupational roles, the termination of marital relationships through widowhood, separation or divorce, or the termination of daily parental tasks when the final child moves away from home. At such points of role transition, the circumstances that prevailed in the roles from which people have departed are likely to influence the nature and intensity of the effects of the departure. For example, new retirees probably include both people reluctantly leaving satisfying labor as well as those impatiently escaping frustrating job conditions. The responses to the shared event of retirement of such diverse people are expected to be quite different. Of course, just as important as past role conditions in shaping the effects of transitional events are the conditions which prevail in the roles into which people newly move. We are now focusing on such role exits and entrances and examining how conditions in both the role one is leaving and the role one is entering may mediate the effects of transitions on adult psychological functioning. For example, we are comparing the newly single and newly retired with those whose marital or occupational status has not changed in the four years between interview points. Our analysis will be aimed at explicating the conditions under which negative psychological outcomes are least and most likely. Thus, it should both advance our theoretical understanding of role transitions and suggest social policies that could positively influence the impact of such events on adult lives.

Many life events, of course, involve changes in an ongoing role rather than role entrance or exit. Our data permit us to trace the impact of events occurring in one role, not only on subsequent conditions in that role, but on subsequent conditions in other roles as well. The identification of such spillover effects into distant roles should bring a more subtle and more complete understanding of the ways in which life events may eventuate in a wide range of outcomes. Our specific plans focus in particular on the set of

parental role changes indicated by changes in activity of youngest and oldest children, and examine the impact of such events on parental experience and occupational involvements, as well as on their marital relationships. Ultimately, we seek to learn how changes in occupational and family life may reciprocally affect one another.

The multi-role nature of our data also permits us to see how conditions in several roles may interact to affect psychological adaptation. For example, a rewarding marital relationship may buffer the effects of parental problems; or occupational frustrations may exacerbate the depressive consequences of marital stress. Similarly, the effect of an occupational promotion may depend on the total set of other role responsibilities the person faces: a single parent with small children may experience more distress than exhilaration from an expansion of job responsibilities. Thus, we plan to explore the joint effects of multiple life changes and role conditions on psychological outcomes--an important expansion of prior analyses.

In summary, linking life events, role conditions, and psychological functioning over time enables us to focus on the following: (1) the role of prior and subsequent conditions in mediating the effects of role transitions; (2) the spillover of life events from one role area to others; and (3) the combination of conditions in several adult roles as they jointly affect psychological functioning.

We shall also be continuing to explore the ways that people attempt to manage life events and role conditions through their own coping responses. Here we are particularly concerned with processes over time: how established coping repertoires are brought to bear on subsequent events; how the effectiveness of coping strategies may vary with the particular event or role problem toward which they are directed; and how events and their repercussions may themselves bring about alterations in people's characteristic coping strategies. In these analyses we shall be able to address a major question in the field of adult development and aging: are such coping dispositions stable properties of individuals or are they heavily influenced by the degree of stability in their environment?

Finally, threaded through all of these issues is a strong interest in the way in which life events, the conditions under which they arise, the coping strategies they prompt, and their outcomes vary with and are organized around the social and economic characteristics of people. Such social characteristics as race, marital status, gender, social class, education, and age thus are among the conditions we shall be taking into account in all of our efforts to understand the relationship between life events and psychological functioning.

Taken together, these analyses should importantly advance our understanding of the impact of life events and life conditions on adult lives. They should also help to expose major gaps in our current knowledge, and thus identify critical problems for future studies to explore.

References

- Brown, B.B. Social and psychological correlates of help-seeking behavior among urban adults. American Journal of Community Psychology, 1978, 6, 425-440.
- Cronbach, L., & Furby, L. How should we measure "change"--or should we? Psychological Bulletin, 1970, 74, 68-80.
- Derogatis, L., Lipman, R., Rickles, K., Uhlenhuth, E., & Covi, L. The Hopkins symptom checklist: A self-report symptom inventory. Behavioral Science, 1974, 19, 1-15.
- Gourash, N. Help-seeking: A review of the literature. American Journal of Community Psychology, 1978, 6, 413-424.
- Harris, C.W. (Ed.) Problems in measuring change. Madison, Wis.: University of Wisconsin Press, 1963.
- Joreskog, K.G. A general method for analysis of covariance structures. Biometrika, 1970, 57, 239-251.
- Joreskog, K.G., & Sorbom, D. Statistical models and methods for analysis of longitudinal data, in D.J. Aigner and A.S. Goldberger (Eds.) Latent variables in socioeconomic models. North Holland Publishing, 1976.
- Lieberman, G.L. Children of the elderly as natural helpers: Some demographic differences. American Journal of Community Psychology, 1978, 6, 489-498.
- Lieberman, M.A., & Mullan, J.T. Does help help? The adaptive consequences of obtaining help from professionals and social networks. American Journal of Community Psychology, 1978, 6, 499-517.
- Menaghan, E.G. Seeking help for parental concerns in the middle years. American Journal of Community Psychology, 1978, 6, 477-488.
- Menaghan, E.G. Assessing the impact of family transitions on marital experience: Problems and prospects, in H. McCubbin (Ed.) Family stress, coping, and social support. New York: Springer, in press, 1980.
- Pearlin, L.I. Sex roles and depression, in N. Datan and L. Ginsberg (Eds.) Lifespan developmental psychology: Normative life crises. New York: Academic Press, 1975.
- Pearlin, L.I. Status inequality and stress in marriage. American Sociological Review, 1975, 40, 344-357.
- Pearlin, L.I., & Johnson, J.S. Marital status, life-strains and depression. American Sociological Review, 1977, 42, 704-715.
- Pearlin, L.I., & Lieberman, M.A. Social sources of emotional distress, in Roberta Simmons (Ed.) Research in community and mental health. Greenwich, Connecticut: JAI Press, in press, 1979.
- Pearlin, L.I., & Radabaugh, C.W. Economic strains and the coping functions of alcohol. American Journal of Sociology, 1976, 82, 652-663.
- Pearlin, L.I., & Radabaugh, C.W. The sociological study of a social problem: A reply to Roman. American Journal of Sociology, 1977, 83, 991-994.
- Pearlin, L.I., & Schooler, C. The structure of coping. Journal of Health and Social Behavior, 1978, 19, 2-21.
- Rosenberg, M. Society and the adolescent self-image. Princeton, N.J.: Princeton University Press, 1965.
- Rosenberg, M., & Pearlin, L.I. Social class and self-esteem among children and adults. American Journal of Sociology, 1978, 84, 53-77.
- Schreiber, S.T., & Glidewell, J.C. Social norms and helping in a community of limited liability. American Journal of Community Psychology, 1978, 6, 441-453.
- U.S. Bureau of the Census. Census of Housing: 1970 Block Statistics; Financial Report HC (3)-68, Chicago, Illinois-Northwestern Indiana Urbanized Area. Washington D.C.: U.S. Government Printing Office, 1972.
- Werts, C.E., & Linn, R.I. A general linear model for studying growth. Psychological Bulletin, 1970, 73, 17-22.
- Wiley, D.E., & Harnischfeger, A. Post hoc, ergo propter hoc: Problems in the attribution of change. Report No. 7. Chicago: Studies of Educative Processes, University of Chicago, 1973.

CONTINUED

5 OF 10

The New York Longitudinal Study (NYLS)

Alexander Thomas, M.D.

Professor of Psychiatry

New York University Medical Center

Stella Chess, M.D.

Professor of Child Psychiatry

New York University Medical Center

Objectives

This longitudinal study was initiated in 1956, and active data collection and analyses are still in progress.

The major overall objective of the NYLS has been the systematic delineation of individual differences in the behavioral characteristics of very young infants, the study of the vicissitudes of such characteristics over time, and the determination of their functional significance for normal and deviant psychological development. The nature of the longitudinal data has also permitted a number of other types of analyses.

The NYLS was launched at a time when psychological theory in this country was dominated by an almost exclusive emphasis on the intra- and extrafamilial environmental determinants of behavioral development. This trend was not limited to any one school of thought. It was manifest both in those inquiries basing themselves on psychoanalytic conceptualizations and in those deriving from various positions in academic psychology. The role of organismic characteristics received some attention, mainly through the view that developmental level was a fundamental factor in structuring the child's reactions to the environment. However, the concept of developmental level in the main referred to general laws of responsiveness and to the time sequences in which universals in personality organization

were achieved, but not to the issue of individuality or uniqueness of functioning. A number of reports had appeared in the preceding 25 years, reporting significant individual differences in the behavioral characteristics of infants, or suggesting that intrinsic differences in "ego tendencies" or central nervous system organization might have important influences on psychological functioning. However, such findings or speculations were not supported by any systematic, long-term investigations on the relationship between individual differences in early childhood and later course of behavioral development (see Thomas, Chess, & Birch, 1968, p. 5 for a brief review of this literature).

The emphasis on environmental influences did, of course, produce extremely important evidence from many directions on the effect of differences in environmental stimuli and life experiences on a child's physical, behavioral, emotional, and cognitive development. These studies also provided a necessary correction of the static mechanical constitutionalist view that development represented the mere unfolding and elaboration of fixed characteristics already present in the newborn infant.

However, at that time, in the early 1950s, we found ourselves progressively dissatisfied with the dominant one-sided environmentalist views. As clinicians, we were repeatedly impressed by our inability to make a direct correlation between environmental influences, such as parental attitudes and practices, and a child's psychological development. Furthermore, even in cases where parental dysfunction was obviously responsible for a child's behavior problem, there was no consistent correlation between the parental approach and the specific pathology manifested by the child. As mental health professionals, we became increasingly concerned at the dominant professional ideology of the time, in which the causation of all child psychopathology, from simple behavior problems to juvenile delinquency to schizophrenia itself, was laid at the doorstep of the mother. As several others pointed out at the time (Bruch, 1954), the guilt and anxiety created in mothers whose children had even minor behavioral deviations were enormous. Like innumerable other parents, we were struck by the clearly evident individual differences in our children, even in the first few weeks of life. And finally, a review of the literature revealed that the available evidence actually did not support the exclusively environmentalist position (Orlansky, 1949).

These considerations led us to the hypothesis that the prevailing views in developmental theory failed to give sufficient, if any, attention to the functionally significant role of individuality in behavior patterns of the very young infant. To test this hypothesis, a systematic antero-spective longitudinal study of a good-sized sample was required. This led to the initiation of the NYLS in 1956.

Sample

The sample was gathered through our direct or indirect contact with mothers who were willing to cooperate in a long-term study of normal child development during pregnancy or immediately after birth.

The sample was unselected, except that all resided in New York City or one of its suburbs, and that the method of contact ensured families of middle- or upper-middle-class background. These characteristics of the families made it relatively homogeneous and reduced, to some extent, social and geographical difference as a source for variance in behavioral development.

The cumulative collection of families was started in March 1956 and completed over a six-year period. Ninety families were enrolled, and all children born to the families during the six-year period were included in the sample. The total initial sample comprised 138 children, with 50 families contributing one child, 31 two children, 6 three children, and 2 four children to the study sample. (Previous reports incorrectly gave our initial sample number as 141, the result of a mechanical tabulation error.)

Five children from five families were lost to the study in the first few years because of long-distance changes in residence. No further sample attrition has occurred since then. The 84 families retained in the study are predominantly Jewish (78%), with some Protestant (15%) and Catholic (7%) families. There is one black and one Chinese family. Almost all parents were born in the United States. Forty percent of the mothers and 60% of the fathers had both college educations and postgraduate degrees, and only 9% of the mothers and 8% of the fathers had no college at all. With only three exceptions, all of the fathers worked either in one of the professions, or in business at a management or executive level. Eighty percent of the mothers had occupations similar to the fathers, and the remaining 20% had been employed as office workers or secretaries. At the time of the birth of the first child enrolled in the longitudinal study, half of the mothers were less than 31 years old, with an age range of 20 to 41 years. The median age of the fathers was 33.6 years; the youngest was 25 years and the oldest, 54 years of age. The length of marriage at the child's birth ranged from one to nineteen years, with half of the group married for at least 5.3 years.

Expressed parental attitudes were markedly homogeneous, that is, primarily permissive and child-centered. It was considered quite acceptable and even desirable for the father to participate actively in the care of the child, though the actual degree of paternal involvement varied widely from family to family. Five percent of the mothers resumed their occupations on a full-time basis, and 32% on a part-time basis by the time the child was two years old. However, the working mothers all described the care of the children as their major responsibility. Attendance at nursery school was considered a socially desirable experience for the child

by almost all the parents, and 89% of the children were enrolled in various private nursery schools at three or four years of age.

The family pediatricians were actively consulted by most of the parents for guidance in the routine care of the baby. The group as a whole was oriented toward the desirability of obtaining psychiatric advice and, if necessary, treatment for a child with deviant behavioral functioning. Upon enrollment in the study, each family was informed that the resources of the research team were freely and immediately available to them for consultation and advice should any evidence of deviant behavior develop in the child at any age-period.

The uniformity of expressed attitudes was not duplicated in the actual child-care practices and patterns of parent-child interactions present within the families. Considerable variability existed with regard to child-care practices and made possible the study of different patterns of interaction between various types of parental functioning and different styles of temperamental organization in the children. In addition, although the parents showed uniformity in expressed child-care attitudes, they by no means revealed the same preferences for given patterns of temperament, nor showed similar tolerances for different temperamental styles in their children.

All the pregnancies and deliveries were essentially uncomplicated, except for one premature birth, with severe neonatal respiratory distress. Another child with uneventful pregnancy and delivery also had a difficult neonatal course and repeated respiratory infections during the first months of life. These two children, both boys, later showed definite neurological and behavioral evidence of brain damage. One other subject, a girl, also showed deviant development beginning in the second year and findings of brain damage, with no identifiable etiology. Outside of these three cases, none of the sample showed any serious physical deviations or disorders throughout childhood.

Data Collection

The following principles have been followed from the beginning to ensure the maximum validity, objectivity, reliability, and relevance of the behavioral data:

1. Behavioral descriptions were obtained from informants, whether parents, teachers, or other sources, in factual descriptive terms. The focus was on the details of daily activities, with a concern not only for what the subject did, but how he or she did it. Statements as to the presumed meaning of the behavior were considered unsatisfactory for primary data. When such interpretative statements were made, the interviewer pressed for an actual description. Thus, if a parent reported that "the

baby hated his bath," or that "my child loves to be with his brother," the interviewer pressed for an actual description of behavior. If a nursery school teacher stated that "this child doesn't like any change in schedule," several examples were elicited, with a detailed description of the manner in which the dislike was expressed. Similarly, a staff observer who reported that a youngster "was afraid to ask the teacher for help" was instructed to spell out in detail the incidents observed and describe the behavior interpreted as "fear."

Special attention was given to tracing the details of a child's first reaction to a new stimulus and the behavior on subsequent exposures to the same stimulus, until a consistent long-term response was established. Such stimuli might be simple, as the first bath or a new food, or they might be complex, as the move to a new home, the introduction of a new person into the household, or the first contact with nursery school. The sequence of responses to new situations, whether simple or complex, gave especially rich information on a child's temperamental characteristics.

3. Data collection emphasized descriptions of the subject's current behavior or behavior of the preceding few months. This minimized the problem of distortion of recall so prevalent with retrospective reports (Robbins, 1963; Wenar, 1963; Chess, et al., 1966).

4. The contamination of data collection by "halo effects" was minimized by using different staff members for different phases of the data collection and data rating for any specific subject.

5. Protocol revisions as required at succeeding age-stage levels were pretested on samples not included in the longitudinal study itself.

The types of data collected have included:

A. Parent interviews.

This represented the prime source of data in the first three years of life. For the goals of our study, it was essential to obtain data on a child's behavior over time, such as the details of the continuity or change in reactions to novel stimuli over days and even weeks. In addition, valuable information on behavior in specific situations often could not be scheduled to coincide with the limited time of an observer's presence in the home. It is only the parent who is in the position to observe the child's behavior continuously over time. To have an observer live in the home is impractical for any substantial sample, and introduces special methodological problems of its own. Parental reports are subject to various biases of selective and distorted observation and recall. While direct observations by an outsider may minimize these issues, this method has its own methodological problems (Horowitz, 1975, p. 470).

For our study, the need to gather information on the young child's behavior in many situations of daily living and on the sequences of

behavioral responses over time required the use of the parent as a primary data-gathering source. At older age periods, as the child's functioning outside the home increased, it became possible to utilize other sources of data besides the parent. We did do a validity check on the parental interviews in 23 of the children in the NYLS, who ranged in age at the time from two to eighteen months. Direct observations were made of the child's behavior over a two- to three-hour period within two weeks of a parent interview. The observations were made by two staff members who had no knowledge of the parent interviews. In eighteen cases, observations were made by both observers at different times. In the remaining five cases, only one observation was made. The observers' behavioral observations were scored for temperamental characteristics as were the parent interviews, and the scores compared. The direct observation scores showed agreement with the parent interviews at the .01 level of confidence, and with each other at the .05 level. Thus, the observers' behavioral protocols agreed more fully with the overall characterization of the child derived from the parent interviews than they did with one another. This should not be surprising, since immediate circumstances present at each time-limited period of episodic observation would be expected to enhance differences more than they would affect the interrelation between any single period and overall functioning over time. Our own judgment that parents can give accurate descriptive reports of their children's behavior has been confirmed by other research workers (Costello, 1975, pp. 45-46; Super & Harkness, 1978; Dunn, personal communication).

B. School observations and teacher interviews.

Data were gathered yearly on the level and style of behavioral functioning, starting in nursery school and extending through the first grade. Both direct observations and teacher interviews were utilized. The observations lasted one hour each and, whenever possible, were scheduled for the indoor free play period. The staff observer had no previous knowledge of the child's behavior, and the child did not know he or she was being observed. The teachers were interviewed during the latter part of the school year, again by a staff member who had not been involved in any other data-gathering procedures on that child. An initial teacher interview was conducted after the first week of the first school year to obtain information on the initial adaptation to the first formal school experience. All observational and interview data, as with the parent interviews, emphasized concrete descriptions of specific behaviors. The teacher interview utilized a semi-structured protocol, modeled after the parent interview protocol, and modified appropriately for the school setting.

C. Psychometric testing.

All children had standard IQ testing done at three years, utilizing the Stanford-Binet, Form L, and at six years, utilizing the WISC. Approximately 60% of the subjects were also tested at 10-14 years of age, again using the WISC.

D. Behavioral observations during standard test situation.

Direct observations were done of each child's behavior during the three-year and six-year psychometric testing. An independent observer who did not participate in the testing procedure itself made a running written account of the child's motor and verbal behavior. The narrative included the initial warm-up period, the formal testing period, and a final period of free play and clean-up.

E. Parental practices and attitudes interview.

When each child was three years old, a structured interview to elicit information regarding parental practices and attitudes was held with each mother and father separately but simultaneously. These interviews were conducted by two staff members who had had no previous contact with the parents. Immediately after the interview, each parent filled out the Schaefer and Bell Parental Attitude Research Instrument (PARI).

F. Psychiatric evaluation.

The staff interviewers reported to the senior research child psychiatrist (S.C.) any expression of concern by parent, teacher, or testing psychologist over a child's behavior. If the psychiatrist judged that the reported behavior warranted clinical study, a psychiatric evaluation was arranged and done by her. These clinical evaluations were conducted in a manner identical with a standard consultation carried out either with a private patient or with a case presenting at a child guidance clinic. A judgment was then made as to whether a behavior disorder existed. If so, a specific clinical diagnosis was made, a treatment plan formulated and discussed with the parents. Follow-up evaluations were done in each clinical case as frequently as judged desirable.

G. Other data.

Achievement scores in reading and arithmetic were obtained from the various elementary schools attended by the subjects. The Wide Range Achievement Test (WRAT) was administered to approximately 60% of the sample between 10-14 years of age. Special sensory neurological or psychological studies were undertaken when a child's history or psychiatric findings suggested the need.

Data Collection Schedules

A pilot exploration of the feasibility of data collection in the neonatal period was undertaken. The wide variability of the behavior in any one neonate from day to day and even hour to hour made the issues of data analysis complex and demanding. With our limited resources at the

beginning of the study, we decided to start data collection at two to four months of age. Parental interviews started at that age-period were conducted at three-month intervals for the first eighteen months of life, then at six-month intervals until five years, and yearly thereafter until eight or nine years. School observations and teacher interviews were done yearly from nursery school through the first grade. IQ test was done at three and six years on all subjects, and on the majority at 10-14 years. Follow-up on the children with behavior disorders was done as frequently as clinically indicated.

After eight to nine years, there was a lull in data collection, except for the follow-up of the behavior disorder cases, the clinical evaluation of additional subjects referred for psychiatric consultation, the psychometric testing of the majority of the sample between 10-14 years, and the gathering of school academic achievement records.

Another active phase of data collection occurred when the subjects were sixteen years of age. For the first time, a direct interview was done with almost all subjects, covering functioning in various areas, recollection of special past events, and subjective self-evaluation. Independent interviews with the parents during this period covered the same issues as with the subjects.

Currently, we are completing a new wave of data collection at age 20-22 years. Again, this involves separate interviews with subjects and parents. The subject interviews are extensive and intensive, covering both current and past issues, objective events and subjective feelings, attitudes and goals. A temperament questionnaire and the Offer Self-Image questionnaire are included.

Summary of Results

1. Origins of temperament. A number of studies, both from our own research unit and other centers, have contributed data which bear directly or indirectly on this issue. The findings suggest "an appreciable, but by no means exclusive, genetic role in the determination of temperamental individuality in the young infant. Prenatal or perinatal brain damage does not appear to influence temperament in any striking fashion. The data also indicate that parental attitudes and functioning, as shaped by the sex of the child or special concerns for a premature infant, at the very most have a modest etiological influence . . . sociocultural factors appear to have some influence . . . chronic anxiety preceding or at least starting in pregnancy may also be significant" (Thomas & Chess, 1977, p. 152).

2. Consistency of temperament over time. As is true for any other psychological attribute, temperament is not immutable. Quantitative analyses for the first five years showed significant inter-year correlations from one

year to the next for all categories except approach/withdrawal, distractibility and persistence. As the time span for the comparison was increased, the number of significant correlations decreased. Qualitative analysis of the behavioral data from infancy through adolescence showed five patterns: (1) clear-cut consistency; (2) consistency in some aspects of temperament at one period and in other aspects at other times; (3) distortion of the expression of temperament by other factors, such as psychodynamic patterns; (4) consistency in temperament but qualitative change in temperament-environment interaction; and (5) change in conspicuous temperamental trait. Any individual subject might show a combination of several of these five possibilities (Chess & Thomas, 1977; Thomas & Chess, 1977, Ch. 12). Currently, we are engaged in a quantitative comparison of temperament in the preschool and early adult years.

3. Temperament and parental attitudes and practices. The development and course of behavior disorder could in many cases be traced to parental unwillingness or inability to accept the child's temperament as normal (Thomas, Chess, & Birch, 1968). Success or failure of parent guidance as a therapeutic procedure also correlated with this same issue, most conspicuously with the distractible, low persistence children (Thomas, Chess, & Birch, 1968, Ch. 16).

4. Temperament and school functioning. Temperament scores at age five years for 116 of the NYLS subjects were compared with academic achievement scores through the sixth grade. Significant relationships were found between the non-adaptability and withdrawal categories and low achievement scores. There were also scattered significant correlations beyond the .05 level of confidence for activity level, sensory threshold, intensity of reaction and persistence (Chess, Thomas, & Cameron, 1976). A study of another sample of 93 children in a suburban middle-class kindergarten showed a significant distortion of the teachers' estimates of the children's intellectual levels which correlated with their judgments of the children's temperament (Gordon & Thomas, 1967). In a number of individual subjects in the NYLS, specific temperamental characteristics appeared to significantly influence the child's adaptation to the school environment, the teacher's estimate of academic ability, or both (Chess, Thomas, & Cameron, 1976).

5. Temperament and behavior disorder. Extensive quantitative and qualitative analyses of this issue have been done with the NYLS data and are reported in detail for the childhood years in our volume on this subject (Thomas, Chess, & Birch, 1968). The findings indicate that features of temperament play significant roles in the genesis and evolution of behavior disorders in childhood. We are currently engaged in a similar analysis in the early adult age-period. Of special interest has been the finding that the temperamentally Difficult Child has the strong possibility of unfavorable influence on parental attitudes and practices and the greatest vulnerability to behavior disorder development in childhood (Thomas, Chess, & Birch, 1968, Ch. 7). This vulnerability is further increased if the Difficult Child temperament is present in a mentally retarded (Chess & Korn, 1970), a physically handicapped (Thomas & Chess, 1977, Ch. 5), or a

brain-damaged child (Thomas & Chess, 1975).

6. Early experience and later development. Developmental theory has been strongly influenced by the formulations of the psychoanalytic and behaviorist movements on the deterministic influence of early childhood for later psychological functioning. Our own findings, as well as those from other major longitudinal studies, challenge this view. "Physical medicine has long since freed itself from simplistic linear developmental models. A physically healthy start in life is better than a sickly childhood. But we would never expect that the healthy six-year-old is invulnerable to future illness or that the undernourished frail youngster is doomed to a life of chronic sickness. With psychological development, the genetic, maturational, and environmental contingencies and their interplay are even more complex and subject to variation and change than in physical development" (Chess, 1979, p. 111).

7. Development in middle childhood. The middle childhood period is one of continued development and psychologic change. In most of the NYLS subjects, sexual interest and activity were evident. The psychoanalytic concept of middle childhood as a "latency period" would therefore appear to be a confusing and inappropriate designation for the six- to twelve-year age-period, and should be abandoned (Thomas & Chess, 1972).

8. Development in adolescence. The data from the behavioral course of the NYLS subjects in adolescence indicate that "no a priori conclusions can be drawn as to the significance or absence of adolescent turmoil in any specific youngster. Each case has to be evaluated individually and in the perspective of overall behavior and functioning" (Thomas & Chess, 1976, p. 542).

9. Cross-cultural comparisons. A number of comparisons have been made between the NYLS subjects of middle-class native-born families and the subjects of unskilled and semiskilled working-class Puerto Rican families (PRWC). On psychometric testing at three years, the NYLS children as a group tended to deal with the test items as task demands, while the PRWC children responded to the test demands as a social situation. This finding appeared to reflect differences in child-care practices in the parents of the two groups (Hertzog, et al., 1968). Sleep problems occurred in a number of the NYLS group in the preschool years, but not in the PRWC children. High activity level temperament, on the other hand, created little risk for behavior disorder development in the NYLS sample, but significant risk for the PRWC children. Again, this appeared to reflect differences in child-care practices, as well as the physical space available for the high activity children (Thomas, et al., 1974).

Final Comment

The opportunity to follow the youngsters in the NYLS from early infancy into early adult life has provided a wealth of data on many issues in developmental psychology, as summarized above. Overall, we have been impressed by the range and variety of the behavioral repertoire of the young infant. This behavior makes it possible for the child to play an active role in development from birth onward. Any global concept of "personality" becomes untenable as one identifies the many different behaviors in different interactions that each child and adolescent exhibits. All of us play many roles as we go back and forth from one life situation to another. All our roles are interrelated, all are parts of our individuality, and no one can be entitled "the real self" at the expense of the others, and endowed with some mystique of personality.

References

- Bruch, H. Parent education or the illusion of omnipotence. American Journal of Orthopsychiatry, 1954, 24, 723-732.
- Burks, J., & Rubenstein, M. Temperamental styles in adult interaction. New York: Brunner/Mazel, 1979.
- Cameron, J.R. Parental temperament, children's temperament, and the risk of childhood behavioral problems. American Journal of Orthopsychiatry, 1977, 47, 568-578.
- Carey, W.B. Measurement of infant temperament in pediatrics, in J. Westman (Ed.) Individual differences in children. New York: Wiley, pp. 293-306, 1973.
- Cattell, R.B. Personality: A systematic and factual study. New York: McGraw-Hill, 1950.
- Chess, S. Developmental theory revisited; findings of longitudinal study. Canadian Journal of Psychiatry, 1979, 24, 101-112.
- Chess, S., & Korn, S.J. Temperament and behavior disorder in mentally retarded children. Arch. Gen. Psychiat., 1970, 22, 122-129.
- Chess, S., & Thomas, A. Temperamental individuality from childhood to adolescence. Jour. Amer. Academy Child Psychiat., 1977, 16, 218-226.
- Chess, S., Thomas, A., & Birch, H.G. Distortions in developmental reporting made by parents of behaviorally disturbed children. Jour. Amer. Child Psychiat., 1966, 5, 226-234.
- Chess, S., Thomas, A., & Cameron, M. Temperament: Its significance for school adjustment and academic achievement. New York University Educational Review, 1976, 73, 24-28.
- Clarke, A.D.B. Predicting human development: Problems, evidence, implications. Bull. Brit. Psychol. Soc., 1978, 31, 947-961.
- Costello, A. Are mothers stimulating, in R. Lewin (Ed.) Child alive. London: Temple Smith, 1975.
- Dunn, J. Personal communication, 1975.
- Gordon, E.M., & Thomas, A. Children's behavioral style and the teacher's appraisal of their intelligence. Jour. School Psychol., 1967, 5, 292-300.
- Guilford, J.P. Personality. New York: McGraw-Hill, 1959.
- Hertzig, M.E., Birch, H.G., Thomas, A., & Mendez, O.A. Class and ethnic differences in the responsiveness of preschool children to cognitive demands. Monographs Society Research in Child Development, 1968, 33, 1-69.
- Horowitz, F.D. (Ed.) Review of child development research. Chicago: University of Chicago Press, 1975.
- McDevitt, S.C., & Carey, W.B. The measurement of temperament in 3-7 year old children. Jour. Child Psychol. and Psychiat., 1978, 19, 245-253.
- Orlansky, H. Infant care and personality. Psychol. Bull., 1949, 46, 1-48.
- Persson-Blennow, I., & McNeil, T.F. A questionnaire for measurement of temperament in six-month-old infants: Development and standardization. Jour. Child Psychol. and Psychiat., 1979, 97, 107-113.
- Robbins, L. The accuracy of parental recall of aspects of child development and child rearing practices. Jour. Abnormal Social Psychol., 1963, 66, 261-270.
- Sameroff, A.J. Early influences on development: "Fact or fancy?" Merrill-Palmer Quarterly, 1975, 20, 275-301.
- Super, S.M., & Harkness, S. The infant's niche in rural Kenya and metropolitan Boston, in Anthropological studies of infancy: Some recent directions. Symposium, meeting of American Anthropological Association, Los Angeles, 1978.
- Thomas, A., & Chess, S. Development in middle childhood. Seminars in Psychol., 1972, 4, 331-341.
- Thomas, A., & Chess, S. A longitudinal study of three brain damaged children. Arch. Gen. Psychiat., 1975, 32, 457-465.
- Thomas, A., & Chess, S. Evolution of behavior disorders into adolescence. Amer. Jour. Psychiat., 1976, 133, 539-542.
- Thomas, A., & Chess, S. Temperament and development. New York: Brunner/Mazel, 1977.
- Thomas, A., & Chess, S. The dynamics of psychological development. New York: Brunner/Mazel, 1980.
- Thomas, A., Chess, S., & Birch, H.G. Temperament and behavior disorders in children. New York: New York University Press, 1968.
- Thomas, A., Chess, S., Birch, H.G., Hertzig, M.E., & Korn, S. Behavioral individuality in early childhood. New York: New York University Press, 1963.

Thomas, A., Chess, S., Sillen, J., & Mendez, O.A. Cross-cultural study of behavior in children with special vulnerabilities to stress, in D. Ricks, A. Thomas, & M. Roff (Eds.) Life history research in psychopathology, vol. 3. Minnesota: University of Minnesota Press, pp. 53-67, 1974.

Wenar, C. The reliability of developmental histories. Psychom Med., 1963, 25, 505-509.

Longitudinal Predictive Study of
Success or Failure in Birth-Control Programs*

Stuart Oskamp and Burton Mindick
Claremont Graduate School

*We gratefully acknowledge the help of countless people in conducting this research: especially the 2000 or so young women whose contraceptive attitudes and behavior we have studied; the teams of clinic staff members, interviewers, coders, secretaries, and research assistants who contributed their skills; and the Center for Population Research of NICHD which supported the research under grant HD-08074 and contract HD-82842. More extended reports of the results may be obtained from either author. Sections of this paper are quoted from Oskamp, Mindick, Berger and Motta, 1978; and Oskamp and Mindick, 1981.

**Now at the Department of Human Development and Family Studies, Cornell University.

Longitudinal Predictive Study of
Success or Failure in Birth-Control Programs

Stuart Oskamp and Burton Mindick
Claremont Graduate School

In this era of overpopulation, inflation, and worldwide shortages of food and energy, birth control programs are crucially important-- both to the individuals involved and to the future of the world. In many nations the population explosion threatens to outrun any possible increases in food production or economic advance, thus depressing the standard of living for all. Though the United States is more fortunate, since we have recently experienced a marked drop in current fertility rates to below-replacement levels, nevertheless unplanned and unwanted pregnancies and births present a major social problem to our society. To cite one example, the number of teenage illegitimate births in the U.S. has risen dramatically since 1960 and now represents more than one-third of all children born to teenage mothers (Baldwin, 1976; Chilman, 1979). Research has shown that these children are very often unwanted and frequently become abused children, while their young mothers suffer many negative consequences in limited education and employment, poverty, welfare dependency, and marital instability. Both mothers and children often become continuing burdens to our society (David & Baldwin, 1979).

Lack of birth-control planning, or failure in carrying out birth control plans, is largely responsible for the social problems of illegitimacy, abortion, and ill-advised forced marriages. These problems could be greatly reduced if contraceptive programs were more widespread and successful. Young women who are sexually active but never use birth control methods have about a 60% probability of getting pregnant over a several-year period in contrast to about a 6% risk of pregnancy for those who use a medical contraceptive method regularly. Despite these

risks, careful surveys of U.S. teenagers show that about 30% of sexually active adolescents use no contraceptive at all, about 40% use it regularly, and over 10% use it ineffectively (Zelnic & Kanter, 1978). Thus, a crucially important question for research is: Why isn't birth control used more often and more effectively?

For over six years we have been investigating this question by studying factors affecting young women's success or failure in using birth control. Most of this research has used a longitudinal design, following each individual's experiences and behavior for a period of several years. Going beyond the usual demographic indices, we have concentrated most of our attention on cognitive, attitudinal, personality, and interpersonal factors affecting contraceptive outcomes. Reports of our findings can be found in Mindick (1978), Oskamp, Mindick, Berger, and Motta (1978), Oskamp and Mindick (1981), and a summary of methodological problems in the area is in Mindick and Oskamp (1979).

A review of past research on personality and attitudinal barriers to contraception (Oskamp & Mindick, 1981) has highlighted the following factors which are frequently related to failure in contraceptive efforts:

- (a) a view of oneself as not sexually active,
- (b) low socialization,
- (c) poor personal efficacy; passivity, and feelings of external control,
- (d) cognitive immaturity--especially, limited sexual and contraceptive knowledge,
- (e) weak planfulness and future time perspective,
- (f) low social adjustment; anxiety, and neuroticism,
- (g) poor communication about sexual and contraceptive matters,
- (h) certain specific attitudes about sex and birth control,
- (i) weak behavioral intentions regarding use of contraception,
- (j) past difficulties in contraception, and
- (k) transitions in life stages and relationships.

Study One--Method

Sample

Participants in the study were new intake patients who came to contraception clinics at four affiliated Planned Parenthood agencies in the Los Angeles suburban area between November, 1973 and July, 1974. Over

700 randomly-selected patients were approached at their initial clinic visit and asked to participate. About 10% were eliminated as research subjects for reasons such as: not really being a new patient, not completing the clinic procedures and obtaining birth control, being pregnant at intake or before starting contraception, being illiterate or unable to cope with the measuring instruments, or refusal to participate in the study. (Only 2% refused, though participation was voluntary and did not affect the patients' receipt of clinic services.) That left 645 patients in the Research Group, all of whom were interviewed and tested during the waiting periods between the various clinic intake procedures, and were carefully followed-up for the next three years. In addition, the remainder of the new intake patients constituted a Control Group (N=431) who were not tested, interviewed, nor specially followed-up except at one point during the following three-year period. Realizing the importance of male partners in birth-control programs, we attempted to obtain data from the male partners of our Research Group patients, but the number who were available and willing to participate (less than 10%) was too small to produce dependable findings.

Description of Research Group

The Research Group patients ranged in age from 14 to 50 with a mean of 21.6, and their mean education was just under 12 years. There was a substantial proportion of ethnic minorities: 18% black, 19% Latin or Spanish surnamed, and 1% Asian. The family take-home pay (as recorded by the clinic) averaged about \$4500 per year for a mean family size of two persons, but there was a wide spread of income figures. Only 25% were married, 62% single, and 13% formerly married. However, 50% had had one or more pregnancies, 19% had had at least one therapeutic abortion, and 35% had living children.

Fully 44% of these patients had not been using any birth-control method in the period before their clinic intake visit, while 37% had been using pills and 19% some other method. The method prescribed by the clinic doctors was overwhelmingly pills (83%), with 5% IUDs, 4% diaphragms, and 7% condoms and/or foam. Only 17% of the patients stated that they wanted no more children, while 63% said they would like a child or children later on.

Research Instruments

Each Research Group patient had a 15-20-minute interview with an experienced female clinic staff member who had been specially trained for this project. Following the interview, subjects received three fairly short questionnaires to answer. Spanish-speaking patients were interviewed and tested in Spanish. Several different sets of questionnaires were used during the course of the research so that a greater number of predictive instruments could be tried out. Consequently, the number of subjects varied rather widely for different measures. The instruments utilized were:

- (a) Future Events Test (FET), a measure of planfulness adapted from the work of Stein, Sarbin, and Kulik (1968);
- (b) Personal Values Abstract (PVA--Gough, 1972);
- (c) Rotter's (1966) Internal-External Locus-of-Control Scale;
- (d) Miller's (1973) Personal Style Inventory (PSI), developed to measure personality variables related to pregnancy-planning behaviors;
- (e) Miller's Sexual and Contraceptive Knowledge Questionnaire (no date);
- (f) A questionnaire on knowledge and attitudes concerning sexual and contraceptive matters, containing the most discriminating items used by Goldsmith, et al. (1972) with teenage Planned Parenthood patients;
- (g) Shipley-Hartford vocabulary scale, a brief measure of verbal intelligence (Buros, 1972, p. 321).

The interview with each patient probed many diverse areas: the patient's relationship with her parents, her general relationship with her sexual partner, her satisfaction with sexual activity, the couple's openness in discussing sexual matters, worries or fears about the practice of birth control, and the patient's behavioral intentions about continuing her contraceptive program, etc. In addition, the interviewers rated a number of the patient's characteristics and predicted her likely contraceptive success.

Behavioral indices were also investigated. These included past birth-control usage, current method of contraception, order of arrival for the initial visit to the clinic, previous pregnancies and abortions, length and frequency of sexual activity, age at first intercourse, number of sexual partners, extent of discussion of sexual topics with parents, and transitions in life stages and relationships.

Demographic variables such as ethnic background, age, education, and socioeconomic status were also used as predictor measures.

In addition to the published summary scales for the questionnaires listed above, ten post-hoc scales were constructed, each made up of useful predictor items from one of the important topic areas hypothesized to be related to contraceptive success. All items for each scale were required to be conceptually related, to have significant positive intercorrelations, to be answered by a relatively large number of overlapping subjects, and to have approximately equal variances (after recoding, if necessary). (These same requirements were also applied to the large number of a priori scales constructed in Study Two.) The number of items in these scales ranged from 3 to 9, and their reliability, using Cronbach's (1970) alpha coefficient

of homogeneity, ranged from .31 to .90. Six of the scales had alphas below .50, while three had alphas above .70.

Follow-Up Procedure and Criterion Classification

At these clinics, patients using oral contraceptives (83% of the Research Group) normally return about three months after intake to get another supply of pills, and IUD patients are expected to return about one month after the IUD's insertion for a check on its placement. Diaphragm users and condom and/or foam patients are not expected to return until the time for regular medical exams, which are scheduled at six months and one year after intake for all patients. Thereafter, patients may return at six-month or one-year intervals, or more often if they prefer.

The clinic routinely sends follow-up letters to patients (unless they have specified no mail contacts) reminding them to make an appointment for the six-month, one-year, and subsequent medical exams. Of course, patients can come in at any time for help with medical problems, side effects, and method changes, or to get further birth control supplies.

Additional follow-up procedures beyond the regular clinic routine were established for patients [in the study]. All nonreturners who had agreed to allow mail contacts were sent a second letter about two months after the first one. It again reminded them about returning to the clinic, and it requested information about whether they were getting birth-control help elsewhere or had decided to stop using birth control. If these letters proved to be undeliverable, they were returned to us by the dead-letter section of the post office. Later, patients who had neither returned to the clinic nor answered the letters were followed up by telephone (if they had indicated that phone contacts were acceptable). Up to five phone calls at various hours were made, and if the patient was reached, a brief and diplomatic phone interview asked about her current contraceptive status using questions similar to those in the follow-up letter.

Based on these follow-up procedures, a criterion classification system was established involving several different measures of contraceptive success or failure. It included behavioral indices of patient attrition, amount of delay in return visits to the clinic, and the ultimate outcome measure, occurrence of unwanted pregnancies. These measures were determined for each patient as of five different points in time after her clinic intake (1-3 months, 6 months, 1 year, 2 years, and 3 years). The final 3-year classification of criterion status, made about three and one-half years after clinic intake, showed the following distribution of categories for patients in the Research Group:

1. unwanted pregnancies (N=76)--about 80% of these cases opted for abortions,
2. pregnancies in which wantedness was questionable (N=15)--this group included six clear cases of method failure where an IUD was still

in place, and two cases where a doctor had advised discontinuing birth control since a tipped uterus supposedly prevented any risk of pregnancy; all 15 cases were omitted from further analyses to reduce error variance,

3. wanted pregnancies (N=22),
4. unknown status; no recent contact nor information to indicate classification in any other category (N=272),
5. moved away, and no additional information (N=27),
6. discontinued using birth control (N=35),
7. getting birth control elsewhere (N=98),
8. returned to clinic at about three years, but not completely regular and prompt for full three years (N=44),
9. regular, prompt returners over whole three-year period (N=56).

Two independent raters displayed 92% agreement using this classification scheme. To correct for chance levels of agreement, a more precise index of inter-coder agreement for nominal scale judgments was computed, as recommended by Scott and Wertheimer (1962, p. 194). For a random sample of 285 judgments by each rater, the value of the π coefficient was .905.

Several different analyses have been made of the above criterion groups. In particular, the unwanted pregnancy group (1 above) has been compared with the regular return group (9 above), with all other patients (3-9 above, omitting group 2), and with the patients in groups 3, 7, 8, and 9 combined. This latter combined group was termed "birth planners" since they had successfully achieved their birth-control goals--that is, they either continued their contraceptive program successfully for three years without becoming pregnant, or they continued it successfully until they decided to have a child. Frequently, such a decision to have a child followed a change of status such as getting married or completing schooling. Combining the wanted pregnancy group together with the other three nonpregnant groups for this analysis was justified partly on the rational grounds that they had all achieved their contraceptive goals, and partly on the basis of empirical tests which showed that they resembled the nonpregnant groups more than they did the unwanted pregnancy group.

Study Two--Method

In planning this second study, we had the benefit of our first study's results for the one-year criterion point. The second study dealt solely

with teenagers, focusing on the important early stages of their sexual and contraceptive experience. It was planned as a more extensive and intensive investigation of a broad range of factors which might affect their contraceptive success or failure. In most respects, the research design was similar to that of Study One. Both studies were guided by the same general hypotheses, as previously summarized; but the specific questionnaires used were quite different, as indicated below.

Sample

Sites for the study were eleven different family-planning clinics. In addition to the four clinics used in Study One, the sites included two other private suburban clinics run by Los Angeles Planned Parenthood, and five public county-run clinics in different areas of Orange County (a part of greater Los Angeles). A consecutive sample of approximately 700 new contraceptive patients below the age of 20 was contacted at the intake visits [for a period of six months] starting early in 1977. Because of differences in the demographic characteristics of the patient populations and in the follow-up procedures which we were allowed to use, only the private Planned Parenthood data are comparable to Study One. Therefore, the approximately 200 Orange County patients are not included in this report.

Description of Subjects

The patients in Study Two had a mean age of 17.8 and had completed an average of 11.2 years of school. They included 13% blacks, 13% Latin or Spanish surnamed, and 1% Asians--a bit lower percentage of minorities than in Study One. About 71% were still living with their families, but . . . clinic policies in recording the data caused most of these girls to be considered as financially independent. As a result, their average family income was reported as being \$2100 per year (probably much underestimated) for an average family size of 1.2 (also underestimated). However, the various indices of their parents' income and education . . . suggest that many of them were in the middle-class range.

As with [the teenage portion of the Study One sample], almost all these teenagers were still single--92% never married, 7% married, and 1% separated or divorced. Also, much as before, 30% of them had been pregnant, 21% had had at least one therapeutic abortion (more than in Study One), and 8% had living children (somewhat less than [the teenagers] in Study One). Only about 6% had never had intercourse, but 43% had never used birth control. Over half (62%) had not been using contraception recently, while about 20% had been using pills and about 18% some other method. Their clinic prescription was predominantly pills or pills plus foam and/or condom (81%--somewhat less than [the teenagers] in Study One), with only 4% being given an IUD, 7% getting a diaphragm, and 8% receiving only foam and/or condom. As before, a heavy majority (72%) of these young women said they wanted a child or children later on, while only 4% were definite in

wanting no more children.

Instruments

Data were gathered from clinic records, extensive research questionnaires, and brief interviews by specially trained female research staff members. The major types of variables studied were listed above prior to the description of Study One. An attempt was made to develop short scales having from 2 to 10 items to measure each major variable, and in some cases multiple alternative scales or items were used. Specific scales and items were adapted or selected from our own previous research and from studies by Ager (1976), Ball (1973), Bendig (1956), Cole (1975), Goldsmith, et al. (1972), Jessor and Jessor (1975), Kantner and Zelnik (1973), Kar (1976), Kothandapani (1971), Miller (1976), Rosenberg (1965), Rotter (1966), Slagle, et al. (1974), Stein, et al. (1968), Vener and Stewart (1974), and several other investigators.

The items chosen for use were organized by topic into one brief interview and six questionnaires. All subjects received the interview, which dealt mostly with family occupational and educational information, and two questionnaires, which concerned (a) sex and birth-control knowledge and (b) sex and birth-control experiences. Each subject also received two of the other four questionnaires, with the forms in use being alternated each week. These forms concerned (c) sexual attitudes, (d) birth-control attitudes, intentions, and inhibiting extraneous events, (e) personality traits and related attitudes, and (f) parental and peer relationships and communication regarding sex, birth control, and deviant behavior. Thus, questions on these last four topics were answered by only about half of the subjects. The interview and four questionnaires took most subjects about an hour to complete, but only 1% of the patients refused to participate.

Reliability of Predictor Scales

For each major concept [to be investigated], one or more scales was assembled, each containing 2-10 items [and each meeting the requirements stated under Study One]. The reliability of the 66 resulting scales was determined using Cronbach's (1970) alpha coefficient of homogeneity. In general, the results were quite satisfactory, particularly in view of the very short length of most scales. Quite a number of even the 2-item scales had alpha coefficients in the .70s and .80s. Overall, there were 13 scales with alphas of .80 or more, 38 scales with coefficients between .50 and .80, and only 15 relatively poor scales with alphas below .50.

Follow-Up Procedure and Criterion Classification

The same general follow-up procedure as in Study One was used. . . .

About two years after their initial clinic visit, the 465 Planned Parenthood teenage patients were followed up, first through a check of clinic records, and then by mailing the same questionnaire as in Study One to patients who had not returned to the clinic for their 1-year visit nor contacted the clinic subsequently. Because these young women were in a highly mobile period of their lives, it was decided that attempts to telephone the patients who had not returned the questionnaire would not be productive.

The final criterion classification was based primarily on the 18-month return visit point for which data were available on almost all patients. The same basic indices and categories of contraceptive success or failure were used as in Study One, with some minor refinements. Patients were classified into the following criterion categories:

1. unwanted pregnancies ($N=47$),
2. pregnancies where wantedness was questionable ($N=7$, including 2 cases of possible method failure)--omitted from further analyses to reduce error variance,
3. wanted pregnancies ($N=10$),
4. unknown status; no recent contact nor information about their criterion status ($N=216$),
5. moved away, and no additional information ($N=5$),
6. miscellaneous communications, but no recent return to clinic ($N=6$),
7. discontinued using birth control ($N=14$),
8. getting birth control elsewhere ($N=25$),
9. reported abstinence as their latest method, and had not gotten pregnant since intake ($N=7$),
10. returned to clinic at around 18 months, but not completely regular in all return visits ($N=41$),
11. regular, prompt returners for full 18 months ($N=87$).

As in Study One, a criterion group of "birth planners" was defined, in this case made up of categories 3, 8, 9, 10, and 11 (combined $N=170$).

Discussion

Methodological Issues

In addition to the usual advantages and disadvantages of longitudinal research, there are several methodological issues which are important in this research. The first is generalizability. Since the subjects were clinic patients, whose awareness of and motivation for contraception would be expected to be higher than the general population's, our findings may not apply to all women of child-bearing age. In particular, the factors which predict success or failure in avoiding pregnancy are apt to be somewhat different for girls who have not yet acknowledged that they are sexually active, and also for older women and/or married women. However, our samples did cover a broad socioeconomic range and were approximately representative of the ethnic makeup of the Los Angeles area. Generalization to other similar clinic populations does seem warranted, since our samples were large and consisted of representative, unselected clinic patients. Moving beyond female patients, we made an attempt to obtain data from male partners of our subjects, but the number available and willing to participate was too small to yield stable data.

Concerning comparability of samples, the Research Group and the Control Group were randomly chosen and demographically identical. The outcome groups displaying contraceptive success or failure also turned out to be very similar demographically, so comparisons of their personalities, attitudes, experiences, etc., were largely unconfounded with demographic differences.

Another issue is the possible disadvantages of outcome measures, such as contraceptive success or failure, as compared to more directly behavioral measures (e.g., regularity of using contraception). It is apparent that outcomes such as pregnancy can be affected by many factors besides the subject's contraceptive behavior--for instance, her fecundity, the extent of her opportunities for sexual contact, accidents (as seen in the occurrence of pregnancies with IUDs in place), and males' actions (as in forceful rape, refusal to use condoms, concern with contraceptives, vasectomy, etc.). However, some of these factors are rare and others will tend to balance out over long time periods. The major advantages of outcome measures are that they can be objectively determined and that they are largely dependent on the subject's behavior. In contrast, behavioral measures such as regularity of contraceptive usage must often be determined from self-reports, and there are strong pressures on subjects to give socially desirable reports rather than objectively truthful ones. We did obtain such reports, but whenever possible we gave preference to criterion measures which were more objective, such as regular return for clinic appointments.

In general, the use of objective behavioral or outcome measures has been much too rare in contraceptive research. The great majority of psychological studies have used birth-control attitudes or intentions as

the criteria to be predicted, rather than actual behavior. Our findings demonstrate that though attitudes and intentions are important, they represent just one of several domains which must be considered in attempting to understand contraceptive behavior and outcomes.

As another methodological refinement, we took great care to make criterion classification as accurate and complete as possible. For instance, wanted and unwanted pregnancies were carefully distinguished, as recommended by Cobliner, Schulman, and Smith (1975). In addition, questionable cases such as method failures and iatrogenic problems were excluded from the data analyses in order to keep the criterion groups as homogeneous as possible.

A final question might be the adequacy of follow-up in these studies. The relatively large number of patients who could not be contacted was rather disappointing. However, there were many reasons for this result which suggest that a fuller accounting would have been almost impossible to obtain. First, there were important ethical guidelines which prevented the clinic or the research staff from contacting any patient who had requested no mail and/or no phone calls. Many of the teenagers in the sample (and some of the older women) were very concerned that no one in their household should learn of their contact with Planned Parenthood. In numerous cases, this extended to giving the clinic false addresses, phone numbers, and even false last names--thus making follow-up nearly impossible. Many of the poorer women did not have a phone, and in other cases we found it disconnected or transferred to another party when we tried to follow-up by phone. Sometimes the parents or husbands of the patients were suspicious of our anonymous callers and refused to relay messages.

Probably the most basic problem in follow-up was the patients' mobility--a majority of these young women were at a very mobile stage of their lives in a state with a highly mobile population, and many came from especially transient socioeconomic subgroups. Consequently, many of our follow-up letters came back via the dead-letter office or were returned marked "addressee unknown." We were able to trace a substantial number of women who had moved through use of the public records of the state Department of Motor Vehicles--the most extensive of our follow-up efforts. For women who reported they were going elsewhere for birth control or who had had their records transferred to another clinic, we requested reports of their status from their new doctor or clinic. The only available follow-up method which we did not use was face-to-face personal inquiries directed to former neighbors and friends. We felt that the intrusiveness involved in that method was too great to justify it, and in addition its payoff was likely to be small.

Conclusions

In summarizing the value of these studies, we would stress that they share both the strengths and weaknesses of naturalistic, longitudinal,

multivariate research designs. They have investigated socially important questions, studied them in real-life situations, and tracked the research subjects for several years in order to develop stable measures of their behavior. On the other hand, in this kind of research there are many complex interrelationships among the variables, and there is no possibility of experimentally controlling the conditions and experiences of the subjects. Since the design is correlational in nature, it is more difficult to demonstrate causal connections between variables. However, despite these inevitable limitations, we have found support for a number of hypotheses based on past research studies, and we have had some success in predicting important contraceptive behaviors on a longitudinal basis.

Finally, from a practical point of view, these methods and findings may not only have predictive value, but they may also suggest new approaches to planning treatment strategies in family-planning agencies, both public and private. A basic goal in such intervention programs must be to break the web of relationships between poor socialization, poor personal efficacy and adjustment, insufficient information, poor planning, unfavorable attitudes and intentions, and responsiveness to external pressures toward deviancy--all of which can contribute to inadequate contraceptive behavior in which society and the [unsuccessful contraceptive] both lose. Future research should continue to be directed toward breaking this chain of relationships.

References

- Ager, J. Comparison of participants and dropouts from a contraceptive teen program. Paper presented at the American Psychological Association meeting, Washington, D.C., September 1976.
- Argyris, C. Dangers in applying results from experimental social psychology. American Psychologist, 1975, 30, 469-485.
- Baldwin, W.H. Adolescent pregnancy and childbearing--Growing concerns for Americans. Population Bulletin, 1976, 31(2), 1-34.
- Ball, G.W. A method of identifying the potential unwed adolescent. Unpublished doctoral dissertation, University of California at Los Angeles, 1973.
- Bendig, A.W. The development of a short form of the manifest anxiety scale. Journal of Consulting Psychology, 1956, 20, 384.
- Buros, O.K. (Ed.) The seventh mental measurements yearbook. Highland Park, N.J.: Gryphon Press, 1972.
- Chilman, C.S. Adolescent sexuality in a changing American society: Social and psychological perspectives. Washington, D.C.: U.S. Government Printing Office, 1979.
- Coblener, W.G., Schulman, H., & Smith, V. Patterns of contraceptive failures: The role of motivation re-examined. Journal of Biosocial Science, 1975, 7, 307-318.
- Cole, S.G. Critical factors in family planning participation. Final progress report for NICHD grant, Texas Christian University, 1975.
- Cronbach, L.J. Essentials of psychological testing (3rd ed.). New York: Harper & Row, 1970.
- Cronbach, L.J. Beyond the two disciplines of scientific psychology. American Psychologist, 1975, 30, 116-127.
- David, H.P., & Baldwin, W.P. Childbearing and child development: Demographic and psychosocial trends. American Psychologist, 1979, 34, 866-871.
- Goldsmith, S., Gabrielson, M., Gabrielson, I., Matthews, V., & Potts, L. Teenagers, sex and contraception. Family Planning Perspectives, 1972, 4, 32-33.
- Gough, H.G. Manual for the personal values abstract. Palo Alto, Calif.: Consulting Psychologists Press, 1972.

- Jessor, S.L., & Jessor, R. Transition from virginity to nonvirginity among youth: A social-psychological study over time. Developmental Psychology, 1975, 11, 473-484.
- Kantner, J.F., & Zelnik, M. Contraception and pregnancy: Experience of young unmarried women in the United States. Family Planning Perspectives, 1973, 5(1), 21-35.
- Kar, S.B. Consistency between fertility attitudes and behavior: A conceptual model. Paper presented at American Psychological Association meeting, Washington, D.C., September 1976.
- Kothandapani, V. A psychological approach to the prediction of contraceptive behavior. Chapel Hill, N.C.: Carolina Population Center, 1971.
- McGuire, W.J. The yin and yang of progress in social psychology: Seven koan. Journal of Personality and Social Psychology, 1973, 26, 446-456.
- Miller, W.B. The personal style inventory. Unpublished manuscript, Stanford University, 1973.
- Miller, W.B. Some psychological factors predictive of undergraduate sexual and contraceptive behavior. Paper presented at the American Psychological Association meeting, Washington, D.C., September 1976.
- Miller, W.B. Sexual and contraceptive knowledge questionnaire. Unpublished manuscript, Stanford University, no date.
- Mindick, B. Personality and social psychological correlates of success or failure in contraception: A longitudinal predictive view. Unpublished doctoral dissertation, Claremont Graduate School, Claremont, California, 1978.
- Mindick, B., & Oskamp, S. Longitudinal predictive research: An approach to methodological problems in studying contraception. Journal of Population, 1979, 2, 259-276.
- Oskamp, S., & Mindick, B. Personality and attitudinal barriers to contraception, in D. Byrne & W.A. Fisher (Eds.) Adolescents, sex, and contraception. New York: McGraw-Hill, 1981, in press.
- Oskamp, S., Mindick, B., Berger, D., & Motta, E. A longitudinal study of success versus failure in contraceptive planning. Journal of Population, 1978, 1, 69-83.
- Rosenberg, M. Society and the adolescent self-image. Princeton, N.J.: Princeton University Press, 1965.
- Rotter, J.B. Generalized expectancies for internal versus external control of reinforcement. Psychological Monographs, 1966, 80(1, Whole No. 609).

- Scott, W.A., & Wertheimer, M. Introduction to psychological research. New York: Wiley, 1962.
- Slagle, S.J., Arnold, C.B., & Glascock, E. Self-competence: A measure of relative risk of unwanted pregnancy? Paper presented at the American Psychological Association meeting, New Orleans, 1974.
- Stein, K.B., Sarbin, T.R., & Kulik, J.A. Future time perspective. Journal of Consulting and Clinical Psychology, 1968, 32, 257-264.
- Vener, A.M., & Stewart, C.S. Adolescent sexual behavior in middle America revisited: 1970-1973. Journal of Marriage and the Family, 1974, 36, 728-735.
- Zelnik, M., & Kantner, J.F. Contraceptive patterns and premarital pregnancy among women aged 15-19 in 1976. Family Planning Perspectives, 1978, 10, 135-142.

A Longitudinal Study of Personality and Cognitive Development

Jack Block and Jeanne H. Block**
University of California, Berkeley

** The research reported is being supported by a National Institute of Mental Health Research Grant MH 16080 to Jack and Jeanne H. Block and by a National Institute of Mental Health Research Scientist Award to Jeanne H. Block. The ordering of the authors' names is immaterial; they could as well be reversed. As with other "projects" we have parented and nurtured, we both have contributed, differently but equally. We wish to express our appreciation to the children and their families who continue to participate in our longitudinal study. Much has been asked of them; they have responded with a grace and enthusiasm that we find affirming. We are grateful to the teachers, both at the Harold E. Jones Child Study Center and in the cooperating Bay Area elementary schools who have devoted many hours to completing the Q-sort descriptions upon which our analyses depend heavily. The administrative staff in these schools also have facilitated our research efforts in many ways for which we are thankful. We are deeply appreciative of the research staff, Myrna Walton, Jolinda Taylor, Judy Casaroli, Mark Haarz, Joy Moore, Mimi Rosenn, Rachel Melkman, Suzie Schmookler, Tammy Socher, Jon Feshbach, Elaine Simpson, Marjorie Hayes, Jackie Heumann, and Betty Goodman, whose diligence, sensitivity and commitment during the process of data collection were essential to the success of our research venture. In addition to participating in data collection, Drs. Anna von der Lippe and Ellen Siegelman importantly contributed to the development of the assessment battery for the three-year-olds. Finally, we wish to thank our colleague, Professor David Harrington, for the many responsibilities he assumed on this project over the years, participating in data collection, supervising the scoring and coding of the data, managing the complex data bank, and for his many psychological perceptions.

A Longitudinal Study of Personality and Cognitive Development

Jack Block and Jeanne H. Block
University of California, Berkeley

Introduction

Many years ago, we began studying two personality parameters which we chose to call ego-control and ego-resiliency. In 1968, we initiated a longitudinal study of approximately 130 children whom we individually assessed at the ages of three, four, five, seven and most recently eleven years. At each assessment point, every child was administered a battery of from 26 to 43 widely ranging experimental procedures, requiring 10 or 11 testing sessions at ages three and four, 4 or 5 sessions at ages five and seven and 9 sessions at age eleven. In addition, we have developed extensive assessment data on the parents of these children and on parent-child interaction styles.

Goals of the Study

We had four major goals for our research:

1. To investigate two parameters of personality functioning, ego-control and ego-resiliency, with regard to their developmental course over the childhood years.
2. To explore the relations of ego-control, ego-resiliency and their interaction, to cognitive functioning, affective differentiation, moral development and interpersonal behaviors.
3. To identify parental, environmental and experiential factors associated with differences in ego-control and ego-resiliency.

4. To assess the predictive utility of these two constructs measured in early childhood for understanding personality characteristics, interests, achievement orientation, attitudes and adaptations in the preadolescent and adolescent years.

The Sample

The children included in our study were drawn from the two nursery schools constituting the Harold E. Jones Child Study Center at Berkeley over the three-year period 1969-71. Extensive individual assessments were conducted at ages three, four, five, seven and eleven. The numbers of children participating vary by year, ranging from a high of 130 at age four to a low of 104 at age seven. This seems to be a low attrition rate given the great mobility characterizing American society. Earlier analyses of subject loss showed no differential attrition as a function of socio-economic level or ethnic origin.

One of the participating nursery schools is a university laboratory school, administered by the University of California; the second nursery school is a parent cooperative, administered by the Berkeley Public Schools. The two schools, jointly considered, attract children from heterogeneous backgrounds with regard to education, socio-economic level and ethnic origin. Although the sample overrepresents the middle and upper middle class, the SES range is wide. Sixty-one percent of the children are White, 31% are Black and the remaining 8% represent other ethnic groups, primarily Oriental and Chicano.

Both the mothers and the fathers of the children are also included in the study. When their children were three-years-old, the parents each provided descriptions of their child-rearing orientations; when their children were four-years-old, the parents each interacted with their children in a teaching situation; when the children were six-years-old, the mothers were interviewed and completed both self-descriptive and child-descriptive adjective Q-sorts. Currently, additional parental data are being collected.

Operationalizing the Constructs of Ego-Control and Ego-Resiliency

Our conceptual expectations shaped the choice of tests and experimental situations included in the assessment battery so that they indexed ego-control and ego-resiliency at each age level.

These experiment-based indices of ego-control and ego-resiliency are presented in Table 1.

Insert Table 1 about here

In addition to the experiment-based indices of ego-control and ego-resiliency, a second, entirely independent method was used to objectify these concepts. This involved comprehensive personality descriptions of the child provided by the nursery and elementary school teachers using the California Child Q-Set (J.H. Block & J. Block, 1969). The operationalization of our concepts of ego-control and ego-resiliency, as well as evidence in support of its construct validity, are presented elsewhere (J.H. Block & J. Block, 1980).

To summarize these several validation efforts to evaluate the usefulness and generality of the ego-control and ego-resiliency concepts, we suggest that we have been able to demonstrate for both sexes over time across different samples of different ages and across methods of experimentation, observation and self-report the presence in behavior of two context-responsive personality subsystems that separately and in conjunction appear to have consequentiality.

Results

By considering a subset of the CCQ findings, we have found interesting reciprocal implications of ego-control and ego-resiliency. If we attend only to the CCQ-superitems conjointly related to ego-control and ego-resiliency, classifying these items according to the particular pattern of ego-control and ego-resiliency relationships manifested, a vivid psychological portrait emerges of the individuals representing each of these patterns (see Table 2).

Insert Table 2 about here

While the correlation between ego-control and ego-resiliency as measured at age three is of near zero order ($r = -.10$), the position of an individual child in this two-dimensional psychological space has strong implications for manifestly different patterns of interpersonal functioning. For the undercontrolling child, the presence of ego-resiliency tends

to temper the expression of impulse without suppressing spontaneity, engagement and enthusiasm. For the undercontrolling child with little ego-resiliency, however, impulse is unmodulated and we see a restless, externalizing, impulsive, easily-disrupted child, a syndrome that fits more than a little the description of the hyperactive child (Whalen & Henker, 1976).

For the overcontrolling child, the presence of ego-resiliency results in a high degree of socialization that fits and feels well, a relative absence of anxiety and intimidation in reacting to and acting on the world. The overcontrolling child with little ego-resiliency appears victimized, immobilized, anxious, overwhelmed by a world apprehended as threatening and unpredictable. There is appreciable evidence of psychopathology in the Brittle Overcontroller, as evidenced by behavioral mannerisms, inappropriate affect and immobilization when confronted by stress.

Thus, within the individual, the reciprocal interactions of ego-control and ego-resiliency have much consequence for the nature of perceptual and behavioral premises, the development of character structure and the quality of functioning in the interpersonal world.

With respect to the CCQ-superitems reflecting orientations toward the self and the world, children who were ego-resilient at age three were seen as task oriented and as emphasizing of autonomy and independence at age four while less resilient children were seen as critical and devaluing of both self and others. Undercontrol at age three relates to an emphasis on and pleasure derived from competition, a manifestation of engagement and desire for active mastery also seen in the T-data as salient components of undercontrol. Children who were described at three as overcontrolling were seen at age four as enjoying solitary play and oriented toward fantasy, manifestations of reserve and psychological "distancing."

The CCQ-superitems in the realm of social behaviors also are substantially influenced by early ego-control and ego-resiliency. In general, positive patterns of social interaction (e.g., empathy, protectiveness, relatedness, directness) are associated with early ego-resiliency while negative interpersonal behaviors (e.g., teasing, manipulateness) are associated with early undercontrol.

Ego-resiliency as assessed at age three has significance for the later evaluated social stimulus value of four-year-olds and, to a lesser extent, seven-year-olds. Early ego-resiliency is associated with positive social stimulus value: with being seen as popular, interesting and physically attractive at age four. The ability to modify behavior in response to the demand qualities of a situation, a hallmark of ego-resiliency, benefits social interactions where differentiated responsiveness

to moods, interests, needs and sensitivities of others is a necessary condition for communication and relatedness. It is not surprising that the more resourceful, responsive, resilient children evoke positive evaluations and reactions from both adults and peers. Negative reactions in the form of scapegoating and victimization, on the other hand, are elicited by overcontrolling children. One might conjecture that these negative, aggressive behaviors represent extreme attempts by their peers to evoke responses from these inhibited, shy, constrained children.

In summarizing the results from the data, abundant evidence for behavioral coherencies with regard to ego-control and ego-resiliency across time was found. We also found that these concepts have implications for interpersonal behaviors, for social evocateness, for patterns of personal adaptations and for psychopathology.

Coda

In the last 15 years or so, the field of developmental psychology has been dominated by issues of cognitive development. The journals record this general emphasis; due homage especially should be paid to Piaget for his important contribution. But for all the importance and centrality of cognitive psychology to the understanding of many aspects of behavior, it should also be realized that cognitive psychology, as generally conceived and generally studied, deals with only a small portion of the mental lives, the experience, and the consequently forged behaviors of individuals.

The psychological world of the individual is surprisingly seldom occupied by the purely cognitive problems with which psychologists have been prone to concern themselves. Cognition in everyday life is not oriented around such problems as pouring water into containers of different shape or the factors influencing the swing of a pendulum. Paradigmatic and essential to study though such cognitive problems are, they are problems without a social or interpersonal context, placed before an individual presumed to function without passions and without highly personal, often behaviorally preempting affect-optimizing criteria.

Thus, psychological thinking about how cognitive structures are created and transformed has derived primarily from consideration of how children interact with the physical world. In the version of the physical world provided to the child, the child observes a "reality" that is impressively orderly and follows clearly inferable rules. The child can test the nature of this reality by actions that elicit direct, immediate, unequivocal feedback from the world. Further, and for entirely

useful reasons, the cognitive problems employed are selected to be distant from the central motivational system of the child so that the "cool" process of cognitive structural development can be perceived most purely, without the intrusive influence of "hot" motivations.

But children live in an interpersonal world as well, a world which depending on time, place and person can be far more central, psychologically than the physical world. They must construct intake, output and integrational structures for dealing with this interpersonal world. The social world is complex and perhaps ultimately fractious, behaving in ways only fuzzily comprehensible. Efforts by the child to test the nature of the interpersonal reality by acting upon it have erratic or dim results; social feedback is often indirect, delayed (sometimes forever in the interpersonal sphere) and equivocal, permitting only the uneasiness of uncertain inference instead of the pleasure of certain deduction afforded by the physical world. Because there is little or no feedback with which to build cognitive structures for dealing with the social world, the child must evolve and apply not-so-cognitive structures for the prediction of behavior. In doing so, the principles that come into play to govern such structures, construed in the absence of unmistakable, unambiguous feedback, are principles that are less than rational. They include the forms of irrationality called "primary process modes of thought" (Freud, 1900), the "cognitive illusions" observed by Tversky and Kahneman (e.g., 1974), and attributional errors in social judgment (e.g., Jones, Kanouse, Kelly, Nisbett, Valins & Weiner, 1972), among others.

"Irrational" modes of perception, action and cognition are due, at least in part, to intrinsic information-processing limitations of the human mind. But also, these intrinsic limitations develop the possibility for individuals to be influenced by strong motivations, pervading fears and prevailing aspirations. The extent and the way in which reigning passions will affect development of the not-so-cognitive structures can be expected to be a joint function of the efficacy of the reality-testing available to the developing child and the motivational stresses being endured.

The processes underlying the development of not-so-cognitive structures (premises, orientations, attitudes on the world, deuterolearnings, scripts, etc.) are the processes underlying the development of personality or ego structures. Conceived in these terms, the timeliness of extended study of character development becomes apparent; there is a promise of a fruitful integration of cognitive and personality psychology. In that integration, reasoning from the relationships reported earlier, we anticipated that concepts very much akin to ego-control and ego-resiliency will be found useful.

References

- Block, J. An experimental investigation of the construct of ego-control. Unpublished doctoral dissertation, Stanford University, 1950.
- Block, J. The challenge of response sets. New York: Appleton-Century-Crofts, 1965.
- Block, J. Lives through time. Berkeley, California: Bancroft Books, 1971.
- Block, J. Recognizing the coherence of personality. Unpublished manuscript, University of California, Berkeley, 1975.
- Block, J. Advancing the science of personality: Paradigmatic shift or improving the quality of research? In D. Magnusson and N.S. Endler (Eds.), Psychology at the crossroads: Current issues in interactional psychology. Hillsdale, N.J.: Lawrence Erlbaum Associates (Wiley), 1977.
- Block, J. & Block, J.H. An investigation of the relationships between intolerance of ambiguity and egocentrism. Journal of Personality, 1951, 19, 303-311.
- Block, J. & Petersen, P. Some personality correlates of confidence, caution, and speed in a decision situation. Journal of Abnormal and Social Psychology, 1955, 51, 34-41.
- Block, J. & Thomas, H. Is satisfaction with self a measure of adjustment? Journal of Abnormal and Social Psychology, 1955, 51, 254-259.
- Block, J. & Turula, E. Identification, ego-control and adjustment. Child Development, 1963, 34, 945-953.
- Block, J.H. An experimental study of a topological representation of ego structure. Unpublished doctoral dissertation, Stanford University, 1951.
- Block, J.H. Issues, problems, and pitfalls in assessing sex differences: A critical review of the Psychology of sex differences. Merrill-Palmer Quarterly, 1976, 22, 283-308.
- Block, M.H. & Block, J. An interpersonal experiment on reaction to authority. Human Relations, 1952, 5, 91-98.

- Block, J.H. & Block, J. The California Child Q-set. Institute of Human Development, University of California, Berkeley, 1969. (In mimeo).
- Block, M.H. & Block, J. Fire and young children: A study of attitudes, behaviors, and maternal teaching strategies. Technical Report for Pacific Southwest Forest and Range Experimental Station, Forest Service, U.S. Department of Agriculture, 1975.
- Block, M.H. & Block, J. The role of ego-control and ego-resiliency in the organization of behavior. In W.A. Collins (Ed.), Minnesota Symposia on Child Psychology, Volume 13. New York: Lawrence Erlbaum, 1980.
- Block, J.H., Block, J. & Folkman, W. Fire and children: Learning survival skills. U.S. Department of Agriculture Forest Service Research Paper, PSW-119, 1976.
- Block, J.H. & Martin, B. Predicting the behavior of children under frustration. Journal of Abnormal and Social Psychology, 1955, 51, 281-285.
- Cattell, R.B. Personality and motivation structure and measurement. New York: World Book Company, 1957.
- Cattell, R.B. Personality and mood by questionnaire. San Francisco: Jossey-Bass, 1973.
- Cohen, J. & Cohen, P. Applied multiple regression/correlation analysis for the behavioral sciences. Hillsdale, N.J.: Lawrence Erlbaum, 1975.
- Flavell, J.H. The development of role-taking and communication skills in children. New York: John Wiley & Sons, 1968.
- Freud, S. (1900) The interpretation of dreams. In Standard Edition, Vol. 4-5, 1953. London: Hogarth Press, 1953.
- Harrington, D., Block, J.H. & Block, J. Intolerance of ambiguity in preschool children: Psychometric considerations, behavioral manifestations and parental correlates. Developmental Psychology, 1978, 14, 242-256.
- Jones, E.E., Kanouse, D.E., Kelley, H.H., Nisbett, R.E., Valins, S. & Weiner, B. Attribution: Perceiving the causes of behavior. Morristown, N.J.: General Learning Press, 1972.

- Kafry, D. Fire survival skills: Who plays with matches? Technical Report for Pacific Southwest Forest and Range Service, U.S. Department of Agriculture, 1978.
- Keller, A., Ford, L.H. & Meacham, J.A. Dimensions of self-concept in preschool children. Developmental Psychology, 1978, 14, 483-489.
- Krebs, D. & Sturupp, B. Altruism, egocentricity, and behavioral consistency in children. Paper presented at the meeting of the American Psychological Association, New Orleans, September, 1974.
- Lewin, K. A dynamic theory of personality. New York: McGraw-Hill, 1935.
- Lewin, K. Principles of topological psychology. New York: McGraw-Hill, 1936.
- Lewin, K. The conceptual representation and the measurement of psychological forces. Durham, N.C.: Duke University Press, 1938.
- Lewin, K. Field theory in social science. New York: Harper, 1951.
- Piaget, J. The construction of reality in the child. New York: Basic Books, 1954.
- Piaget, J. & Inhelder, B. The child's conception of space. London: Routledge and Kegan Paul, 1956.
- Rubin, K.H. & Schneider, F.W. The relationships between moral judgment, egocentrism, and altruistic behavior. Child Development, 1973, 44, 661-665.
- Rutherford, E. & Mussen, P.H. Generosity in nursery school boys. Child Development, 1968, 39, 755-765.
- Selman, R.L. The relation of role taking to the development of moral judgment in children. Child Development, 1971, 42, 79-91.
- Tversky, A. & Kahneman, D. Judgment under uncertainty: Heuristics and biases. Science, 185, September 27, 1974.
- Whalen, C.K. & Henker, B. Psychostimulants and children: A review and analysis. Psychological Bulletin, 1976, 83, 1113-1130.
- Witkin, H.A. & Goodenough, D.R. Field dependence and interpersonal behavior. Psychological Bulletin, 1977, 84, 661-689.

Table 1

Components Selected on an A Priori, Conceptual Basis of the Experiment-Based
Composites of Ego-Control and Ego-Resiliency
at Ages Three, Four, Five, and Seven

	Ego-Control		
Age Three	Age Four	Age Five	Age Seven
Actometer scores, composited over four readings	Actometer scores, composited over four readings		
Delay of Gratification Score (reflected)	Delay of Gratification Score (reflected)	Delay of Gratification (reflected)	
Barrier Tasks, Intensity scores, composited over Barrier Door and Block Tower tasks	Barrier Tasks, Intensity scores, composited over Barrier Drawer and Block Box tasks	Barrier Puzzle Task, Intensity Score	
Satiation & Cosatiation Trials	Satiation & Cosatiation Trials	Satiation & Cosatiation Trials	
Percept Recognition, mean trial first guess offered (reflected)	Percept Recognition, mean trial first guess offered (reflected)		
Competing Set - Contextual errors	Competing Set - Contextual errors		
Motor Inhibition - mean time, standard condition, two trials on each of three tasks	Motor Inhibition - mean time, standard condition, two trials on each of three tasks	Motor Inhibition - mean time, standard condition, two trials on each of three tasks.	

-404-

Table 1
(Continued)

Age Three	Ego-Resiliency Age Four	Age Five	Age Seven
Motor Inhibition - Regression-adjusted scores in Slow condition, composited over three tasks	Motor Inhibition - Regression-adjusted scores in Slow condition, composited over three tasks	Motor Inhibition - Regression-adjusted scores in Slow condition, composited over two tasks	Digit Span Backwards, Total score
Dual Focus - Regression-adjusted score		Dual Focus-Regression-adjusted score	Spatial Visualization, Correct, nonegocentric responses
-465- Level of Aspiration - Discrepancy score (reflected)	Level of Aspiration - Discrepancy score (reflected)		
Distractability - Power score	Distractability - Power score		Rod and Frame Test, Accuracy score
Percept Recognition - Mean trial of correct recognition (reflected)	Percept Recognition - Mean trial of correct recognition (reflected)		
Incidental Learning - Total recall		Incidental Learning - Total recall	
Sigel Object Sorting Task - Mean number of objects in Scorable categories	Sigel Object Sorting Task - Mean number of objects in Scorable categories	Unstructured Object Sorting Task - number of objects in Scorable categories	Physiognomic Perception Test, Weighted maturity score

Table 1
(Continued)

Ego-Control			
Age Three	Age Four	Age Five	Age Seven
Level of Aspiration, mean estimate over four trials	Level of Aspiration, mean estimate over four trials		
Lowenfeld Mosaic Test - Planfulness (reflected)		Lowenfeld Mosaic Test - Planfulness (reflected)	
Curiosity Box - Initiation Time (reflected)		Concept Evaluation Test - Yes score	Concept Evaluation Test - Yes score
Ego-Resiliency			
Age Three	Age Four	Age Five	Age Seven
Barrier Tasks - Number of Alternative solutions, composited over Barrier Door and Block Tower	Barrier Tasks - Number of Alternative solutions, composited over Barrier Drawer and Block Box tasks	Barrier Task - Number of Alternative solutions in Puzzle task	Peabody Picture Vocabulary Test, Time/Difficulty Correlation

Table 1
(Continued)

Age Three	Ego-Resiliency	Age Five	Age Seven
	Concept Evaluation Test - Consistency of standards score	Concept Evaluation Test - Consistency of standards score	Concept Evaluation Test - Consistency of standards score
	Alternative Hypotheses Task - Trial at which criterion for recognition of pattern change achieved (reflected)		
		Anticipation of Consequences - Total score	Anticipation of Consequences - Total score

Table 2

CCQ-items Significantly Associated with the Four
Ego-Control/Ego-resiliency Conjunctions

Resilient Undercontroller	Resilient Overcontroller
Energetic, active	Compliant
Curious, exploring	Calm, relaxed
Recoups, resilient	Empathic
Interesting, arresting	
Brittle Undercontroller	Brittle Overcontroller
Restless, fidgety	Inhibited, constricted
Undercontrolling of impulse	Worrying, anxious
Externalizing, vulnerable	Intolerant of ambiguity
Brittle, narrow margin of integration	Rigidly repetitive under stress
Manipulative	Interpersonally reserved
	Withdraws under stress
	Manifests inappropriate affect
	Manifests behavioral mannerisms

A Longitudinal Study of Children
from Kindergarten into the Adult years*

Philip E. Kraus

Professor Emeritus, Hunter College of the City University

Distinguished Professor, Touro College, New York

*This report is a condensation of Yesterday's Children: A Longitudinal Study of Children from Kindergarten into the Adult Years (New York: Wiley-Interscience, 1973).

A Longitudinal Study of Children
from Kindergarten into the Adult years

Philip E. Kraus

Professor Emeritus, Hunter College of the City University

Distinguished Professor, Touro College, New York

This is a twenty-year longitudinal study, begun in 1953 and concluded in 1973, of a previously unstudied population which typifies the vast majority of the children found in America's public schools. The study was authorized by the New York City Board of Education to study the effects of the school experience upon children typical of those found in the city's schools.

The study was designed to answer questions such as the following:

- Are there discernible patterns in children's learning processes?
- How early does the level of achievement become fixed?
- How do current practices of retention and acceleration affect children?
- How early is the promise of giftedness identified?
- To what extent do traumatic experiences affect learning and adjustment?
- How early are serious symptoms of maladjustment displayed?
- What is the significance of "separation-from-parent" anxiety and school phobia when starting school?
- What is the effect of mobility upon learning and adjustment?

In answering these and similar questions it was hoped that data might be collected which would provide perspective to look back, to identify symptoms of disturbance, and also to note encouraging predictors of achievement and success.

Actually, these objectives did not measure "the effects of the school experience" for we had no comparable group which went without schooling. Even after 25 years, we can only wonder which of the successes of our subjects can be attributed to the school and which of the failures can be laid at its feet. What did evolve was a developmental study of children who were under constant observation along with periodic reactions of parents and teachers.

The Sample

The research population consisted of all the children who entered the kindergartens of two New York City schools in September, 1953.

Although it is recognized that the results of a study of children in only two schools may not be applicable to all New York City school children, these schools were, nevertheless, sufficiently different in character and location so that they represented a picture of two broad segments of the school population.

One school in the study is located in the borough of Brooklyn, where it drew most of its children from a low middle-class housing project. Parents here were white-collar workers, skilled laborers, civil service employees, and a few businessmen. Most of the children in this school went home for lunch and, as was to be expected, only one or two qualified for inclusion in the free lunch program. This school is referred to here as the Jefferson School. All of its children moved on to the same junior high school.

The other school is located in the upper part of Manhattan on the fringe of Harlem. The occupations of the parents of children in this school ranged from unskilled labor to varying degrees of skilled labor. Although the community was not classified among those on the lowest socioeconomic level, the school did draw most of its children from depressed areas. During their seven years in the elementary school, one-third of the approximately 1500 children were eligible for free lunch. Approximately 85 percent of the kindergarten and first-grade class was black; this same ratio continued from 1953 through 1963 when the children completed the junior high school grades. In an effort to create more integrated schools, these children had been dispersed among three junior high schools at the end of the sixth grade. For purposes of this study, this school is referred to as the Washington School.

Of the 274 children with whom the study began, 115 attended the Jefferson School, and 159 were enrolled in the mostly black Washington School. It would have been unrealistic to expect that the 274 children in the first grade and the additional 20 who were admitted in the second grade would all remain in their original elementary schools through the sixth grade and would be available for study through the ninth grade. In fact, the selection of the Washington School population for intensive study was questioned by those who knew of the school's high mobility rate.

It was therefore very gratifying to find that 85 (70%) of the original 122 children in the Jefferson School remained to complete the sixth grade, as did 80 of the 172 (47%) in the Washington School. Moreover, 80 of the former and 68 of the latter completed the ninth grade in junior high schools accessible to the project staff. Longitudinal data for the elementary school years was thus available for the 165 children who formed our basic group, while 148 of these were studied through the ninth grade. Data were recorded for 85 of our children who were graduated from high school, although only 64 responded to our questionnaire at that time. It must be noted that computing longitudinal data for 148 children for nine years, and for 85 children for 12 years, is a difficult endeavor. Furthermore, there were several hundred children who were not part of the basic population who entered the two schools after the second grade and for whom data were also gathered. These data completely supported the conclusions drawn from the basic population material.

After the subjects had reached the ninth grade, all funding stopped. Contacts with the subjects for the next ten years were carried on solely by the writer, aided on several occasions by small foundation grants for secretarial services.

The three years that children in the study attended high school were difficult research years not only because of the complete absence of staff but also because of the dispersal of the children. When the Jefferson School children left their junior high school, they enrolled in 16 different New York City high schools while the Washington School children, because of their location in a different borough of the city, began their tenth grade in 22 different high schools. (There were 86 high schools in New York City at that time.) Contacts from then on became limited to mailed questionnaires, telephone calls, and occasional visits to the few schools having a sizable number of these children.

The resulting attrition of subjects in the study was inevitable. Some did not answer requests for information, some were absent so much that the schools could give little information about their performance

and progress, and others dropped out of school completely. Efforts to trace them were either rebuffed or unsuccessful. Drop-outs in the Jefferson School population (white) were almost all children who had moved away and consisted of achievers as well as poor students. In the Washington School population (black) mobility was much greater and the high school drop-outs were some achievers from upward mobile families, but mostly poor students or serious behavior problems. Nevertheless, a questionnaire mailed at their completion of the twelfth grade brought responses from 16 Washington School black subjects and from 48 Jefferson School white subjects. These, of course, were the children who had remained through graduation from high school. In general, children who were successful, especially from the Jefferson school, welcomed our inquiries, and up to this writing some spontaneously notify us of their progress, their marriages, and other changes in their personal lives.

By coincidence, four girls of the Jefferson School population and one boy from the Washington School subsequently enrolled in Hunter College, with which the writer was affiliated, thus providing frequent personal contacts and consultation, and many delightful nostalgic conversations about their childhood. Showing them their childhood writings and drawings was invariably a richly rewarding experience.

Most of the material described in the study is based on group data that were obtainable only from the kindergarten through the ninth grade, or up to the age of 14. Findings beyond that time are included where pertinent in the Summary of Findings.

The research population progressed through the grades in accordance with the following calendar:

Elementary School

1953-1954	Kindergarten
1954-1955	First grade
1955-1956	Second grade
1956-1957	Third grade
1957-1958	Fourth grade
1958-1959	Fifth grade
1959-1960	Sixth grade

Junior High School

1960-1961	Seventh grade
1961-1962	Eighth grade
1962-1963	Ninth grade*

Senior High School

1963-1964	Tenth grade
1964-1965	Eleventh grade
1965-1966	Twelfth grade

*Those who had been placed in special progress (SP) classes completed the ninth grade in 1962.

At the end of the first grade, children in the Jefferson School achieved a mean IQ score of 112.46 and those in the Washington School, a mean IQ score of 98.59. These scores remained quite consistent in subsequent testing in the third, sixth, and eighth grades.

The Instruments

As in most studies of this kind, the research staff experienced much anxiety that data that had not been anticipated might subsequently be needed. Because of this fear, there was a tendency to leave no aspect of growth unrecorded, even though it was acknowledged that some data elements might not be needed.

Standardized tests were available for measuring achievement and intelligence. The measurement of behavior and personality traits was much

more difficult. Many published scales were examined, but few were applicable to our population or warranted the administration time required. Original scales and questionnaires were therefore devised which were thought to be simpler and which subsequently provided significant information.

In selecting tests for measuring achievement, an attempt was made to use, wherever possible, those that were being administered on a city-wide basis. In the grades not included in the general testing program, supplementary achievement tests were administered by the project staff. The following standardized tests were used:

For Measuring Achievement

Reading

The New York Readiness Test (Grade 1)
New York Test of Growth in Reading (Grade 2)
Gates Primary Reading Test-Type (Grade 2)
Metropolitan Achievement Tests-Primary, Elementary,
Intermediate and Advanced (Grades 3-9)
Stanford Achievement Test-Intermediate (Grade 5)

Mathematics

New York Inventory of Mathematical Concepts (Grades 1-5)
Metropolitan Achievement Tests-Intermediate (Grade 6)

For Measuring Intelligence

Pintner-Cunningham Primary Test (Grade 1)
Otis Quick Scoring Mental Ability Tests-Alpha (Grade 3)
Otis Quick Scoring Mental Ability Tests-Beta (Grade 6)
Pintner General Ability Test (Grade 9)
Wechsler Intelligence Scale for Children
Revised Stanford-Binet Intelligence Scale
Henmon-Nelson Tests of Mental Ability

It was hoped that during the next few years we could obtain individual IQ scores for each of our children. This proved to be an illusion and although local colleges were most cooperative, less than half the children could be examined. The Henmon-Nelson Test was administered only to children whose teachers had recommended them for admission to Special Progress classes in junior high school but whose IQ scores were below the required minimum.

For Assessing Behavior and Personality

Haggerty-Olson-Wickman Behavior Rating Schedules
California Test of Personality
Ohio Social Acceptance Scale

The Haggerty-Olson-Wickman Behavior Schedule was used with only a few classes but was abandoned as being too cumbersome and too time-consuming for teachers. Another research project under way in the Jefferson School utilized the Ohio Social Acceptance Scale, the results of which were given to us. All of these scales were used only once and furnished little information that was not already becoming available through our own original scales and questionnaires which were designed to guide the project staff and teachers in their observations of children and to obtain information in their study of child growth. These instruments were tested and refined to meet the following criteria:

They should supply information about each child that could be obtained in no other way.
They must in no way threaten the teacher.
They must impose little or no clerical burden upon the teacher.
They must be easy to use.

The following rating scales were thus devised:

Teacher's Estimate of Reading Development--to determine each child's achievement and interest in reading, as judged by the teacher.
Teacher's Estimate of Intelligence--to determine each child's level of potential as estimated by the teacher.
Attitude Toward Classroom Activity--to measure children's reactions to, and participation in classroom activity.
General Adjustment Rating Scale--to note the extent of behavioral adjustment of each child, and to record teachers' recommendations for special services where indicated.

These are described more fully in the pages following. In addition to scores from standardized intelligence and achievement tests, additional information was gathered.

Information Obtained from Children

A. Samples of drawings and paintings--Twice each year, children drew or painted for the Project Assistants a picture of "myself," of "a person" of "my family," and of "anything." For some children, the self-portraits and family portraits provided interesting supplementary personality data. This material was also used for comparison with the drawings of children two decades later.

- B. Samples of handwriting--It was hoped that these would provide a record of the development of muscular coordination. Children were asked to use their best handwriting in writing on the following topics:

Grade 3--If I had \$10
Grade 4--The Most Fun I Ever Had
Grade 5--If I Could Fly
Grade 6--If I Were a Teacher

In Grades 1 and 2, sentences were copied from the chalkboard. Handwriting itself proved to have no significance and in no case were there any radical unexpected changes. With all children, fine muscular coordination improved with physical maturity.

- C. Children's questionnaires--These were devised to obtain reactions and attitudes of children toward various phases of their school environment as well as their personal lives. Until reading and writing ceased to be a problem, the questions were read orally by the project staff members in individual conferences with children and their answers were recorded. For the most part, the same questions were used on each subsequent questionnaire.
- D. Sociometric measures--These measures were based upon responses to the following questions:

Grade 1--Whom would you like to sit next to?
Grade 2--Whom would you like to sit next to?
Grade 3--Whom would you like for your partner on class trips?
Grade 4--If you could choose anyone in the class, a boy or a girl, as a partner for a class trip, whom would you choose?
Grade 5--If you had your choice, which child of your class would you like to have with you in your class next September? If you couldn't get this one, whom would you choose second? Third?
In the fifth grade the Ohio Social Acceptance Scale was used along with the sociogram. This scale endeavors to determine not only the degree to which a child is accepted by classmates but the extent to which he or she reaches out and seeks their friendship.

Grade 6--If you could choose, which boy or girl in your present class would you like to have in your class in junior high school? If you couldn't have this child, whom would you choose second? third? If you are eager to add more names you may do so in order of your choice.

Information Obtained from Teachers

After orientation conferences with Project Assistants, teachers provided the following data:

- A. Teacher judgment of readiness in areas listed on the New York Reading Readiness Test--This scale is part of the New York Reading Readiness Test and was used only in the first grade before formal reading instruction had begun.
- B. Teacher's Estimate of Reading Development--This scale proved to be less reliable than the achievement tests. Teachers frequently estimated the reading levels of their pupils in terms of the basic reader which they were using. Where the level of the reader used was limited to that of the grade, the true ceiling of children's abilities was rarely identified. Average or poor readers were more often correctly identified than were superior readers. Nevertheless, this scale frequently provided insights into the standards that teachers used in evaluating pupil progress.
- C. Teacher's Estimate of Intelligence--This scale was, of course, of most value before the first intelligence test had been given in the first grade. However, after IQ scores were available, many teachers were still able to sense abilities that were not supported by IQ scores, whose validity they questioned. This independence of teachers in judging intelligence gave clues to disparities between performance and IQs and suggested children who warranted closer observation and study.
- D. Attitude Toward Classroom Activity--This scale indicated, for the most part, the close relationship between ability and interest.
- E. General Adjustment Rating Scale--This was a five-point scale on which teachers rated each child's behavioral adjustment. This scale consistently proved to be of great value.

F. Lists of children with adjustment problems--These were obtained twice a year, once in November and once in March. One of the purposes here was to provide teachers with spaced opportunities to identify their problem children and to discuss them with the project staff.

G. Lists of gifted and talented children--These were obtained each year in an effort to determine how early and how frequently the promise of superior potential was identified.

For certain children the following data were gathered:

A. Anecdotal material--Anecdotal records were at first kept for a random sampling of 25 children in each of the two schools. These children had been selected in the kindergarten. However, after the second year of the study, it was felt that time could more profitably be spent in keeping detailed anecdotal records for all those children whose behavior or learning patterns warranted it. Such records were kept throughout the period of study.

B. Notations for deviates--All children were regularly observed but notations were made for those who showed deviations in physical, emotional, social, or intellectual growth. For example, continuous records were kept for the tallest children, the shortest children, left-handed children, and all who showed noticeable deviations from the norm.

C. Other data--Other data that were gathered included copies of school records, health and test cards, and pertinent health information that was supplied by nurses and doctors.

Notes concerning the socioeconomic background of each of the two school communities, their resources, and changing patterns were recorded.

The excellent relationship that developed between project staff members and the schools, the children, and parents facilitated the gathering of information and helped immeasurably in obtaining enthusiastic cooperation. Two staff members (or Project Assistants as they became known) were assigned to the larger Washington School and one to Jefferson School, where they spent approximately three days a week. As children progressed each year from grade to grade, and teacher to teacher, Project Assistants moved with them. Great care was taken never to interfere with school procedures or school placement.

Our children are now grown men and women. Among them are teachers and students of law, dentistry, and medicine; technicians and nurses, secretaries, clerks, salesmen and budding junior executives, housewives and mothers, mechanics and skilled laborers. Among them, also, are the failures in life--the drifters and the adult delinquents. It is futile to speculate which aspects of their lives were shaped by school experiences. Certainly the organizational changes that occurred during their schooling seemed to have no measurable effect. How do we really know when the school has succeeded and for what it can be held accountable?

Our findings do not support DeLone's (1979) conclusions concerning the high correlation between economic background and subsequent occupational success. Given children of equal ability, the most significant predictor of adult achievement was parental attitudes toward school and toward education. This was particularly true among black families.

References

- Cloward, R. & Jones J. Social class: Educational attitudes and participation. In A.H. Passow (Ed.), Education in depressed areas. New York: Teachers College Press, 1963.
- Delone, R. Small futures. New York: Harcourt Brace Jovanovich, 1979.
- Harris, A.J. How to increase reading ability. New York: Longmans, Green and Co., 1947.
- Harris, A.J. & Lovinger, R.J. Longitudinal measures of the intelligence of disadvantaged Negro adolescents. School Review, 1968, 76, 60-66.
- Jose, J. & Cody, J.J. Teacher-pupil interreaction as it relates to attempted changes in teacher expectancy of academic ability and achievement. American Education Research Journal, 1971, 8 (1), 39-49.
- Kagan, J. & Moss, H.A. Birth to maturity. New York: John Wiley and Sons, 1962.
- MacFarlane, J.W. Some findings from a ten-year guidance research program. Progressive Education, 1938, 15, 529-535.
- Rosenthal, R. & Jacobson, L. Pygmalion in the classroom. New York: Holt, Rinehart and Winston, 1968.
- Silberman, C.E. Crises in the classroom. New York: Random House, 1970.
- Weber, L. The English infant school and informal education. Englewood Cliffs, New Jersey: Prentice-Hall, 1971.
- Weiner, M. & Murray, W. Another look at the culturally deprived and their levels of aspiration. Journal of Educational Sociology, 1963, 36, 319-321.

The Child's Political World:

A Longitudinal Perspective*

James Lare (Occidental College)

Stanley W. Moore (Pepperdine University, Malibu)

Kenneth A. Wagner (California State University, Los Angeles)

*The authors wish to express their appreciation for financial support from the Cleveland Foundation and the Spencer Foundation of Chicago, as well as the assistance provided by Joan Wagner, Jeri Steinberh, Robert Cendejas, and Carla Ryhal.

The Child's Political World:
A Longitudinal Perspective

James Lare (Occidental College)

Stanley W. Moore (Pepperdine University, Malibu)

Kenneth A. Wagner (California State University, Los Angeles)

The present study is designed to fill four significant gaps in our knowledge of the political socialization process. First, it generates much-needed empirical data on young school-aged children. In the first year those in our panel were at least two years younger, and in some cases four or five years younger, than those questioned in any of the major studies of early political socialization. The children were interviewed individually, and open-ended questions were used extensively to encourage complete and clear responses. Second, this study is longitudinal; the same children were interviewed near the end of each school year from kindergarten through fourth grade. Third, since the children were interviewed during the years 1974 through 1978, and because many of the questions and themes used in earlier studies were included in the interview schedules, these data provide bases for comparison with the findings reported in the pioneering socialization studies of the early 1960's. Fourth, the usefulness of cognitive-developmental theory to explain the growth of civic awareness among these young children is systematically evaluated.

We began to explore the political world of the child at a critical time of life. In Western society, up until a child enrolls in school, he or she is normally exposed to a limited group of parental and other adult authorities. This often results in the child's developing an absolute view concerning laws and authorities (Hyman, 1959). When the child enters school, however, divergent viewpoints of other adult authority figures are introduced. This broadened exposure leads to changes in the child's posture toward rules, laws and authority. Whether these early orientations to the political realm are critical or only marginally relevant in determining later adult attitudes has been a matter of debate. Adequate longitudinal evidence is not currently available to determine the strength of these relationships. But the findings of developmental psychology indicate that

attitudes and learned behavior appear to be a product of interaction with an individual's environment and that early influences on the developing personality are likely to affect both the style and the content of later interactions (Ginsburg & Opper, 1969; Piaget & Inhelder, 1969; Piaget, 1965). Consequently, these early cognitive and affective orientations toward political authorities, processes, and issues may well be influential throughout life.

Procedural Questions

To answer a number of these procedural questions a pilot study was conducted by on the researchers in a non-targeted school district. A day of interviewing by this researcher and five university students demonstrated that the age and sex of the interviewer did not significantly affect the responses of kindergarten children. There was some indication, however, that informal attire and the dropping of academic titles such as "professor" or "doctor" helped achieve the kind of rapport essential for satisfactory interviews. It was also found to be desirable to remove the interviewee from the classroom setting so he or she would not be distracted by curious peers observing the interview process. Further, the researchers decided to conduct most of the interviews for this five year longitudinal study themselves because the use of multiple and changing interviewers would increase the likelihood of errors. The possibility of using the teachers to conduct the interviews was rejected because of the great imposition on their time and the likelihood that their teaching would be influenced by their knowledge of questions to be asked of their pupils in subsequent years. Teaching to prepare students for future interview questions would obviously affect the rate of the children's growth in civic awareness as measured by our interview schedule. It was also assumed that each researcher would take care to establish rapport with the children prior to beginning the kindergarten interviews. If these efforts to establish rapport were successful, the children would accept, and indeed welcome, the return of the increasingly familiar interviewer. For example, one researcher spent three days in each kindergarten classroom before interviewing any of the children. Efforts such as this have been rewarded with good interviewer-interviewee rapport as evidenced by the relaxed demeanor of the children, their friendliness in greeting the interviewers each year, and their continuing responsiveness to the open-ended questions in the interview schedule.

Questionnaire Design and Administration

The interviews were conducted annually during the last four to eight weeks of the school year. A major constraint was the need to accommodate ourselves to each child's attention span. In kindergarten the children were asked 57 questions in one interview session lasting 10 to 15 minutes. Starting in the first grade, the interviews were usually conducted in two parts because the number of questions had increased to 91. By fourth grade the interview schedule contained 176 questions and required 35 to

45 minutes to administer. Before conducting the third grade interviews an experiment with pencil and paper responses to a limited number of fixed choice items was attempted with third graders who were not in the panel. Because we found these children were more likely to circle the "Don't know" (DK) response when it appeared on the paper in front of them than was the case when they were asked "yes" or "no" questions orally, we elected to continue utilizing oral interviews exclusively. The tendency for respondents to "overselect" the "DK" response when it is available to be circled on a written questionnaire was further shown to exist when some of the fifth grade children answered a series of questions from a written questionnaire while the others responded to an oral administration of the same questions. Twelve of the items listed a "DK" response alternative on the written questionnaire. Those children responding to the written questionnaire selected the "DK" response 6.8 percent more per question than those responding orally to the same question. By contract there was an average difference of only .45 of one percent greater use of "DK" by those responding to 35 questions when the "DK" alternative was not listed on the written question when compared to the oral group.

Related to our concern about the children's limited attention span was the further need to avoid demoralizing these young children by compelling them to respond "I don't know" to a large number of questions. For that reason kindergartners and first graders were not asked all of the questions that might appropriately be asked of third and fourth grade children. Even in the relatively shorter kindergarten and first grade interview schedules, special care was taken in the sequencing of the questions to offer the children regular opportunities to give some kind of affirmative response. In addition, there was an increasingly heavy reliance on open-ended as opposed to fixed-choice questions in each year of the study. Open-ended questions are defined as those eliciting more than a simple "yes" or "no" response ("Do we have a king in our country?") or implying a single correct answer ("Who is the President now?"). A very few questions were either deleted or modified once they were incorporated into the questionnaire.

The wording of questions was another major concern. We wanted to test what the children really knew and felt about the political process, not the scope and limits of their vocabulary. In developing a vocabulary for the interview schedule we wished to avoid ambiguity, loaded words, and, above all, misunderstanding between researchers and respondents. We faced a typical research tradeoff: if we used only language that the children in kindergarten could be expected to understand, our questions would be limited in subject matter, yet if we waited until most children understood more complex political terminology, we would not be able to ascertain the trend as the first few children, and then more, came to hear about and comprehend a particular political phenomenon. In addition, if we were to use one term intelligible to the kindergartner, and a different term when the children were older, we would create a comparability problem. The Educational Developmental Laboratory's A Revised Core

Vocabulary: A Basic Vocabulary for Grade 1-8 was used in selecting key terms (Taylor, Frackenpohl & White, 1969). For example, "How is the country ruled?" would be too difficult for many primary school children, much less for five or six year olds. Therefore, in our first question we asked, "Who is the boss of our country?" "Boss" has the advantage of being in the vocabulary of most kindergartners whereas alternatives like "leader" or "ruler" might not. We also intentionally introduced terms, such as "politician," that would not be in the kindergartner's vocabulary. One reason for using "advanced vocabulary" was to check on the child's truthfulness in telling whether or not he or she knew what the word meant, but more importantly it enabled us to determine time and circumstances in which understanding of complex political phenomena began to emerge.

What is the effect of repeatedly interviewing the same children year after year? Would they be sensitized to government and politics and become more knowledgeable simply because of the stimulation of being repeatedly interviewed? A longitudinal design controls for most internal invalidities, but it is vulnerable to the problem of sensitization, the possibility that repeated testing will affect the results (Campbell & Stanley, 1963). Researchers want to be able to generalize their findings beyond their sample, indeed to the entire target population. In order to check on the possibility that children in the panel were being sensitized and thus no longer typical of primary school children, the researchers incorporated three control features into the research design. The responses of first grade children who were members of the kindergarten sample were compared with the responses of first grade classmates who were not interviewed during the kindergarten year. Secondly, in third grade the researchers again made a check against the possibility of testing sensitization (while simultaneously checking the effects of a written questionnaire vis-a-vis a verbal one) by presenting a 21-item subsection of the verbal interview schedule as a written questionnaire to seven intact third grade classrooms. This population of 167 children included 53 of the panel. Thirdly, more than forty percent of our panel in fifth grade responded to a 54-item written segment of the questionnaire. This section of the questionnaire was also administered to 154 of their classmates who were not members of the panel.

Analysis indicated that responses of the first grade panel and non-panel children were similar; as third and fifth graders, the children in the panel were still not sensitized to the point of becoming more knowledgeable in responding to cognitive questions than third graders and fifth graders not in the panel. There is evidence, however, that the third and fifth grade children in our panel were likely to view government as more pervasive than non-sample third and fifth graders.

Categorization of Questionnaire Content

Most of the political objects the children were asked about may be classified according to the triad of "Eastonian categories": the national political community, the regime, and the authorities of government (Easton, 1971; Easton & Hess, 1962). In addition, the interview schedule raises questions similar to those used in earlier studies concerning political geography (e.g., "What city do you live in?") (Greenstein, 1965; Jahoda, 1963) and public vs. private sector roles (e.g., "Does a mailman, milkman, policeman, etc., work for the government?") (Easton & Dennis, 1967). A category of questions unique to this study concerns a series of contemporary public policy issues: Vietnam, Watergate, crime, the energy crises, inflation and unemployment. Another way of classifying these same interview questions is according to those which are cognitive, affective, evaluative, and self-reporting. "What is inflation?" is a cognitive query; "Is inflation good or bad?" calls for an affective response; "Why is it good or bad?" calls for an evaluative judgment; and whether the child has ever heard the word 'inflation' asks for a self-report.

Description of the Panel

The panel of 243 children used in this study may be described as an "availability sample." Each of the principle researchers approached the school authorities in his local district and requested permission to interview from one to three complete kindergarten classes during May and June of 1974. The conditions for interviewing the children varied from district to district. One school board required that the parents return a signed permission slip before a child could be interviewed, and this only after one member of the board publicly vouched for the researcher; another simply asked that a notice be sent home with each child and parents who did not wish their child interviewed could contact either the school or the researcher with that request. In one district the approval of the Committee on Research was all that was necessary.

The children comprising the original kindergarten sample resided in five school districts located in four communities separated by as much as sixty miles; two inner suburbs within ten miles of Los Angeles' civic center and two outlying suburbs located almost fifty miles from the central city. The first interviews with 218 kindergartners were conducted in late May and early June of 1974 (Moore, et al., 1976). One year later 186 (85.3 percent) of these children were re-interviewed and 135 of their first grade classmates were questioned for the first time. In the spring of 1978, when most of these children were completing fourth grade, 149 of the 218 interviewed first in kindergarten and 94 of the 134 initially interviewed in first grade remained to provide the panel of 243 children on which this longitudinal analysis is based. They represent 69 percent of all the children interviewed during the course of this study. The demographic characteristics of the 110 children who were lost from the sample closely resemble those of the 243 youngsters who comprise the total used in this study (see Table 1).

The original panel of 218 pupils consisted of six intact kindergarten classes and portions of three other classes in the district requiring permission slips. Those interviewed did not usually include all kindergarten children in their respective schools: two researchers selected two of the four available classes--one used the two morning sessions from the school while the other used one morning and one afternoon class from the same school. A third interviewed all children in two kindergarten classes at one school and the fourth was permitted to interview one-third of the three kindergarten classes at one school. Assignment to these classes at each school was random in that no criterion, other than equal distribution of sexes, was employed. Given the random assignment of pupils to the kindergarten classes, it was decided to interview all the children in the first grade classes at the schools used for the kindergarten interviews. The researchers assumed that the cognitive and affective responses of the newly interviewed and the re-interviewed first grade pupils would be similar. This assumption was confirmed when analysis showed that there were no significant differences in the responses of the children interviewed for the first time in 1975 when compared with those given by pupils in the original sample for any of the variables used in this analysis. The addition of the new students and the comparison of their responses with those provided by members of the original panel served two functions: first, it permitted enlarging the panel without impairing validity, and second, the analysis provided a check on the test-retest influence of the original questioning.

Table I

A Comparison of Panel Members and Dropouts on Three Demographic Characteristics

		Panel Members (N=243)	Lost from Panel (N=110)
Sex	Male	50.2	55.4
	Female	49.8	44.5
Race	Anglo	81.5	87.3
	Non-Anglo	18.5	11.8
	NA	0.0	0.9
Sibling Order	Only Child	5.8	12.7
	Oldest Child	24.3	25.5
	Middle Child	25.1	25.5
	Youngest Child	44.9	34.5
	NA	0.0	1.8

The extent of school mobility on the part of these 243 children seems quite consistent with the general pattern for Southern California. During the five year period of this study the children dispersed from the original 5 schools to 47. One-third of the children have attended more than one school during their first five years of formal education, with 10 percent moving from a public to a private school, the majority of these transferring to parochial schools that did not offer kindergarten programs. The eighteen children who have attended three to five schools in five years are usually youngsters from broken homes or else upwardly frequent mobile families who move with each change of employment or rise in family income.

Implications for Education, Socialization Theory and for Democracy

The lessons of this research for the primary school educator are both clear and noteworthy. There is the implication in our findings that it would enhance the rate and the quality of political learning among primary school children if teachers were sensitized to the threshold cognitions which are open new vistas of political awareness for a young child. If a youngster, during the kindergarten and first grade years, were encouraged to identify the most salient political authorities (e.g., the policeman, the judge, and the President) and with the government and its functions, these latter abstractions would become more meaningful and subsequently the child could more readily associate other public authorities with the realm of government.

The Western democratic tradition includes the notion of a separation of political power and the concept of checks and balances between separated powers. It would therefore seem desirable to have children from a rather young age becoming increasingly sensitive to the importance of legislative and judicial as well as executive authority. The children in our panel, however, persisted even through the fourth grade, in giving predominant attention and importance to the President. Eighty-six percent of the fourth graders perceived the President as "the boss of the country"; 79 percent believe he does the most to run our country; and 38 percent think he alone makes the laws. By contrast, not one child thought that Congress was "the boss of the country"; only one child (in fourth grade) believed that Congress does the most to run our country; and six children (2.5 percent) said that Congress alone makes the laws. Just 17 percent of the fourth grade children acknowledge a division of labor in the legislative process by suggesting that the President must cooperate with "others" in making the laws. This early and persistent concept of Presidential supremacy and the absence of a notion of shared political power prevent these young children from developing an appreciation for the complex bargaining and brokerage roles that both the President and Congress are frequently called upon to play.

Students of political socialization, as well as others utilizing a design similar to the one used in this investigation, should note that a carefully planned administration of the questionnaire on an annual basis does not noticeably sensitive the children (at least during the first six years of school) to provide responses unlike those given by non-panel classmates. The "don't know" responses, however, are inflated when they are available as an option to be circled on a written questionnaire when compared with an oral questioning where no options are suggested. An oral administration of the questionnaire also seems to elicit responses that are more honest, in that the children responding orally are only about half as likely as panel and non-panel children who had the 59 question subsection administered in written form, to claim in self-report questions that they often read the front page news or a weekly news magazine. Finally, where open-ended responses are requested the answers are more complete when the questions are asked orally.

There are a number of findings in this study which have important implications for the future stability and vitality of our modern democratic society. Most striking perhaps is the significantly lower level of partisan awareness we found among primary school children in the mid-1970's when compared with the levels found fifteen or twenty years ago. A mere eight percent of our fourth graders indicated that they preferred one party over another in 1978, and 16 percent could tell us President Carter's affiliation. In Greenstein's (1965) pioneering study, when children were given party names on written questionnaires, more than 60 percent of fourth graders indicated a preference for one of the two major parties. If privatistic and transitory orientations toward candidates who are promoted so professionally by the media have replaced the more stable and easily inherited party identification of an earlier generation, the new pattern may diminish stability in the system by inducing more rapid shift from one "packaged" politician to another at both the national and the state levels. The image of politics created by extensive use of the electronic media has diminished the influence of party machinery and indeed the appeal of the party label itself.

There is some indication from the responses of our panel that there is a strong subject as opposed to participant orientation among the primary school children of the 1970's. They do not hesitate to assert that obeying the laws is more important than voting in an election; they feel more positively toward the policeman than toward any other governmental authority figure; and they are increasingly less willing, as they advance through the primary grades, to tell us "what they would like to do for other people if they were the boss of our whole country." Particularly striking in this pattern of subject orientation is the significantly great visibility of the judicial process in the eyes of young children as compared with the relative invisibility of the legislative process. The formalized, constraining and non-partisan character of the judicial process contrasts with the less formal and highly partisan behavior that often prevails among both elected legislators and their activist supporters.

Encouraging implications for the future of our democratic system lie in the increasing evidence that young children appreciate the fallibility of governmental institutions and authorities. Significantly higher percentages of children in the 1970's were prepared to acknowledge that the government and the President make mistakes than was the case a decade or two earlier. Yet despite this awareness of fallibility in public institutions, these children, who were coming to their political awareness during the Watergate era, still want to vote and they also understand why we have laws and why they must be enforced. Generally speaking, they appear to be at a higher stage of cognitive and affective development with respect to law than they are with most other facets of governmental and political activity. If a tolerance for ambiguity and even paradox is a requisite for good citizenship in a democratic society, then the young child who places obedience to a fallible government above voting and yet insists that he wants to vote may be unconsciously developing a political orientation that is appropriate for a complex government and conflict-ridden world.

References

- Andrain, C. Children and civic awareness. Columbus, Ohio: Merrill, 1971.
- Campbell, D. & Stanley, J. Experimental and quasi-experimental designs for research. Chicago: Rand McNally and Co., 1963.
- Connell, R.W. The child's construction of politics. Melbourne: Melbourne University Press, 1971.
- Easton, D. The political system. New York: Alfred A. Knopf, 1971.
- Easton, D., & Dennis, J. Children in the political system. New York: McGraw-Hill, 1969.
- Easton, D., & Hess, R.D. The child's political world. Midwest Journal of Political Science, 1962, 2, 229-246.
- Garcia, F.C. The chicano political experience. North Scituate, Mass.: Duxbury Press, 1977.
- Garcia, F.C. The political socialization of chicano children. New York: Praeger Publishers, 1973.
- Ginsburg, H. & Opper, S. Piaget's theory of intellectual development: An introduction. Englewood Cliffs: Prentice Hall, 1969.
- Goodman, M.E. Emergent citizenship: A study of relevant values in four-year olds. Childhood Education, 1959, 35, 248-251.
- Greenstein, F. Children and politics. New Haven: Yale University Press, 1965.
- Hess, R. & Torney, J. The development of political attitudes in children. Garden City, New York: Doubleday, 1967.
- Hyman, H. Political socialization. New York: Free Press, 1959.
- Ippolito, D., Walker, T., & Kolson, K. Public opinion and responsible democracy. Englewood Cliffs: Prentice-Hall, 1976.
- Jahoda, G. The development of children's ideas about country and nationality, Part 1: The conceptual framework. British Journal of Educational Psychology, 1963, 33, 47-60.
- Jaros, D. Socialization to politics. New York: Praeger, 1973.
- Kohlberg, L. Moral stages and moralization. In T. Lickona (Ed.), Moral development and behavior. New York: Holt, Rinehart and Winston, 1976.
- Kohlberg, L. Stage and sequence: The cognitive developmental approach to socialization. In D.A. Goslin (Ed.), Handbook of socialization theory and research. Chicago: Rand McNally, 1969.
- Langton, K.P. Political socialization. New York: Oxford University Press, 1969.
- Moore, S.W., Wagner, K.A., Lare, J. & McHargue, D.S. II The civic awareness of five and six year olds. The Western Political Quarterly, 1976, 29(3), 410-424.
- Niemi, R. et al. The politics of future citizens. San Francisco: Jossey-Bass, 1974.
- Piaget, J. The moral judgement of the child. New York: Free Press, 1965.
- Piaget, J. & Inhelder, B. The psychology of the child. New York: Basic Books, 1969.
- Renshon, S. (Ed.) Handbook of political socialization. New York: Free Press, 1977.
- Runkel, P.J. & McGrath, J.E. Research on human behavior: A systematic guide to method. New York: Holt, Rinehart and Winston, Inc. 1972.
- Schwartz, S. Preschoolers and politics. In D. Schwartz and S. Schwartz (Eds.), New direction in political socialization. New York: Free Press, 1975.
- Sigel, R. & Brooks, M. Becoming critical about politics. In R. Niemi, et al., The politics of future citizens. San Francisco: Jossey-Bass, 1974.
- Stevens, A. The acquisition of participatory norms: The case of Japanese and Mexican-American children in a suburban environment. The Western Political Quarterly, 1975, 28, 281-295.
- Taylor, S.E., Frackenpohl, H. & White, C.E. A revised core vocabulary: A basic vocabulary for grade 1-8. Research and Information Bulletin No. 5 Revised Huntington, New York: Educational Developmental Laboratories, Inc. March, 1969.
- Vaillancout, P. & Niemi, R. Children's party choices. In R. Niemi, et al., The politics of future citizens. San Francisco: Jossey-Bass, 1974.
- Wagner, K.A., Lare, J., & Moore, S.W. Developing notions of law and morality in five to eight year olds. In The Proceedings of the Eighth Annual Piagetian Conference (1978), Los Angeles: University of Southern California, 1979.

Schooling of Young Children*

Doris R. Entwisle and Leslie Alec Hayduk**

The Johns Hopkins University

*This work was supported by OEG-3-71-0122, NIE-G-74-0029 and a Guggenheim Fellowship to Doris Entwisle

**Now at the University of Western Ontario

Schooling of Young Children

Doris R. Entwisle and Leslie Alec Hayduk

The Johns Hopkins University

Previous sociological research on educational attainment, with few exceptions, deals with students enrolled in secondary schools or college. The present research was undertaken to shed light on the social and psychological determinants of schooling for much younger students, specifically children in the first three grades of elementary school. Several lines of reasoning suggest that very early educational attainment is critical for later attainment and that the time has come for serious sociological study of students at the beginning of their educational careers.

First, developmental theory as well as common observation suggest that later achievement depends heavily upon achievement in the elementary school. Several bodies of research (Kraus, 1973; Husen, 1969; and others) suggest that by third grade children are launched into educational trajectories which largely govern their future educational and occupational attainments. Second, the social structure of elementary schools in the United States, with their self-contained classrooms and close neighborhood ties, probably accentuates any effects of social determinants on attainment because at that point in the life cycle, effects of school, neighborhood, and family are most consonant. Third, although differences among students attributable to the school they attend are judged to be almost trivial at the secondary level (see e.g., Coleman, 1966 and Mosteller and Moynihan, 1972), this by no means implies that schools have no effects, nor does it foreclose the possibility that there may be larger differential effects of schools earlier in students' lives. For example, Heyns' (1978) analysis indicates that schooling effects for sixth and seventh graders in Atlanta tend to offset differences in family background, and we (Entwisle and Hayduk, 1978) have observed that in two schools where

children's tested ability levels were not far apart at the start of first grade, by the end of first grade in one school children were reading far above grade level while in the other school they were already behind. There are, in fact, many good theoretical and practical reasons to support the study of the social determinants of elementary schooling.

The fundamental assumptions guiding our work are drawn from a broadly symbolic interactionist or Meadian perspective. The first few grades of school provide students with a social arena in which they learn the important skills of reading and arithmetic. In the process of learning they develop ideas about their own academic competence. We assume that the social structure of the school mediates the social feedback stimulating students to achieve. Simply put, achievement occurs in large part because students are responsive to evaluations by significant others (parents, peers, teachers) as well as to evaluations they make of themselves. Although psychological theories of learning uniformly conceive of an individual as learning in isolation (Estes, 1970), we believe that the social matrix in which the student functions is critical for both the cognitive and affective development of young children.

The basic strategy of this research was to select three elementary schools as exemplars of common types of schools in the United States--a white middle-class suburban school, an integrated lower-class urban school, and a black lower-class urban school--and then to study children enrolled in these schools as they passed through the first, second, and third grades. Achievement in reading and arithmetic were the two cognitive outcomes studied, and the children's marks in conduct and their ideas about their own competence in reading, arithmetic, and conduct were the affective outcomes studied. Multivariate models that include both social and individual-level variables were constructed to describe the process of early schooling, and then the adequacy of these models was tested using data from the three schools. Such an approach draws upon the tradition of sociological studies of secondary-school attainment but, with one partial exception (Heyns, 1976), the purpose of this study, and also its execution, are very different from previous sociological studies of schooling.

Variables and Their Measurement

As mentioned, for the elementary school child some of the most important consequences of schooling are the child's mastery of the basic skills, how well and how quickly he or she learns to read and to do simple arithmetic. Other consequences have to do with ideas children forge about their own competence. Eventual educational attainment, or even short-run attainment on the next rung of the educational ladder, may depend as much on young children's conceptions

of their own ability as on their absolute level of achievement because children who are self-confident and who believe they will succeed when they start school act in ways to make their beliefs come true.

To test these notions we proposed a set of cyclic structural equation models and collected data to test the models from several hundred children, starting from the time they begin first grade. As mentioned, the guiding conceptual scheme in formulating the models is that, as children mature and pass through educational institutions, schooling comes about as a consequence of social interaction between the child and the set of significant others. This interaction occurs in a definite time frame. We used this frame as a guideline in formulating the models and in collecting the data.

Two kinds of outcomes are considered, marks given by teachers (M) and children's academic expectations (E). Since marks in reading and arithmetic constitute the heart of the elementary curriculum, especially in the first three grades, children's marks in reading and arithmetic were taken as the two main cognitive outcomes of schooling. The child's conduct is conceptualized as one of the affective outcomes of schooling. Children are socialized to be attentive, to be responsive to peers and teachers, to enjoy schoolwork. These attitudes and dispositions are the general basis for assigning marks in conduct. Three further affective outcomes are children's expectations for their own performance in reading, arithmetic and conduct. These expectations, we assume, also are a consequence of schooling. The three expectation measures and the associated marks (reading, arithmetic, conduct) are the major variables we wish to explain.

To account for children's expectations and their marks, the models invoke the following variables: the child's race; the child's sex; IQ as measured on any one of several standardized tests; parent's estimate of the child's general ability to do schoolwork (ABIL); parent's specific expectations for the child's level of performance in reading, arithmetic, and conduct (PE); peer's ranking of the child's ability to read (PEER); and the number of days the child is absent during the school year (ABS).

A word is needed about teachers' marks vs. standardized test scores. Precisely because teachers are likely to be influenced by a child's appearance and demeanor and because teachers vary in the marking standards they apply, school systems routinely evaluate children's reading and arithmetic skills with standardized tests. In the models to be discussed in this chapter we included only teachers' marks. We demonstrate elsewhere (Entwisle and Hayduk, 1980), however, that children's marks in reading and arithmetic, as given by the teachers included in this study, did predict the children's performance on standardized tests of reading and arithmetic. Perhaps more important, the structural models function equally well and in a similar fashion to explain teachers' marks or standardized achievement. We concentrated on teachers' marks, however, because marks were available for all children in the study whereas standardized test scores were not.

A brief description of variables follows.

RACE. This variable was coded "0" for white, "1" for black. The few Orientals and Hispanics were classed as white.

SEX. This variable was coded "1" for boys, "2" for girls.

IQ. These were scores from a variety of tests given in the different schools at various times. (A summary of the nature of these tests and the time of their administration is given in Table 5.1, Entwisle and Hayduk, (1980).)

ABIL., Parent's Estimate of the Child's Ability to do Schoolwork. Parents of each child (usually the mother) were asked, "How do you rate your child in school ability compared with other children in this school?"

(Check one)

- () Among the Best
- () Above Average
- () Average
- () Below Average
- () Among the Poorest

The responses were coded from 1 (Among the Best) to 5 (Among the Poorest).

PE, Parent's Expectations. The parents of each child were presumed to have expectations for how well their child would perform in reading, arithmetic, and conduct. Parents were asked to fill out a questionnaire (or if necessary they were queried by an interviewer who filled out the questionnaire for them) indicating their expectations. Figure 1 reproduces part of the questionnaire that tapped parents' expectations. When necessary, interviewers interpreted for parents the marking standards used by the school and answered questions parents asked.

Interviewers whose race usually matched that of the parent gathered data from parents only once a year, in the fall, shortly before the first report card was issued. (See Figure 2.) Parents' interviews preceded children's interviews in any given semester. Parents' expectations were coded from "1" (high) to "4" (low).

E, Children's Expectations. Just before their earliest report cards were issued in first grade, children were queried about their own expectations for their forthcoming marks in reading, in arithmetic, and in conduct. Children were queried again just before the end of first grade, and also twice in the second and third grades. (See Figure 2.) Each child was presumed to have some notion of how well he or she would perform in each area. Using a large stylized plastic replica of a report card and cardboard squares containing large numerals, she explained to the

child what report cards are, what marks are, how marks are coded, and if necessary what "reading," "arithmetic" and "conduct" meant. She then asked each child to "play a game" in which the child tried to guess what marks the forthcoming report card should show in reading, arithmetic, and conduct, by placing the cardboard squares on the big "report card." Great care was taken that the children understood the task and understood the meaning of both "marks" and "report cards."

This kind of interview provided a separate measure of children's expectation in reading, arithmetic, and conduct (E_R, E_A, E_C) for each semester of the first three grades in school. Children's expectations were coded in the same manner as parents'--from "1" high or A, to "4" low or D.

Expectations were always procured shortly before the relevant report cards were issued. The large plastic report card replica was used only as long as necessary, consistently for the very first expectation interview, infrequently at the end of first grade, and never beyond first grade because by second grade all children could report their guesses verbally.

M, Marks. Children's marks in reading, arithmetic, and conduct (M_R, M_A, M_C) were taken from school records. Marks on the first report card in first grade are "T1 marks." Marks on the last report card in first grade are "T2 marks." (See Figure 2.) During second and third grades, marks were again recorded in the middle of the year and at the end of the year. The mid-year marks were recorded either one-third or one-half way through the year, depending on whether the school used three or four marking periods per year. All marks, like expectations, were recorded on a scale from 1 to 4, with 1 high.

PEER, Peer Rating. The average peer rating of a child was obtained in different ways over the course of the study. Mainly it was procured by noting the order in which children chose up sides "to play a reading game." Two captains, a boy and a girl, were first selected by the teacher according to whom she judged to be best in the class in both reading and social maturity. Captains then took turns picking children to be on their team to play a relay game that involved simple reading.

This variable was scored separately for each classroom as follows. First all children were ranked in order in which they were chosen with two sets of ranks because of the two teams. That is, the two captains were ranked 1.5, the two children chosen first by each captain were ranked 3.5, and so on. Then each child's rank was normalized to lie between zero and one, with the highest scored zero, the lowest scored one, and the others located at equal intervals between. In a class with 22 students, 2 would receive zero, two would receive one, and the 18 lying within the two extremes would receive scores of .10, .20, .30 and so on. (In the last two cohorts in all schools peer popularity

scores were obtained by similarly scaling the number of children who chose any particular child as their best or second best friend. In the first two cohorts in the middle class school a more complicated procedure was used by which each child chose six classmates on two rounds. Then the same procedure for scaling was used.)

ABS, Absence. This is the total number of absences for the entire school year as recorded on the year-end report card.

Social Context. An important part of the conceptual scheme undergirding this research is that the child's social milieu--the racial and socioeconomic composition of the school and neighborhood--may affect the schooling process. Schooling effects may vary according to social context since, for example, middle class children may be more responsive than lower-class children to parents' expectations, or absences may impair performance in one place more than another. To take account of social context, three schools were selected to typify three common types of elementary schools in the United States in the 1970's--a white suburban middle-class school, an integrated urban lower-class school, and a black urban lower-class school. Characteristics of these three schools and the clients they serve are summarized from data provided by the Maryland Accountability Program Reports for 1973-74 and 1974-75 (Table 1). Models to explain the schooling process were estimated separately for each of the schools, so social context as a variable occurs between models.

All data on school facilities point to reasonable comparability among schools--enrollment, pupil-staff ratios, and teachers' qualifications are highly similar from one school to another. By contrast, all data pertaining to students or their families point to lack of comparability between the all-white suburban school and the two urban schools, but to close comparability between the two urban schools. On both median education and median family income, the suburban school significantly outranks its urban counterparts and the percent disadvantaged is negligible. Summary data for pupils (non-verbal ability test and grade equivalent scores on vocabulary, reading comprehension, and arithmetic in third grade) show higher figures for the suburban school on all pupil-test measures but comparable patterns for the two urban schools. (Similar data collected during the first grade (see Entwisle & Hayduk, 1978) show smaller differences at first grade between the integrated lower-class and white middle-class schools than those in Table 1 which are for third grade.)

Sampling Design

This research is unusual in terms of both the length and nature of its sampling. The sampling plan is not easy to describe because it does not fit neatly under any of the labels commonly used. It is

"longitudinal" in more than one sense of that term because children in three different schools were followed for as long as three years and because a set of causal processes, assumed to operate in a particular temporal sequence, are represented by relations among variables measured at times closely matched to the causal assumptions.

Variables were measured at many times over the three-year period. For example, achievement in arithmetic, initially measured at the time the first report card was issued, can be seen as a consequence of parents' expectations (measured earlier in first grade) or of children's expectations for themselves (measured prior to the issuing of report cards). But since achievement in arithmetic is measured also at the end of first grade and on four subsequent occasions, later achievement can be seen as a consequence of earlier achievement (among other things). In every case the time when a variable is measured can reflect some notion of causal priorities. The frequent and repeated measurement of variables in this study stands in marked contrast to studies of secondary students where measures are generally secured on only one or two occasions. Also directly asking informants (parents, teachers and peers) other than students for such information is fairly unusual.

Children in the three schools were followed for a minimum of one semester and a maximum of six semesters. A "cohort" consists of all the children starting first grade in a particular year in any school. Each cohort was followed until it completed third grade or until the study ended. The white middle-class school was studied for five school years (from 1971 to 1976), the integrated lower-class school for four school years (1972 to 1976) and the black lower-class school for three school years (1973 to 1976).

For the white, integrated, and black schools, 34%, 23% and 23% of the children were in attendance for a full three years (190, 140 and 95 children, respectively). A further 97, 79, and 30 completed the first two grades, and 102, 151 and 130 completed the first grade only. The annual incrementing of the samples, adding more first-grade children for each successive year the school was in the study, is the major reason for differences in children's longevity in the study. That is, children who entered first grade in 1975 could be followed for only one year, until the study ended in June 1976. Those who entered in 1974 could be followed for a maximum of two years. Proportionately fewer children appear in the higher grades in the lower class schools because these schools were included in the research later (1972 and 1973 for the integrated and black schools respectively). However, the minimum case base for any estimate is two cohorts, and these minimum numbers occur for third grade only in the two lower class schools. All of the other aggregations serving as a basis for any parameter estimate consist of three or more cohorts.

In aggregating cohorts we assume that inter-cohort differences are small compared to other differences. Comparisons between schools based on aggregated data are blind to any differences caused by a 1971-76 duration for first-grade data collection in the middle-class school vs. the 1972-76 or 1973-76 durations in the two lower-class schools, for example. Also some cohorts may be more socially cohesive than others, and although there was considerable stability in the staff at all three schools, some cohorts had different teachers from others. It is possible that cohort differences could mask other differences, but if cultural context as it is defined in this study (social class of neighborhood and integrated status of school) is an important influence on schooling, it should be large enough to outweigh variation between schools due to teachers or due to period differences.

Another way to define the sample of children and their longevity in the study is through a tally of the number of students entering and leaving each school for each time period, T1 to T6 (See Table 2). The school catchment areas were revised from time to time while the research was in progress, and these changes, made during summers, are responsible for some entries and departures. In addition, families try to gear household moves to the end of the school year.

Coverage. Our aim was to sample exhaustively each school's population in the required grades during the years the study was in progress. The number of children in any school is a moving target, of course, since enrollment changes because of within-semester arrivals and departures. However, students arriving at any time during a semester were very likely to be included for that semester because most of the measures were taken just before the end of each semester.

Not many students were missed for other reasons. For instance, if necessary several call-backs were made to interview absentees. Interviewers returned to schools at least three times to interview children for expectations, and efforts to locate parent respondents were equally intense. Only an illness that kept a child away from school for several weeks would lead to failure to obtain a measure of that child's expectations, and if marks for the current semester were available, even such a chronic illness would not lead to exclusion of the child from the panel, merely to a "missing value" on the expectation variable.

Starting in the fall of 1974 when it became necessary to obtain informed written consent from every child's parent or guardian, in all schools there was some reduction in coverage. This reduction was mainly because we had difficulty in reaching some parents rather than because of refusal. In fact the major reason for exclusion of some children after September 1974 was failure to obtain informed consent from parents. Pilot work for this research started in 1969 and from then until June 1974 it was necessary to secure approval from school authorities but not from individual parents for a child to be included

in the research. In the schools selected for this study there were no teachers who declined to participate, and therefore all classrooms in appropriate grades were in the sample.

Occasionally a student was too shy to respond to the interviewer. In such a case the interviewer did not press the child but tried an interview at some later date. In a miniscule number of cases, probably no more than 3 in any school over the entire period of the research, a value is missing on the child's expectation variable for this reason. The child would still be part of the sample, however, because marks and the other information could be obtained.

Summary. The sample consists of 1567 Baltimore children divided as follows: 557 who attended a white middle-class suburban school at some time between September 1971 and June 1976; 604 who attended an integrated lower-class urban school sometime between September 1973 and June 1977. The aim was to include all children who began the first grade in each of the schools starting from September of the year in which the school joined the study, and to follow each child who remained enrolled in the school through the end of the third grade year. The total sample for each school is composed of overlapping two-wave panels of children where the majority, but not all, of those present in one panel are present in the next.

Results to Date

The results now available are 18 matrices of parameters estimates for the full data set (6 models for each of three schools). These matrices are available in Entwisle and Hayduk (1980).

Space does not permit discussion of the findings except for a few outstanding points.

- 1) This work indicates the feasibility of studying schooling in the earliest grades, and there is a heartening consistency in model structure across schools and from year to year. This consistency, together with the respectable proportions of variance accounted for, indicates that models probably have some validity.
- 2) The influence of race appears negligible. The number of "significant" effects associated with it are few and general patterns in the findings point to socioeconomic status as a much more important variable than race in early schooling.
- 3) Children's expectations are considerably less well-determined than marks. They seem to be in a gradual process of formation over the first three years of school. Although they persist to a similar degree in each school, their determinants look different by school.

- 4) Early marks are surprisingly well determined, but more so in the lower-class schools than in the middle-class school. Children appear to have more influence on their marks in the middle-class school than in the lower-class schools.
- 5) Conduct seems to be an important determinant of early school achievement. Since expectations for conduct are responsive to gender, the child's deportment could easily mediate sex differences in performance.
- 6) In the middle-class school in each cycle children's expectations responded to prior marks, whereas in the lower-class schools no such response appeared. The differential feedback pattern argues that only children in the middle-class school attend to evaluative feedback. Lack of attention to feedback could explain why children's expectations are so much higher in the two lower-class schools than in the middle-class school and also why they decline very little over time.

References

- Adams, N. Birth Order: A critical review. Sociometry, 1972, 35, 411-439.
- Coleman, J. S., Campbell, E. Q., Hobson, C. J., McPartland, J., Mood, A., Weinfeld, F. D., & York, R. L. Equality of Educational Opportunity. Washington, D.C.: U.S. Government Printing Office, 1966.
- Cronbach, L. J., & Furby, L. How we should measure "change"--or should we? Psychological Bulletin, 1970, 74, 68-80.
- Entwisle, D. R., & Hayduk, L. A. Too Great Expectations. The Academic Outlook of Young Children. Baltimore: Johns Hopkins University Press, 1978.
- Entwisle, D. R., & Hayduk, L. A. Schooling of Young Children: Cognitive and Affective Outcomes. Final Report. National Institute of Education Project No. NIE-G-74-0029. 1979.
- Entwisle, D. R., & Hayduk, L. A. Schooling of Young Children. Book in preparation. 1980.
- Entwisle, D. R., & Webster, M. Raising Children's Performance Expectations. Social Science Research, 1972, 1, 147-158.
- Estes, W. K. Learning theory and mental development. New York: Academic Press, 1970.
- Goldberger, A. S. Structural equation models: An overview. Structural Equation Models in the Social Sciences. A. S. Goldberger and O. D. Duncan (eds.) New York: Seminar Press, 1973.
- Heyns, B. School Learning. New York: Academic Press, 1978.
- Husen, T. Talent, Opportunity and Career. Stockholm: Almqvist and Wiksell, 1969.
- Joreskog, K. G., & Sorbom, D. LISREL III: Estimation of Linear Structural Equation Systems by Maximum Likelihood Methods. A FORTRAN IV Program. National Educational Resources, Inc., 1976.
- Kraus, P. E. Yesterday's Children. New York: Wiley, 1973.
- Meyer, W. J., & Thompson, G. G. Sex differences in the distribution of teacher approval and disapproval among sixth-grade children. Journal of Educational Psychology, 1956, 47, November 385-396.
- Mosteller, F., & Moynihan, D. P. (eds.) On equality of educational opportunity. New York: Vintage Books, 1972.

Perry, E. Communication of teacher expectations over time. Unpublished paper presented at the American Psychological Association meetings, Chicago, 1975.

Rist, R. C. Student social class and teacher expectations: the self-fulfilling prophecy in ghetto education. Harvard Educational Review, 1970, 40, 411-451.



UNIVERSITY OF SOUTHERN CALIFORNIA

social science research institute

LONGITUDINAL RESEARCH IN THE UNITED STATES:
RELEVANCE TO PRIMARY PREVENTION OF DELINQUENCY

By
SARNOFF A. MEINICK

And
NICOLE HARWAY

VOLUME II-B

DESCRIPTIONS OF LONGITUDINAL RESEARCH

Prepared for the United States Department of Justice,
Law Enforcement Assistance Administration under contract
No. 78-NI-AX-0088. Points of view or opinions stated
in this document are those of the authors and do not
necessarily represent the official position or policies
of the United States Department of Justice.

84660.42

TABLE OF CONTENTS

DESCRIPTIONS OF LONGITUDINAL RESEARCH

VOL. II B

1. The Rutgers Health and Human Development Project: A Longitudinal Study of Alcohol and Drug Use by D. Lester, et al.	505 - 516
2. The Study of Adult Development at Harvard Medical School by G.E. Vaillant	517 - 532
3. The Duke Longitudinal Studies by E. Palmore	533 - 547
4. Older Workers in Rural Areas by E.A. Powers, W.J. Goudy & P.M. Keith	548 - 557
5. The Longitudinal Study of Transitions by D.A. Chiriboga	558 - 577
6. The Consumer Panels of Young Married Couples by E. Crimmins, L. Lannom & P.A. Ewer	578 - 589
7. First and Second Marriages: The First Three years of Married Life by E. Benson-von der Ohe	590 - 618
8. Patterns of Child Rearing (PCR) by R.R. Sears	619 - 632

NCJRS

FEB 4 1982

ACQUISITIONS

9. The Hyperkinetic Child Grows Up: Predictors of Symptoms, Delinquency and Achievement at Follow-up by J. Loney, J. Kramer & R. Milich	633 - 648
10. Life Assessment of a Cohort of Chronic Schizophrenics Discharged Twenty Years Ago by C.M. Harding & G.W. Brooks	649 - 666
11. A Long-Term Follow-up Study of Former Adolescent Psychiatric Patients by W. Quast & A.J. Hafner	667 - 571
12. Long-Term Follow-up of the Major Psychoses by M.T. Tsuang	672 - 681
13. Who's in that Dark Alley? by S. Dinitz & J.P. Conrad	682 - 688
14. A Longitudinal Study of Delinquency and Drop-Outs by D.S. Elliott	689 - 706
15. From Boy to Man - From Delinquency to Crime by M.E. Wolfgang	707 - 711
16. Health Consequences of Drug Use: A Longitudinal Study of Urban Black Youth by A.F. Brunswick	712 - 735
17. Interim Report on Prospective Study of Children at Risk for Schizophrenia by L. Erlenmeyer-Kimling et al.	736 - 749

18. The Consortium for Longitudinal Studies: A Follow-up of Participants in Early Childhood Programs by H.W. Murray et al.	750 - 765
19. Two Decades of Research: The High-Scope Longitudinal Pre-school Evaluations by D.P. Weikart & L.J. Schweinhart	766 - 778
20. Monoamine Oxidase Activity and the Longitudinal Biological High Risk Approach to Schizophrenia and Affective Illnesses by Robert D. Coursey, Monte S. Buchsbaum & Dennis L. Murphy	779 - 806
21. A Longitudinal Study of Personality Development by J. McCord	807 - 818
22. Longitudinal Data from the Seattle and Denver Income Maintenance Experiments: Designs and Uses by M.C. Keeley	819 - 837

The Rutgers Health and Human Development Project:
A Longitudinal Study of Alcohol and Drug Use

David Lester

Robert J. Pandina

Helene R. White

Erich W. Labouvie

National Alcohol Research Center

and

Center of Alcohol Studies

Rutgers University

New Brunswick, New Jersey

Introduction

Since its transfer from Yale University in 1962, the Center of Alcohol Studies has been an integral part of Rutgers University.

Even in its formative years, the Center's faculty realized the importance of longitudinal studies of the development of alcoholism. In 1949, Robert Straus and Selden Bacon (1953) collected data from 15,747 students in 27 colleges and universities across the United States. Following up some of these students became possible 25 years later: a sample (1,572) of these former respondents is currently being studied (Fillmore, Bacon, & Hyman, 1977) with a view to tracing the careers of alcohol users and non-users from early adulthood to middle age, with special emphasis on the sociological and sociopsychological factors antecedent to the emergence, remission, or maintenance of problem drinking.

CONTINUED

6 OF 10

Elucidation of the etiology of alcoholism demands intensive study of many individuals over a substantial portion of their lives. Studies of the variation in characteristics between alcoholics and nonalcoholics may yield certain correlates of alcoholism, but there is always the question whether these characteristics preceded the onset of alcoholism and precipitated the problem drinking or whether these characteristics were produced by the drinking (Lester, 1966; Wallgren & Barry, 1970; Chandler, 1972). Previously reported longitudinal studies in this area (McCord, McCord, & Gudeman, 1960; Robins, Bates, & O'Neal, 1962; Jones, 1968) have been of limited scope, e.g., using a single cohort or a set of measures from one discipline, caveats which also apply to the College Drinking Study mentioned above (see also Zucker, 1976; Kandel, 1978). We believe that the origin of deviant drinking and of its varied subsequent problems (family, medical, etc.) is a complex phenomenon, the understanding and disassembly of which cannot be expected with the use of methods and measures of any one discipline or profession.

Beginning in 1973, therefore, we decided that discussions alone of longitudinal studies, and especially of a prospective multiple cohort study, were insufficient. What was required was preparation of a specific protocol detailing the variety of factors which such a study should encompass, including specific measures of independent and dependent variables within a design which would be amenable to analysis and which would lead to the elucidation of significant clusters of predictive factors.

The Center of Alcohol Studies was designated a National Alcohol Research Center in March 1978. Initial funding was sufficient to provide only for the initiation of the longitudinal study, but we are optimistic that the long-term prospects will allow for ancillary investigations to enrich the major study.

Our objective is to characterize, differentially, the population of problem drinkers by establishing their distinctive biological, social, and emotional makeup and by composing the history of the development of drinking. Although this study is centered on the major dependent variables of alcohol consumption and alcohol problems, it will also explore the use of other drugs (e.g., marijuana, nicotine, etc.). If this study were to be restricted to alcohol, it would assert for alcohol a uniqueness it may possess to only a limited degree. The presence of ethyl alcohol in the natural environment and, indeed, as an endogenously occurring substance in many organisms with abundant enzymes for its metabolism, does confer upon alcohol some special characteristics, not least the voluminous history of its social use by almost all peoples. Its easy availability and widespread social sanction make it an acceptable drug in those situations where it helps people to enjoy their environment or to cope with or escape from it. These functions of alcohol may change over time as other drugs attain acceptability; the decriminalization of marijuana, for example, may provide such a competitor, and we may anticipate other such possibilities. We must, therefore, consider the possibility that an apparent predisposition to alcoholism may, in fact, be a

predisposition to addiction by drugs which alter or manipulate mood. In the context of this study, not to investigate other drug use and misuse would inhibit the range of questions to be asked and possibly screen out the most relevant answers.

We have every expectation that the study will provide a wealth of material concerning the nature of human development and useful baseline information on the acquisition and emergence of a variety of human behaviors, including such behaviors which come under the rubric of mental illness.

Design

The Health and Human Development Project is conceptualized as a multiple time-series design (Labouvie, 1978). It involves a series of cohorts observed at several age levels (see Table 1). The measurement of different cohorts is initiated at different times. Each cohort-specific time series includes a longitudinal and a cross-sectional sequence (Baltes, 1968). All longitudinal sequences are based on samples with a minimum of 75 males and 75 females in each at the first time of measurement. All cross-sectional sequences are based on series of independent control samples obtained from the same cohort; each control sample is observed only once and consists of a minimum of 50 males and 50 females.

Although Table 1 depicts only a series of cohorts from 1961 to 1969, current plans call for continued recruitment of successive cohorts born between 1970 and 1983 and for initiation of measurement on each at the age of 12.

Project Preparations

Preparations for Participant Intake

The first participant entered the study on 7 May 1979. In the months prior to this moment, equipment and supplies were readied and a myriad of organizational details settled. Among the more important of the latter were the training of telephone and home interviewers in conjunction with our subcontractor, the Eagleton Poll, a division of the Eagleton Institute of Politics of Rutgers University, the preparation of a variety of organizational forms and the development of numerous questionnaires for participants, for their parent(s) or guardian(s), for ineligibles (by reason of age), for eligibles unwilling to participate, etc.

Telephone Interview Phase (TIP)

Participation is restricted to New Jersey residents; these participants are sought through telephone interviews conducted by the Eagleton Poll. Telephone numbers are generated randomly and selected by a computer which weights telephone exchanges by population density; households, therefore, with unlisted numbers are included. Since 95% of New Jersey households have a telephone, the sample which results from the TIP can be assumed to be highly representative of the New Jersey population.

The questions utilized in the TIP elicit demographic information about the respondent's household: family composition, race/ethnicity, level of income, occupation, education, and alcohol consumption. The questionnaire acts as a screening device for locating families with an eligible participant; for example, in 1979, eligibility is conferred by birth in 1961, 1964, or 1967. When such a household is identified, the interviewer explains that the family is eligible to participate in a health and human development study focusing on adolescents and young adults. Brief background information about the project is presented: the study is financed by HEW (now HHS) and the family will be compensated in the amount of \$75 and expenses. If the respondent agrees, the interviewer obtains the name and address, telling the respondent that someone more conversant with the project will call back to explain the project in greater detail so that a more informed decision on participation can be made. Names of eligible families are grouped geographically and assigned to a field interviewer, also employed by the Eagleton Poll, for telephone call-back and possible subsequent home interview.

Personal Interview Encounter Phase (PIEP)

A network of field interviewers, located in various regions of New Jersey, are responsible for the telephone call-back. They engage in a thorough discussion of the project and arrange a home interview. If participation is refused, either during call-back or home visit, reasons for the refusal are collected. In addition, those who refuse during the home visit are asked to complete a questionnaire for nonparticipants, comprised of a concise version of all the questions asked of assenters.

During the home visit, consent forms are explained and signed. Questionnaires (described later) are completed by each parent and the participant, and locator forms, to aid in follow-up, are also completed. These questionnaires and forms are mailed to the project (H&HD). A short demographic questionnaire is left for both parents, and the Health History questionnaire is left for parents and participant to complete, being either mailed to H&HD or brought in person at the time of the on-site visit. Through a toll-free call to H&HD, the field interviewer arranges an appointment for the on-site visit. The H&HD telephone is answered, in person, daily until 10:00 P.M.; at all other times, scheduling information is supplied by an answering and recording device, also available to

participants if an appointment change needs to be arranged. Appointments are confirmed a day in advance.

On-Site Procedures

The staffing pattern, physical facility, and task demands of the study place an upper optimal limit of four on the number of subjects who can be processed adequately on any day.

Four schedules guide subjects through the day's routine, coordinating various tests and promoting efficient utilization of staff and facility. The schedules are constructed so that appropriate consideration is given to the demands placed upon participants: thus, tasks requiring completion of questionnaires are interdigitated with behavioral, cognitive, and physiological tasks and with interactive responses to questions at a computer terminal. The purpose of such task dispersion is to counteract fatigue and any response set bias, either of which might interfere with valid and reliable responding.

At the end of the day, the subject engages in a short debriefing with the assigned interviewer; answers are sought to questions about the day's experiences as well as some personal questions to maintain rapport. Participants are paid at this time, while the parents' fee and travel expenses are paid by mail when travel vouchers are received by H&HD. Subjects are finally reminded of their continued involvement with H&HD, that is, follow-up and retesting 3 (or 6) years later.

Tracking Procedures

There are several procedures used to ensure a maximum rate of tracking and recontacting participants. In addition to the addresses of parents and participants, the field interviewers secure completion of locator forms, which provide information as to mailing address(es), work telephone(s), school name(s), social security number, driver's license number, and the address and telephone number of one close friend and one close relative.

During the on-site visit, this information is confirmed and information about possible moves (e.g., college admission, etc.) is also acquired. During the on-site debriefing, the events of the day are discussed in order to leave the participant with a positive feeling for the project and its staff. The payment of \$50 to the participant provides an immediate reward.

Approximately 3 months after the on-site visit, a follow-up telephone

interview is conducted with the participant. Its purpose is to obtain feedback on the whole experience, but more importantly, personal contact is maintained with the participant. (Virtually all participants so contacted have reconfirmed their intention to return in 3 years.) Participants are instructed to send in a card with change of address (if they move) and they are informed that H&HD will pay \$3 when the address change is verified.

During the Christmas holidays, a greeting card will be sent to the participants and their families from the H&HD staff. Telephone calls at 1- and 2-year intervals will be made to maintain contact and to sustain and nurture the participant-H&HD relationship. We expect these efforts, among others, to minimize or eliminate loss of subjects.

A follow-up of nonparticipants will also be made. Its purpose is to convert nonparticipation and, failing this, to obtain each additional information as to evaluate the representativeness of the sample of participants.

Measures and Instruments

We believe alcohol and drug-use behavior, including alcoholism, has a multiple causation--an interaction, over time, of biological constitution and environment. Any attempt at a prospective study which avoided measures from one or another area would be fraught with the danger of missing an important element (or elements) contributing to various drinking outcomes. Thus, our measures are drawn from a variety of areas: sociodemographic, sociocultural, psychosocial, psychological, psychophysiological, behavioral-cognitive, biochemical, and medical. Data are collected by subject self-report, from significant others (parents and spouses), behavioral tests, interaction with questionnaires via computer, physical examination, analysis of blood components, and psychophysiological tests.

Sociodemographic, Sociocultural, and Psychosocial Measures

Data in these areas are collected primarily by self-report questionnaires completed by participants and their significant others. The questions employed were developed after examination of over 50 questionnaires used in other studies of adolescents, drinking and drug-taking, deviance, etc. At home, the participant supplies information about school, leisure time, work, neighborhood, family composition, politics, and completes the Zuckerman Sensation Seeking Scale (Zuckerman, et al., 1964). On-site information is supplied about religion, deviant behavior, feelings about self, coping mechanisms, stress, sex role orientation, and additional information about school and work. The interviewer elicits more information about the family through personal interviews. The participant also

completes a questionnaire asking about relationships with parental figures [including a revised version of the Youth Perception Inventory (Streit, 1978)], siblings, and close friends. Information is sought about the degree of interaction, the social controls exhibited by and the behavior and attitudes of these significant others.

Data regarding the dependent variables of alcohol and drug use are collected through an Alcohol and Drug Use questionnaire (with information obtained about the pattern of use, if any, of 15 substances), and an Alcohol and Drug Experience questionnaire which inquires about attitudes to, motivations for, and consequences of alcohol and drug use.

When appropriate, participants complete questionnaires about work and marriage. Significant others (parents/spouses) respond to home questionnaires which yield sociodemographic, attitudinal, and behavioral data regarding themselves and the participant.

Psychological Profiles

Answers relevant to a psychological profile of the participant are obtained via responses at an interactive computer terminal. Two standard questionnaires are used, a psychiatric symptom checklist, the SCL-90R (Derogatis, 1977), and an abbreviated version of the Personality Research Form (PRF-E) (Jackson, 1974). Besides yielding information adequate for standard comparisons--and having the advantage of immediate on-line entry of responses into the data file--the response time to each question also becomes part of the data file. We believe that these response times may be revealing of other facets of the personality. Participants enjoy this mode of questioning which provides a welcome break from paper and pencil responding.

Psychophysiological and Behavioral-Cognitive Measures

There is increasing evidence that certain psychophysiological and behavioral-cognitive tests can discriminate between alcoholics and nonalcoholics. Certain of these differences undoubtedly arise from the long-term effects of repeated intoxication, but some might be antecedent. We have selected tests in these areas based on considerations of test simplicity (not requiring complicated instructions), seeming independence from intelligence and socioeconomic status, and performability at near asymptotic levels with little practice. Measures of autonomic level and reactivity include the response of the dark-adapted pupil to photic stimuli, and pre- and post-stress measures of skin conductance and finger volume (yielding heart rate and other indices). The stress is immersion of the arm in ice water for 30 seconds. Both locomotor (Heath Rail Walking: Goldstein, et al., 1968) and static ataxia (Schneider, 1970) are assessed with eyes open, the latter with eyes closed. The participant's place on the continuum of field dependence-independence is measured

(Roc-and-Frame--Witkin, et al., 1962) as is perceptual reactance (Petrie, 1967). Impulse control, as judged by Retarded Knob Rotation (Parsons, Tarter, & Edelberg, 1972; Parsons, 1975) is also measured.

Biochemical Determinations

Blood samples are drawn on two occasions, a baseline sample during the medical examination and the second immediately after the cold stressor. These samples are analyzed for cortisol, sex hormones (testosterone, progesterone, and estrogens), and catecholamines. A substantial plasma sample is retained and stored in liquid nitrogen, providing the future possibility of prospective determination of compound(s) which we have not had the prescience to determine when the sample was drawn. Platelets are also being separated from the blood for determination of monoamine oxidase levels.

Health History and Medical Examination

A Health History questionnaire is left during the home visit to be answered by participant and parents and a supplemental form is completed on-site. On-site, a thorough medical examination is conducted by a licensed physician and the results of this examination are made available to the participant's family physician upon written request.

Project Status

In the period 15 April 1979 to 31 October 1979, about 20,000 interviews were conducted by telephone to obtain a pool of eligible participants. Based on this experience, we estimate that about 750 individuals will be both eligible to participate and willing to be involved in the study. These projections indicate that there should be no problem in easily reaching our goal of 450 participants (150 per birth cohort) in the first 12 months of intake.

Ten percent of the households contain an eligible respondent. Of all such households, about 84% agreed to a further, more extensive telephone discussion. Our data, based on some 2,000 eligible households, indicate that those who refuse to be called back are similar to those who agree, as judged by the responses to sociodemographic and alcohol consumption questions asked anonymously in TIP.

Almost half (46%) of all respondents who agree to be called back agree to a home visit. Having taken this step, almost all (99%) agree to enter the study. In terms of race, religion, and alcohol consumption (except wine), participants are similar to refusers; the former, however, tend to

be of somewhat higher socioeconomic status.

Since the first participant entered the study, 278 participants have been tested on-site, and an additional 90 have been scheduled for on-site visits (as of the 31st of October, 1979). At the present rate of processing subjects, we anticipate completing the first wave of on-site testing by the 23rd of February, 1980 and beginning on-site testing of Wave 2 of the study (individuals born in 1962, 1965, and 1968) on the 1st of May, 1980.

Concluding Comments

Three essential elements of this study provide it, we believe, with advantages not possessed by the few longer standing longitudinal drug and alcohol use studies.

It is expected that identification of persons at risk for alcoholism and other alcohol- and drug-related problems, as well as identification of persons not at risk for those problems, will make possible the design of interventions targeted at individuals and also designed to treat rationally a specific prodromal profile of disorder. Clearly, important information on etiology and development will not be obtained for some years, yet important cross-sectional knowledge of the development of drinking and drug use behaviors will be available within the initial period of this project. Knowledge of these problems will help in the design of appropriate prevention strategies for specific problem complexes, and contribute equally to more effective treatment.

In sum, this study appears to embody the expectations voiced many years ago by Jellinek (1942), the first director of the Center of Alcohol Studies, that the combined researches, insights, and working hypotheses of many disciplines were "sufficient to warrant application of the existing knowledge to the investigation of the essential and complex problems of the origin of inebriety and addiction, their prevention and treatment" (pp. xxiii). It is our good fortune that it is at last possible to launch this most needed and fundamental research.

Table 1

Participant Entries into the Health and Human Development Project

Cohort	Age							
	12	15	18	21	24	30	36	42
1961	--	--	1979	L2 C1	L3 C2	L4 C3	L5 C4	L6 C5
1962	--	--	1980	L2 C1	L3 C2	L4 C3	L5 C4	L6 C5
1963	--	--	1981	L2 C1	L3 C2	L4 C3	L5 C4	L6 C5
1964	--	1979	L2 C1	L3 C2	L4 C3	L5 C4	L6 C5	L7 C6
1965	--	1980	L2 C1	L3 C2	L4 C3	L5 C4	L6 C5	L7 C6
1966	--	1981	L2 C1	L3 C2	L4 C3	L5 C4	L6 C5	L7 C6
1967	1979	L2 C1	L3 C2	L4 C3	L5 C4	L6 C5	L7 C6	L8 C7
1968	1980	L2 C1	L3 C2	L4 C3	L5 C4	L6 C5	L7 C6	L8 C7
1969	1981	L2 C1	L3 C2	L4 C3	L5 C4	L6 C5	L7 C6	L8 C7

Cohorts are defined in terms of year of birth. Table entries indicate when measurement is initiated for a given cohort.

L2, L3 . . . L8 = cohort-specific longitudinal sample observed every three years until the age of 24, and every 6 years thereafter.

C1, C2 . . . L7 = cohort-specific series of independent control samples each observed only once.

References

- Baltes, P.B. Longitudinal and cross-sectional sequences in the study of age and generation effects. Human Development, 1968, 11, 145-171.
- Chandler, M. Strategies of research into the etiology of alcoholism. Ann. N.Y. Acad. Sci., 1972, 197, 218-221.
- Derogatis, L.R. SCL-90R (Revised Version) Manual-I. Baltimore: School of Medicine, Johns Hopkins University, 1977.
- Fillmore, K.M., Bacon, S.D., & Hyman, M. Alcohol drinking behavior and attitudes: Rutgers panel study. Report prepared for the National Institute on Alcohol Abuse and Alcoholism under Contract No. ADM 281-76-0015. New Brunswick: Rutgers Center of Alcohol Studies, 1977.
- Goldstein, G., Chotlos, J.W., McCarthy, R.J., & Neuringer, C. Recovery from gait instability in alcoholics. Quart. J. Stud. Alc., 1968, 29, 38-43.
- Jackson, D.N. Personality research form manual. Goshen, N.Y.: Research Psychologist Press, Inc., 1974.
- Jellinek, E.M. Effects of alcohol on the individual; a critical exposition of present knowledge (Vol. I). Alcohol addiction and chronic alcoholism. New Haven: Yale University Press, 1942.
- Jones, M.C. Personality correlates and antecedents of drinking patterns in adult males. J. Consult. Clin. Psychol., 1968, 32, 2-12.
- Kandel, D.B. Convergences in prospective longitudinal surveys of drug use in normal populations, in D.B. Kandel (Ed.) Longitudinal research in drug use. New York: Wiley, 1978, pp. 3-38.
- Labouvie, E.W. Experimental sequential strategies for the exploration of ontogenetic and socio-historical changes. Human Development, 1978, 21, 161-169.
- Lester, D. Self-selection of alcohol by animals, human variation, and the etiology of alcoholism. A critical review. Quart. J. Stud. Alc., 1966, 27, 395-438.
- McCord, W., McCord, J., & Gudeman, J. Origins of alcoholism. Stanford, Calif.: Stanford University Press, 1960.
- National Institute of Mental Health. Alcohol and alcoholism: Problems, programs and progress. National Institute of Mental Health, Dept. of Health, Education and Welfare. Publ. No. (MSM) 72-9127, 1972.

Parsons, O.A. Brain damage in alcoholics: Altered states of consciousness, in M.M. Gross (Ed.) Alcohol intoxication and withdrawal. II. Experimental studies. New York: Plenum, 1975, pp. 569-584.

Parsons, O.A., Tarter, R.E., & Edelberg, R. Altered motor control in alcoholics. J. Abnorm. Psychol., 1972, 80, 308-314.

Petrie, A. Individuality in pain and suffering. Chicago and London: University of Chicago Press, 1967.

Riley, M.W. Sociological research: A case approach. New York: Harcourt, Brace and World, 1963.

Robins, L.M., Bates, W.M., & O'Neal, P. Adult drinking patterns of former problem children, in D.J. Pittman and C.R. Snyder (Eds.) Society, culture and drinking patterns. Carbondale, Ill.: Southern Illinois University Press, 1962.

Schneider, E.W. Characteristics of alcohol ataxia. Ph.D. Dissertation. New Brunswick: Rutgers University, 1970.

Straus, R., & Bacon, S.D. Drinking in college. New Haven: Yale University Press, 1953.

Streit, F. Technical manual: Youth perception inventory. Highland Park, N.J.: Essence Publications, 1978.

Wall, W.D., & Williams, H.L. Longitudinal studies and the social sciences. A social science research council review. London: Heinemann Educational Books, Ltd., 1970.

Wallgren, H., & Barry H. III. Actions of alcohol. Amsterdam, London, New York: Elsevier, 1970, Vol. 2, pp. 774 et seq.

Witkin, H.A., Dyk, R.B., Faterson, H.R., Goodenough, D.R., & Karp, S.A. Psychological differentiation. New York: Wiley, 1962.

Zucker, R. Parental influence upon drinking patterns of their children, in M. Greenblatt and M.A. Schuckit (Eds.) Alcoholism problems in women and children. New York: Grune and Stratton, 1976, pp. 211-238.

Zuckerman, M., Kolin, E., Price, L., & Zoob, I. Development of a sensation-seeking scale. J. Consult. Psychol., 1964, 28, 477-782.

The Study of Adult Development at
Harvard Medical School

George E. Vaillant, M.D.
Professor of Psychiatry, Department of Psychiatry
Harvard Medical School, Boston, Mass.
Research Psychiatrist
Harvard University Health Services, and
Director of The Study of Adult Development,
1493 Cambridge Street
Cambridge, Mass. 02139

*This work was supported by the W.T. Grant Foundation, Inc., the Commonwealth Fund, the Carnegie Corporation, The Field Foundation, the Ford Foundation, the Spencer Foundation, and Grants M1817, MH07286, MG-38798, MH-32885 from the National Institute of Mental Health and (AA-01372) from the National Institute of Alcoholism and Alcohol Abuse.

The Study of Adult Development at
Harvard Medical School

George E. Vaillant, M.D.

The Study of Adult Development at Harvard Medical School has followed 660 adult men's lives (The COLLEGE Sample and the CORE CITY Sample) from 1940 to 1980, and from adolescence into late middle life. Attrition due to men withdrawing or becoming lost to the follow-up has been held at 6%. Due to the length of repeated contacts, these studies vie with the Terman Study and with the Oakland Growth Study (see this volume) as the longest studies of adult development in the world. The Oakland Growth Study sample has been more thoroughly studied, but it is smaller and has suffered more attrition. The Terman sample includes both sexes, is larger and of longer duration (from age 10 to 70); but after childhood most of the Terman subjects have not been interviewed.

The COLLEGE Sample: The Grant Study

In 1938, thanks to a generous gift from the philanthropist, William T. Grant, the Harvard University Health Services under the leadership of Arlie V. Bock, M.D., undertook a study of "healthy," college sophomores. Until 1954 the research program was directed by Clark Heath, M.D. The early years of the study resulted in several dozen publications, most of which are reviewed in the three books, What People Are, Young Men You Are Normal, and College Men at War. (Heath, 1945, Hooten, 1945, Monks, 1951.) From 1954 to 1972 Charles McArthur, Ph.D., was the director of the study. Since 1972, the study has been directed by George E. Vaillant, M.D. The study has always remained the administrative responsibility of the Harvard University Health Services.

The average subject was 18 years of age when the Grant Study began in 1938, and has been studied until the present when the average subject has passed his 60th birthday. In all, 268 men were chosen: a few came from the classes of 1939-41, but 204 came from a 7% sample of the classes of 1942-44. About one-tenth of the sample was chosen by change factors (e.g., 4% by being self-referred, and 2% for being younger brothers of subjects already selected for the study). Nine-tenths of the sample was selected in the following fashion. About 40% of each class were excluded

as a result of mediocre academic achievement. Known medical or psychological difficulties led to the exclusion of 30% more. The names of the remaining 30% of the class were submitted to the college deans who selected from this group about 100 students who they recognized as "sound." From that subgroup -- now narrowed down to 10% of the original class -- one sophomore in five was not actually accepted into the study as a result of schedule conflicts or poor motivation. Once accepted into the study, the COLLEGE subjects were most loyal. During their college years only 10 of the 268 subjects finally selected for the study dropped out. Since then, only two men have withdrawn.

After being accepted into the Study, each man was seen by a psychiatrist for eight interviews. These interviews focused on the man's family, on his own career plans and on his value systems. The psychiatrist made an effort to get to know the men as people rather than patients. No effort was made to look for pathology nor to interpret the men's lives psychoanalytically. The psychiatric interviews included a history of early sexual development, but unfortunately the psychiatrist did not inquire into the boy's friendship or dating patterns. Thus, many early questions relevant to the vicissitudes of middle life went unanswered. (In 1937, when the study began, Anna Freud and Heinz Hartmann had only just published the two monographs that were to launch ego psychology. In Berkeley, Erikson was only beginning the research that was to make the life cycle a popular and comprehensible intellectual concept and the Sullivans and Kleinians had not yet convinced dynamic psychiatry that interpersonal relationships were of crucial importance.)

The COLLEGE Sample subjects were also seen by a family worker who took a careful social history from each sophomore subject and traveled the length and breadth of the United States to meet the parents. In each boy's home, she took a family history that included characterizations of the grandparents, aunts, uncles and first cousins. She also obtained from the mother a history of the childhood development of each boy and a family history of mental and physical illness.

Each COLLEGE subject received an unusually complete two-hour physical exam, including a complete record of his daily habits, past illnesses, and his physical response to stress. Each man was studied by a physical anthropologist, who recorded his somatotype, determined whether his physical habitus was predominately masculine or feminine, and made exhaustive anthropometric measurements. A physiologist also studied each subject and measured his insulin tolerance, his respiratory functions and the physiologic effects of running on a treadmill for five minutes or until near exhaustion. Finally, each man was given tests by a psychologist, Frederick Wells, designed to reflect native intelligence (the Alpha verbal and Alpha numerical), a vocabulary test, a shortened Rorschach test, and a book assembly test designed to assess manipulative dexterity and the comprehension of spatial relationships. In 1950 the man and their wives were interviewed in their homes by Margaret Lantis, a social anthropologist. To many she

administered the Thematic Apperception Test.

As measured by their college board tests, the academic achievement of the chosen students fell in the top 5% to 10% of high school graduates, but their average Scholastic Achievement Test (SAT) score of 584 did not put them beyond the reach of many other able college students. Because one of the criteria for selection had been successful academic achievement, 61% of the study subjects graduated with honors in contrast to only 26% of their classmates. In native ability, however, the study subjects were only slightly superior.

Socio-economically, the COLLEGE Sample made up a privileged group, but not exclusively so. In 1940, one-third of their fathers made more than \$15,000 a year, but one-third made less than \$5,000. One-third of their fathers had had some professional training, but one-half of their parents had no college degree. Fifty percent of the men had had some private education, but often on scholarship. In college, 40% of the men received financial aid, and one-half worked during the academic year to pay a significant part of their educational expenses. Eighty percent were Protestant, 10% Catholic and 10% Jewish. The study contained no blacks.

The COLLEGE Sample was not selected to be representative of any group, but the net was cast in such a fashion as to have a high likelihood of retrieving a large group of men who would lead satisfactory lives -- regardless of the observer's bias. The sample was not comprised of volunteers who "wanted to be studied." The emphasis was on selecting men at the independent end of the independent-dependent continuum. The men, a majority of whom were first-born sons, had deliberately chosen to go to a difficult and competitive college. Then, they had been further selected for their capacity to master this situation. Put differently, the sample had been chosen for its capacity to equal or to exceed its natural ability. The happy-go-lucky but equally stable youngster, who characteristically searched for a good time, was probably underrepresented. The Stoics outnumbered the Dionysians.

The historical accident of World War II forced these men into a common experience that permitted them to be compared with their fellows on grounds other than academic excellence. These men performed well on battlefield (Monks, 1951). Only 11, instead of a statistically expected 77, were rejected for service on physical defects. Instead of an expected 36, only 3 were rejected for psychiatric reasons. The proportion wounded and killed did not differ from that experienced by the armed forces as a whole. Only 10% went into the Army with commissions, but 71% were officers at discharge. Forty-five percent entered the Navy with commissions, and 90% were officers at discharge.

After college graduation the COLLEGE Sample men were followed until 1955 by annual questionnaires. Since then they have been sent questionnaires every two years. These questionnaires were lengthy and designed

to benefit from these men's high verbal skills. They paid special attention to employment, family, health, habits (e.g. vacation, sports, alcohol, smoking, etc.) and political views. They were reinterviewed in 1951 and a random 50% sample was reinterviewed in 1968-70. Complete physical exams were obtained in 1969, 1974 and 1979.

Twenty five years out of college, 95% of the men had married, and 15% of them had divorced. The model man in the COLLEGE Sample had the income and social standing of a successful businessman or physician, but displayed the political outlook, intellectual tastes, and lifestyle of a college professor. Twenty-five years after college, the subjects remained relatively healthier and occupationally more successful than their classmates. Their mortality was 50% less. Four times as many held class offices as would have been expected by chance. Although less intellectually gifted than the Terman sample, their achievements as measured by inclusion in Who's Who in America and American Men and Women of Science were comparable.

A quarter of the men have become lawyers or doctors: 15% became teachers, mostly at the college level, and 20% went into business. The remaining 40% are distributed throughout other professions like architecture, accounting, advertising, banking, insurance, government or engineering. The proportions of Grant Study men in each occupational subgroup were no different from that of their classmates.

Psychologically, the COLLEGE Sample in adult life fared better than the population as a whole, but it is hard to say how much better. Such comparisons are difficult due to limitations in psychiatric epidemiology. Under the criteria of mental health that Srole and associates (1962) used in their epidemiologic survey of the mental health of urban America, 70% of the Grant Study subjects might have fallen in the 19% that Srole and associates considered "well." By the standards of Luborsky's Health Sickness Rating Scale (1962), 20% of the men at age 47 received a score of less than 70 and so might be defined as psychiatrically ill. In college, the Grant Study psychiatrist estimated that 55% of the men could have benefitted from psychiatric consultation. By the time the COLLEGE Sample was 30, 10% had seen psychiatrists, and by age 48, the number had increased to 40%. In Srole's New York City sample only 13% of subjects, aged 20-59, had ever seen a psychiatrist; however, economics and education may have played a far more important role in such differences than psychopathology.

In summary, the COLLEGE Sample subjects were relatively psychologically healthy, but the precise differences between their health and that of any other group is impossible to ascertain. Chosen as they were for academic success and aided by ethnicity, sex, and the G.I. Bill of Rights and the economic climate of 1945-65, the average Grant Study subject was upwardly socially mobile to a degree that may be uncommon in future historical epochs. Nonetheless, the COLLEGE Sample provides a vivid, if historically limited, view of how the male life cycle may progress under favorable circumstances.

The CORE City Sample: The controls from the Glueck's prospective study of delinquency

The men of the CORE CITY Sample were drawn from the 500 11-16 year old boys selected almost forty years ago by Sheldon and Eleanor Glueck (1950) as controls for their prospective study, Unraveling Juvenile Delinquency. From 1940-44 these boys were chosen from Boston inner city schools on the basis of not being known to be seriously delinquent. The account of how the Glueck's obtained the cooperation and support of the Boston School Committee for this research makes fascinating reading (Glueck & Glueck, 1950.) Like the COLLEGE subjects, the CORE CITY men were originally studied by a multi-disciplinary team of physicians, psychologists, psychiatrists, social investigators and physical anthropologists. Virtually all of the CORE CITY men were interviewed at ages 25, 31 and 47.

Until 1970 this study was directed by Sheldon and Eleanor Glueck of the Harvard Law School. In 1970 they deeded their case records to the Harvard Law School Library. Since 1974, the men in this study have been actively followed by George E. Vaillant, M.D., and co-workers, as part of The Study of Adult Development at the Harvard Medical School. Administratively, these files remain the responsibility of the Harvard Law School.

Originally, Sheldon and Eleanor Glueck matched the CORE CITY subjects with a cohort of 500 youths who were remanded to reform school. A boy in the delinquent group was matched with a boy in the non-delinquent group by four variables: age, intelligence, neighborhood crime rate and ethnicity. The 60% of Boston census tracts with the highest rates of juvenile delinquency contained 95% of the subjects. The subjects' average I.Q. was 95. The parents or grandparents of 70% had been born in Italy, Ireland, Great Britain or Canada. The controls had been selected by reviewing the enrollments of neighborhood schools and due to the tact and preparation of the Glueck staff, the refusal rate was kept to 15%.

The fact that the subjects were controls for a study of urban delinquency imposed several sources of bias. Besides the obvious ethnic and intellectual constraints included in the sample selection, a major source of bias was that the sample excluded about 10% of school boys because of their serious delinquency. Thus, just as the COLLEGE Sample probably excluded passive, underachieving but otherwise perfectly health college students, the CORE CITY Sample probably excluded some ambitious, energetic students who manifested early delinquency but who enjoyed good subsequent outcome.

The Glueck's original study methodology involved two parallel investigations of each boy and his family. Questions of social history and current functioning were resolved by comparing those findings obtained from interviews with the boy, his school, and his family with those findings obtained from the exhaustive search of public records, especially those of the Massachusetts Board of Probation and the Boston Social Service Index.

For 30 years the Massachusetts Board of Probation had cross-indexed arrest records throughout the state; and for half a century the Boston Social Service Index cross-indexed every Boston family's contact with a welfare or social agency. (Unfortunately after 1968 as the state and federal authorities assumed increasing responsibility for public welfare, the Social Service Index was discontinued but its files are preserved at The Study of Adult Development.) The Gluecks' painstaking search of probation, mental health and social agency records allowed documentation of familial delinquency, alcoholism, mental illness and mental retardation for three generations.

Certainly, the simultaneous use of interviews with boy, parent and teacher, as well as multiple longitudinal recorded sources of public information revealed far more evidence of psychopathology than could be obtained by cross-sectional data collection alone. In the 1974-78 refollow-up of the men, the same technique was used once again. An effort was made to check interview data against public records and data from mental health, hospital and law enforcement agencies.

However, depending on a methodology of redundant evidence can distort the evidence in two ways. First, an individual from a very large family would have more relatives at risk for different kinds of psychopathology. Second, recent immigration of parents into the United States sharply reduced the availability of information from public records. For this reason, the number of available relatives was recorded for each case as a check on this source of bias in assessing familial histories for psychopathology.

Originally, the CORE CITY men were severely disadvantaged. Half of them lived in clearly blighted slum neighborhoods; half were known to five or more social agencies and a third had I.Q.'s of less than 90. Fifty percent of their homes had no tub or shower (by contrast in 1940, only 16% of Boston dwellings as a whole were without tubs or showers). Indeed, in 1940 only 30% of CORE CITY men's homes had hot water and central heat and electricity and a tub and toilet. Thirty-one percent of the parents of the CORE CITY men were in Social Class V by the criteria of Hollingshead and Redlich (1958) and over two-thirds had recently been on welfare.

By age 47 only 10% of the CORE CITY men were in Class V and as adults only 7% ever depended on welfare. While very few of their parents were in Social Class II (1%) or III (10%), by midlife over half of the men themselves were in Social Classes I-III. Although only a third of the parents of the CORE CITY men attended high school, perhaps half of the college-aged children of the men were attending college.

Like the Grant Study men, these CORE CITY Sample men were helped in upward mobility by being white, enjoying the educational opportunities of the G.I. Bill, and by the economic prosperity of the United States between 1945 and 1965. But the upward mobility of the Glueck Study men is nevertheless instructive in illustrating the paths by which men selected for low

I.Q., deteriorated neighborhood, and inner city schools mastered those difficulties.

In addition to these 456 inner city men selected for non-delinquency, it is important to note that for an average of 18 years the Gluecks also followed 500 youths remanded to reform school. This cohort consisted of virtually all Boston youths sent to juvenile reformatories for the first time between 1940 and 1944. These men were studied in the same depth as the CORE CITY Sample who served as controls. They were personally interviewed at age 25 and 32. In addition, up to age 32, careful records were assembled of their criminal histories and of their treatment by the criminal justice system. The follow-up data for this sample are well described elsewhere (Glueck & Glueck, 1950, 1968; & Glueck, 1977). These men have not been personally followed past the age of 32, but their records of psychiatric hospitalizations and arrests in Massachusetts are current up to age 45.

Nature of the Data in the Two Samples

Nobody in either sample has been permanently lost. The credit for the lack of attrition must go to Sheldon and Eleanor Glueck of the Harvard Law School and to the original staff of the Grant Study at the Harvard University Health Services who from the beginning created trusting relationships with these men. Several subjects have dropped out of the Study, but they are known to be living and most were interviewed at the age of 31. In almost all cases public records and alumni reports have allowed characterization of these dropouts' occupational success, marital status, arrest records and general adjustment over the past decade. Fifty-one men have died, but the Study of Adult Development has death certificates on almost all; and almost all were active participants in the Study until their deaths. Most of the men in both samples were reinterviewed for two hours when they were 31, and 87% of the surviving CORE CITY Sample and 100% of a randomly selected 50% sample of all the participating COLLEGE subjects were reinterviewed at age 47. At the same time almost no man has an absolutely complete data set. This is especially true of the COLLEGE Sample who were contacted at 24 different points in time.

The COLLEGE men spent their young adulthood in World War II and were last interviewed at the height of the Vietnam War. The CORE CITY men had childhoods blighted by the Great Depression, were too young for World War II and were last interviewed during the calmer era of Gerald Ford's presidency. However, with the exception of the COLLEGE men's greater "generation gap," anxiety over their war-protesting, pot-smoking adolescent children, no specific cohort effects were identified in the two samples born almost a decade apart.

Tables 1 and 2 depict the independent assessment of mental health that have been made by different raters. All the childhood ratings (rater reliability, $r = .7$ to $.8$) in Table 1 were made by individuals blind to the course of the men's lives after adolescence. If the estimate of Childhood Strengths were roughly the same for both samples, the Childhood Weaknesses of the CORE CITY Sample, not surprisingly, were many times greater than the COLLEGE Sample. The childhoods of the CORE CITY men were 20 times more likely to manifest problems characteristic of multi-problem families (Vaillant, 1980). If loving parents are not the prerogative of the socially privileged, freedom from multi-problem households is such a benefit.

In Table 1, it can be seen that the men had complete physical examinations, somatotyping and in the COLLEGE Sample unusually complete medical histories taken at the time of initial study. In many cases, projective tests were also obtained. School records were available on all the men, as were extensive family histories of mental illness, alcoholism and delinquency.

Table 2 depicts the measures of outcome obtained for both samples. In general, these variables were obtained from raters blinded to the subject's childhood. For the COLLEGE Sample, a variable like unemployment was measured in weeks, and for the CORE CITY Sample it was often measured in years. Within each sample, unemployment correlated very highly with mental health; and however concrete, such variables have the advantage of being assessed objectively.

Other mental health variables were far more abstract. For example, while representing a far more subjective judgment in both groups of men, Maturity of Coping Styles was estimated for both groups of men (Immature coping styles include hypochondriasis, dissociation, passive aggression, projection, schizoid fantasy and acting out; intermediate styles include isolation, repression, reaction formation and displacement; and mature styles include suppression, humor, altruism, and sublimation (Vaillant, 1977). The Luborsky Health Sickness Rating Scale of global mental health (Luborsky, 1963) has proved surprisingly reliable (inter-rater reliability about $.9$). A modification of Erikson's model (Vaillant, 1980) was used to assess the psycho-social maturation of the men in both samples.

Table 2 also reflects the efforts during the repeated follow-ups of the COLLEGE Sample to assess the effect of emotional factors on physical health. Thus, number of days sick leave taken per year (much more highly correlated with psychosomatic illness), and the effect of emotional stress on the body have been obtained serially for the COLLEGE Sample.

Table 3 describes variables, more or less independent of emotional health, that affect health deterioration. Such data reflect the current focus of the Study of Adult Development in trying to separate the effects of physical health on mental health and of mental health upon physical

health. For each of the 660 men in the Study, a life chart from age 15 to the latest follow-up is being completed. The top line of the life chart depicts the age at which physical health shifts from Excellent to Minor Illness, to Chronic Illness, to Disability, to Death. On lower lines, periods of underemployment, unemployment and retirement will be charted.

The CORE CITY and COLLEGE Samples possess three advantages for the longitudinal study of psychopathology. First, they all were originally selected, interviewed, rated and theoretically regarded as non-deviant controls; but at the same time, each youth was systematically investigated in a manner usually reserved for studies of psychopathology and deviance. Most of the studies of normative development have paid relatively little attention to psychopathology. Second, since the CORE CITY and the COLLEGE men entered the Study so young, they comprise a far more complete universe than does a sample of chronic clinic attendees who usually comprise studies of psychopathology. For example, the Study had clinical data on those men who died young and on those with psychopathology but had not yet come to medical or psychiatric attention. Third, the men have all been followed at multiple points in time without significant attrition due to dropping out, loss or geographic mobility. Too often in follow-up studies of psychopathology there are only "2 points on the graph."

On the negative side, the CORE CITY and COLLEGE Samples provide a very narrow sampling of human beings (male, white, American, circumscribed intellectual range and born between 1919 and 1932.) But if a narrow sampling is a failing of most surviving longitudinal studies, then the Study of Adult Development provides two very different samples that may be fruitfully contrasted with each other.

Table 1

Nature of Available Independent Variables Scored or Obtained:
Staff Blinded to Adult Adjustment

	<u>COLLEGE Sample</u>	<u>CORE CITY Sample</u>
A. Health		
1. Complete Physical Exam	Yes	Yes
2. Treadmill, EEG, Insulin Tolerance, etc	Yes	---
3. Family Medical History - 3 generations	Yes	---
4. Somatotyping & Anthropometry	Yes	Yes
5. Birth Difficulties/Childhood Health (Retrospective)	Yes	Yes
B. Social		
6. Ethnicity for 3 Generations	Yes	Yes
7. Parent's Social Class (Hollinghead & Redlich, 1958)	Yes	Yes
C. Psychological		
8. Dominant Personality Traits	Yes	Yes
9. School Records	Yes	Yes
10. Family History of Alcohol and Mental Illness for 3 generations	Yes	Yes
11. Family History of Mental Deficiency & Delinquency for 3 generations	---	Yes
12. Intelligence Tests	MAT, SAT, Army Alpha	Wechsler-Bellvue, Stanford Achievement Tests
13. Rorschach Tests	Yes	Yes
14. Psychiatric Interviews	Yes	Yes
D. Summary Rating		
15. Estimate of Future Mental Health	Yes	Yes
16. Estimate of Childhood Strengths	Yes	Yes
17. Estimate of Childhood Weaknesses	Yes*	Yes

Table 2

Available Independent Variables Assessed at Age 47+2

	<u>COLLEGE Sample</u>	<u>CORE CITY Sample</u>
A. Global Measures		
Stage in Erikson's Life Cycle (Vaillant, 1980a)	Yes	Yes
Health-Sickness Rating Scale (Luborsky, 1963)	Yes*	Yes
Adult Adjustment (Vaillant, 1977)	Yes	---
Adult Social Class (Hollingshead, Redlich, 1958)	Yes	Yes
Rating of Maturity of Coping Style (Vaillant, 1976)	50%	75%
B. Facets of Mental Health		
Mood-altering Drugs Use	Yes	Yes
Alcohol Abuse (Vaillant, 1980b)	Yes	Yes
Number of Psychiatric/Counseling visits	Yes	Yes
Sociopathy (Robins 19 point scale, 1966)	---	Yes
Social Competence (Object Relations)	Yes	Yes
Adjustment of Children Over 15	Yes	---
Marital Happiness and Stability	Yes (3 points in time)	Yes (1 points in time)
Marital Happiness and Stability (Wife's View)	Yes (2) points in time)	---
Percent of Adult Life Unemployed	Yes	Yes
Blind Rated TAT's (At age 30)	30%	---
C. Psychosomatic Variables		
Days of Sick Leave/Year	Yes (6 points in time)	Yes (1 points in time)
Presence/Absence Psychosomatic Illness	Yes	---
Effect of Emotional Stress on Body	Yes (4 points in time)	---
Recent Physical Exam Age 48+2	Yes	Yes
Recent Physical Exam Age 53, Age 58	Yes	---

Table 3

Theoretically Important Health Affecting Variables Rated at Age 47+2

	COLLEGE Sample	CORE CITY Sample
Obesity (% Overweight)	Yes	Yes
Pack/Years of Smoking	Yes	Yes (est.)
Extent of Alcohol Abuse (Cahalan Scale)	---	Yes
Diagnosis of Alcoholism (DSM III)	Yes	Yes
Combined Age at Death of Parents and Grandparents (TIAL)	Yes	---
Number of Serious Accidents	Yes	Yes
Days in Hospital (before age 47)	Yes	Yes
Exercise Patterns	Yes	---

Project Reports

- Vaillant, G.E., Brighton, J. & McArthur, C.C. Physicians' use of mood-altering drugs -- a 20-year follow-up report. New England Journal of Medicine, 1970, 282, 365-370.
- Vaillant, G.E. Theoretical hierarchy of adaptive ego mechanisms. Archives of General Psychiatry, 1971, 23, 107-118.
- Vaillant, G.E. The evolution of adaptive and defensive behaviors during the adult life cycle. Abstracted in the Journal of the American Psychoanalytic Association, 1971, 19, 110-115.
- Vaillant, G.E. and McArthur, C.C. A thirty year follow-up of somatic symptoms under emotional stress. In M. Roff & L.N. Robins (eds.). Life History Research in Psychopathology, Vol. 2. Minneapolis: University of Minnesota Press, 1972.
- Vaillant, G.E., Sobawale, N.C. & McArthur, C.C. Some psychological vulnerabilities of physicians. New England Journal of Medicine, 1972, 237, 372-375.
- Vaillant, G.E. Why men seek psychotherapy, I: results of a survey of college graduates. American Journal of Psychiatry, 1972, 129, 645-651.
- Vaillant, G.E. & McArthur, C.C. Natural history of male psychological health, I: the life cycle from 18-50. Seminars in Psychiatry, 1972, 4, 417-429.
- Vaillant, G.E. Antecedents of healthy adult adjustment. In M. Roff and D. Ricks (eds.), Life History Research in Psychopathology, Vol. 3. Minneapolis: University of Minnesota Press, 1973.
- Vaillant, G.E. and Vaillant, C.O. The relation of psychological adaptation to medical disability in middle aged men. In Successful Aging, Durham, North Carolina: Duke University, 1974.
- Vaillant, G.E. Natural history of male psychological health, II: some antecedents of healthy adult adjustment. Archives of General Psychiatry, 1974, 31, 15-32.
- Vaillant, G.E. Natural history of male psychological health III: empirical dimensions of mental health. Archives of General Psychiatry, 1975, 32, 420-426.
- Vaillant, G.E. Natural history of male psychological health, V: the relation of choice of ego mechanisms of defense to adult adjustment. Archives of General Psychiatry, 1976, 33, 535-545.

- Vaillant, G.E. Adaptation to Life. Boston: Little-Brown, 1977.
- Vaillant, G.E. Natural history of male psychological health, VI: correlates of successful marriage and fatherhood. American Journal of Psychiatry, 1978, 135, 653-659.
- Vaillant, G.E. Natural history of male psychological health, IV: what kinds of men do not get psychosomatic illness. Psychosomatic Medicine, 1978, 40, 420-431.
- Vaillant, G.E. Natural history of alcoholism I: a preliminary report In S.B. Sells and John Strauss, (eds.), Human Functioning in Longitudinal Perspective, Baltimore: Williams and Wilkins, 1979.
- Vaillant, G.E. and Milofsky, E.S. Natural history of male psychological health, IX: empirical evidence for Erikson's model of the life cycle. American Journal of Psychiatry, 1980, in press. Roy R. Grinker Lecture, Michael Reese Hospital, March 6, 1979.
- Vaillant, G.E. Health consequences of adaptation to life. American Journal of Medicine, 1979, 67, 732.
- Vaillant, G.E. Alcoholism in prognosis. In J.F. Fried & G.F. Ehrlich, A Textbook of Medical Prognosis, New York: Prentice-Hall, 1980.
- Vaillant, G.E. Adolf Meyer was right: psychiatry needs the life chart. Institute of Pennsylvania Hospital, Strecker Memorial Series, Number 16, 1979
- Vaillant, G.E. and Vaillant, C.O. Natural history of male psychological health, VII: effects of mental health on physical health. New England Journal of Medicine, 1979, 301, 1249-1254.
- Vaillant, G.E. Natural history of male psychological health, VIII: antecedents of alcoholism and "orality." American Journal of Psychiatry, 1980, 137, 181-186.
- Vaillant, G.E. The doctor's dilemma. Presented at Alcoholism Treatment: Finding New Directions, Institute of Psychiatry, London, England, April 23-27, 1979. In press, 1980.

References

- Hollingshead, A.B. and Redlich, F.C. Social class and mental illness. New York: John Wiley and Sons, 1958.
- Luborsky, L. Clinicians' judgment of mental health. Archives of General Psychiatry, 1974, 31, 292-299.
- Luborsky, L. Clinicians' judgments of mental health. Archives of General Psychiatry, 1962, 7, 407-417.
- Robins, L. Deviant children grown up. Baltimore: Williams and Williams, 1966.
- Srole, L., Langner, T.S., Michael, S.T., Opler, M.K. and Rennie, T.A.C. Mental health in the metropolis. New York: McGraw Hill, 1962.

Entre to Earlier Publications from the Grant Study

Heath, C.S. What people are. Cambridge: Harvard University Press, 1945.

Hooten, E. Young man you are normal. New York: Putnam, 1945.

Monks, J.P. College men at war. Boston: American Academy of Arts and Sciences, 1951.

Entre to Earlier Publications from Sheldon and Eleanor Glueck

Glueck, S. and Glueck, E. Unravelling juvenile delinquency. New York: Commonwealth Fund, 1950.

Glueck, S. and Blueck, E. Delinquents and nondelinquents in perspective. Cambridge: Harvard University Press, 1968.

Glueck, S. Lives of labor, lives of love. Hicksville, New York: Exposition Press, 1977.

The Duke Longitudinal Studies

Erdman Palmore
Duke Medical Center

The Duke Longitudinal Studies

Erdman Palmore

Duke Medical Center

This chapter describes the general methodological problems and procedures of the Duke Longitudinal Studies of Aging.

The Duke studies deal with "normal aging" in two senses: healthy aging and typical aging. The aged and middle-aged persons studied were relatively healthy in that they were noninstitutionalized, ambulatory, community residents who were willing and able to come to the Duke Medical Center for one or two days of tests and examinations. Second, the more common or typical patterns and problems of aging are focused upon rather than the unusual abnormalities. Findings deal with the typical physical changes that accompany aging, typical health care patterns, typical patterns of mental aging, some common forms of mental illness among the aged, the normal social roles, self concepts, satisfactions, leisure and sexual behavior, and the overall factors related to longevity.

Investigations of normal aging are of crucial importance in advancing the science of gerontology and in helping aged persons develop and enjoy a richer and longer life. When we can distinguish normal and inevitable processes of aging from those which may accompany aging simply because of accident, stress, maladjustment, or atrophy from not using skills, we can better focus our attention and efforts on those factors which can be changed and corrected.

The interdisciplinary nature of the studies is useful because aging affects many interrelated types of behavior and functioning. When specialists from different disciplines work together, the mutual stimulation, correction, and combination of perspectives can result in more accurate, thorough, and comprehensive understanding of the aging process.

Background

The First Duke Longitudinal Study of Aging was initiated in 1955 to investigate processes of aging among a panel of noninstitutionalized males and females 60 years of age and over from the time of initial observation to death. The Second Duke Longitudinal Study began in 1968 with individuals 46 to 70 years of age. These individuals were a random sample of the members of the local health insurance association stratified by age and sex. The studies were conceived as exploratory and multidisciplinary. Their organization was intended to facilitate the accumulation of the widest possible range of observations from investigators with a variety of theoretical perspectives on a common panel of subjects. Medical, psychiatric, psychological, and sociological perspectives were represented among project investigators throughout the project's existence.

No single theory of aging informed the development of the study. The goals included the generation of hypotheses as well as the testing of hypotheses of interest to a variety of investigators who brought quite different theoretical perspectives to bear on the analysis of data.

Design and Sample of the First Study

The centrality of detailed information on medical and psychological functioning which required precise measurements and laboratory evaluation presented a special problem in the development of a panel of older subjects. Two days of examination in the Duke Medical Center were required. The reported experience of survey research indicated that subjects' refusal to participate was a function of age in the later years. Since the proposed examination of panelists was costly and continued participation in the panel was vital, the Duke investigators chose to recruit volunteers for the first study rather than attempt to involve a randomly drawn sample of noninstitutionalized subjects.

A snowball technique was used to create a large pool of volunteers 60 years of age and older. From this pool a panel was selected that reflected the age, sex, ethnic, and socioeconomic characteristics of the older population in Durham, North Carolina. The purpose of this procedure was to maximize the variety of controls which might be introduced subsequently, not to give the appearance of random sampling without its substance. The primary inducement originally offered was a free annual physical examination. A description of the timing of the examinations and panel attrition are summarized in Table 1.

Most panelists seem to have been motivated by the medical examination, but an equally important factor in maintaining the panel appears to have been the opportunity for participants to obtain satisfying social recognition regularly. This recognition took the form of greeting cards on special occasions, occasional contact by telephone with the social workers, and solicitous attention on those days when the panelists were brought to the medical center for observation.

A high degree of identification with the project was reflected in a relatively low rate of attrition for reasons other than by death or serious illness throughout the twenty years of the study. After the second round of examinations, in which 15 persons refused to continue, refusals were minimal.

Several methodological problems associated with this design are immediately obvious. Generalizations from the data necessarily are limited to statements concerning the relationship among factors in the aging process under specified conditions. Statements based on data from this panel about the distribution of characteristics or configurations of characteristics among older people generally are not warranted. Statements about certain aging processes with each panelist used as his or her own control are warranted. Moreover, the length of time between observations is not identical for all panelists. Where this factor may be relevant, precise information about intervals between observations is available for each panelist.

For each panelist, approximately 788 pieces of information were coded in each series of observations for each subject. Of these, 336 were medical; 109, psychiatric or neurological; 109, psychological; and 234, social. A summary of observations is found in Table 2. Medical determinations included history and current status of systems, with special attention being given to cardiovascular ratings, serology, ophthalmology, and audiology. A summary rating of physical functioning was developed for estimating the presence or absence of pathology and associated degree of disability. An index estimating cardiovascular functioning has been developed.

The psychiatric or neurological variables focus on aspects of mental status and functioning, both current and historical, among panelists. Detailed electroencephalograph (EEG) evaluations supplement the findings of the neurological examination.

Psychological variables include Wechsler Adult Intelligence Scale (WAIS) scores, assessment of reaction time in a variety of learning experiments, and summaries of Rorschach responses.

Sociological variables include life style indicators, work status and history, family relationships, and level of activity and of life satisfaction, including an adaptation of the Cavan Activities and Attitudes scales.

The Second Study

The second longitudinal study, or Adaptation Study, was begun in 1968 with 502 persons 46-70 years of age (see Table 3). This study was initiated because of growing awareness that the aging process is not confined to persons over 65. The aging processes were recognized as extending over the entire life span, with special problems of adaptation

during middle age and the preretirement years. Potentially stressful events such as death of spouse, serious illness, menopause, children leaving home, and preparation for retirement often occur before the age of 65. The adaptation study was designed to study responses to such events and the aging processes in general during the ages of 45 through 69. More specifically, the study was particularly interested in the factors or mechanisms which contribute to "successful" aging, defined in a variety of ways.

The study was also designed to be interdisciplinary, to be representative of the broad middle and upper socioeconomic levels of our society, and to allow cross-sectional, longitudinal, and cross-sequential types of analysis. It is an interdisciplinary study because of its inclusion of several types of investigators and of variables representative of the physical, psychological, and social aspects of adaptation to aging. The sample was designed to be representative by drawing on a random sample of all members of the major health insurance association residing in the Durham area, stratified into ten age-sex cohorts. The design allows cross-sectional analysis between the ten age-sex cohorts, longitudinal analysis over a period of six years, and cross-sequential analysis comparing different age cohorts at different points in time.

Variables

The major variables investigated are summarized in Figure 1, "Diagram of Variables for the Second Duke Longitudinal Study." As this diagram indicates, there are three main types of independent variables: physical, psychological, and social. These are conceived of as the resources which should be related to how well a person adapts to aging and to stressful events. A person's physical resources include current mobility and general health, as well as more specific musculoskeletal functioning, cardiovascular capacity, and ability to see and hear. Psychological resources include intelligence and personality characteristics such as tendency to be outgoing, emotionally stable, assertive, conscientious, imaginative, experimental, self-sufficient, etc. Another psychological resource is the ability to concentrate and respond quickly and accurately under stress conditions (Vigilance Test). Abnormal electroencephalograph recordings indicate possible limitations on a person's mental resources. The person's various social resources are measured by counting the number of different social roles that a person is engaged in; the number of role partners he or she interacts with; the number and type of social contacts per month; the number of hours per week spent in various activities; socioeconomic resources such as education, income neighborhood, and parental status; concept of who and what kind of person he or she is, what kind of person he or she would like to be; what kind of life he or she has had and expects to have; and knowledge about the attitudes toward community resources. Attitudes toward personal health and willingness to take the sick role, as well as attitudes toward time, death, achievement,

planning, and internal control may also be important resources in adaptation.

There are two main types of adaptation measures: performance and intrapsychic reactions. Whether or not a person is able to perform so as to maintain financial and social independence is considered an important measure of adaptation. The amount of time spent in productive activities and social activities are other measures of performance. Reaction time and accuracy are a measure of a person's ability to respond to external stimuli. The intrapsychic measures of adaptation are designed to measure whether or not a person is satisfied and happy with his or her adaptation or whether he or she has become more depressed with feelings of meaninglessness, powerlessness, normlessness, and isolation. The psychosomatic symptoms are an indirect measure of a person's adaptation and have been found to be related to unhappiness, psychoneurosis, and other indicators of poor adaptation. Finally, a person's self-concept should be related to how well he or she has adapted.

During the eight hours of examination and interviews, data on more than 1000 variables were recorded and entered on electronic tape for each person.

Sample

In order to be fairly representative of the broad middle and upper socioeconomic levels in Durham, the sample was drawn from the membership list of the major health insurance association in Durham County. This membership list includes the majority of middle- and upper-income-level Durham residents. The list underrepresents the lower class and the poverty stricken, but the purpose of this study was to concentrate on the large "middle mass" rather than attempting representative numbers of the poor, the chronically ill, the illiterate or those in physical or mental hospitals. Illiterate persons were excluded because they could not take the written tests. Institutionalized or homebound persons were excluded because they could not come in for the examinations. Blacks were also excluded from this study because they were being studied separately in another project.

An initial 2000 names and addresses were randomly drawn from the membership lists to serve as a pool for persons to be interviewed and tested. Individuals were contacted and scheduled in a random order until enough subjects had been interviewed and tested to fulfill the target number in each age-sex cohort. Persons were first mailed an introductory letter explaining the purpose and procedures of the study and then were contacted by telephone or face-to-face for those without telephones.

Because the study requires eight hours of examinations and interviews without compensation offered to the subjects (other than a free medical

examination and laboratory tests), about half of the persons contacted either refused to participate or did not schedule a day for the examinations before the target number of subjects for that age-sex cohort had been reached (see Table 4). The remaining persons in the initial pool were either ineligible because they lived outside of the Durham area, were in the wrong age bracket, were Negro, illiterate, housebound, etc., or were not contacted before the target number had been reached.

Because about one-half of the persons contacted were not interviewed and tested, an important question is the degree of representativeness of our interviewed subjects compared to the noninterviewed persons. We carried out a special study of noninterviewed persons in order to estimate how representative our sample is. A random sample of 100 noninterviewed persons was drawn and minimum information on their age, sex, education, occupation, marital status, and health was secured by telephone or home visits. We also obtained tabulations from the health insurance association on the hospital and medical care usage of both the interviewed and non-interviewed persons. The analysis of these data yielded no significant differences between those interviewed and those not interviewed, except that those not interviewed tended to more often be in the extreme health categories (excellent health, poor health) and less in the middle categories (good health, fair health). Thus, while there probably are some undetected differences between the interviewed and noninterviewed, there does not appear to be any gross demographic or health bias in the sample of those interviewed.

A special problem was encountered in the 65-69 year age group. Persons in this age group who are still members of the health insurance association are a somewhat elite group, because they are insuring themselves against only those health care costs not covered by Medicare. These persons are about one-half of all persons over 65. In order to make our sample in this age category more representative we included thirty-two subjects randomly drawn from persons who had been in the Duke Hospital during 1968-1969. It was thought that hospitalized persons would complement the Blue Cross - Blue Shield sample by more nearly representing those in poorer health. As it turned out, there were small differences between the Duke Hospital group and the health association group over 65, with the Duke group having somewhat higher income, education, and intelligence scores. Therefore, our sample of persons 65-69 is somewhat above average and not as representative as the rest of our sample. This age category should be treated with caution in cross-sectional comparisons.

It may be worthwhile to spell out some of the standard cautions one should bear in mind when evaluating these data. First, the samples studied are somewhat unrepresentative of the universe from which they were drawn. The First Longitudinal Study was drawn from a pool of volunteers to approximate the demographic composition of Durham, but they were volunteers nevertheless and therefore differed in various ways from a strict cross-section of Durham. The Second Longitudinal Study was based on a probability sample of members of the local health insurance

association, but we were able to get the participation of only half of those eligible. While the participants appear similar to nonparticipants in basic characteristics, they undoubtedly differ in some ways. In general, the two samples appear to underrepresent the poorer, sicker, and less able members of the community. We can only guess how our findings might change if we were able to get more representative samples.

Second, the universe from which the samples were drawn was limited to fairly normal middle-aged and older persons living in the Durham area. We do not know how well our findings might apply to other geographical areas or to less normal persons such as the institutionalized, although some of the reports discuss this through comparison with other studies.

Third, there are usually errors of measurement, coding, and processing that creep into the data despite our best efforts to catch and correct them. There is also the problem of interpreting some chance fluctuation as significant, or on the other hand, of ignoring an important finding because the numbers were small.

A last caution is the old maxim: "Association does not prove causation." We had no classical experiments in these studies because we wanted to study the aging process in its natural state. On the other hand, it is true that associations which persist with proper controls do provide evidence that a causal relationship may exist.

Themes

When we review the specific findings on normal aging, some general themes emerge which tie together findings from several substantive areas. The same themes emerge from both studies.

Advantages of Longitudinal and Interdisciplinary Study of Aging.

The longitudinal method analyzed changes in the same persons over time in dermatology, ophthalmology, hearing, health care after Medicare, physician and self-health ratings, cognitive functioning, EEG frequencies, functioning and adjustment after widowhood, and self-concepts. The various analyses of factors related to longevity also required a longitudinal design. The advantages of interdisciplinary collaboration is evident in the fact that most of these reports involve the joint analysis of variables from two or more areas: biology, physical medicine, psychiatry, psychology, and sociology.

Patterns of Declining Health and Physical Functioning

This theme is repeated with some variations in the overall declines in immunoglobulin concentrations, skin conditions, vision, hearing, health rating, and sexual activity. This confirms the expected pattern of normal aging.

Exceptions to Physical Decline

Despite the overall declines as measured by averages or group percentages, substantial minorities show no decline and may even have improvement in various measures of skin conditions, vision, hearing, health ratings, vitamin levels, and sexual activity. Thus, the process of physical aging is not necessarily an irresistible and irreversible force. Health and functioning can and do improve for some older persons just as they do for younger persons.

Little or No Decline in Social and Psychological Functioning

This theme is in direct contrast to the theme of declining physical functioning. The first study presented several findings showing little or no decline in activities, attitudes, cautiousness, recall, and general adjustment. The second study reports only small overall declines in intelligence scores (especially among those aged 60-69 and those free of hypertension), reaction times, correct signal detections, and EEG frequencies. It is impressive how many of the normal aged are able to compensate for their growing physical handicaps and maintain fairly stable levels of social and psychological functioning. This evidence tends to refute the disengagement theory that decline and withdrawal are typical, inevitable, and normal.

Wide Variation in Aging Patterns.

This theme takes two forms: documenting the range of individual differences which group averages obscure, and attempting to account for individual and group differences. The aged usually show more individual variability than the young. This is shown dramatically in the report on immunoglobulin concentrations and is reflected in most other comparisons of the old and young. Thus, even when there are significant mean differences between the old and young, there is usually considerable overlap between the two distributions. This means that many older persons have better functioning and better scores than many younger persons. These studies attempt to account for this individual variation by analyzing differences by sex,

race, socioeconomic group, health practices, chronic and acute illness, retirement, children leaving home, etc. More complete discussion of findings are reported by Palmore (1970, 1974).

Thus, the complex and difficult enterprise of exploring and charting the territory of aging continues. We believe this enterprise will eventually allow more older people to safely reach "the promised land" of a better and longer life.

Table 1
Participants in First Longitudinal Study by Round

Round #	Dates	Total Participants	Number with:		
			Physicals	Psychologicals	Social Data
1	3/55-5/61*	270	268	267	268
2	9/59-5/61	183	181	182	182
3	1/64-3/65	178	138	140	172
4	10/63-6/67	138	109	110	136
5	4/68-1/69	110	92	93	107
6	2/70-8/70	108	94	92	101
7	1/72-5/72	81	65	60	74
8	2/73-8/73	68	62	57	54
9	4/74-9/74	57	55	52	55
10	12/74-8/75	56	53	47	51
11	3/76-8/76	44	42	41	41

*Ten of the participants in the first round had their first examinations after the second round of examinations had started for other participants.

Table 2

A Summary of Longitudinal Observations

Medical history (original and interim)	Laboratory studies
Physical examination	Urinalysis
Neurological examination	Blood morphology
Mental status	Blood chemistry
Depression and hypochondriasis	Serologic test for syphilis
Dermatological examination	Cholesterol
Ophthalmological examination	Urea nitrogen
Visual fields	Medical summary
Acuity	Psychological data
Color perception	Rorschach
Depth perception	Aspiration level (TAT)
Color photographs	Wechsler Adult Intelligence Scale
Audiometry	Reaction time
Pure tone	Social history and information
Speech threshold	Retirement data
Electroencephalogram	Activities
Chest x-ray	Attitudes

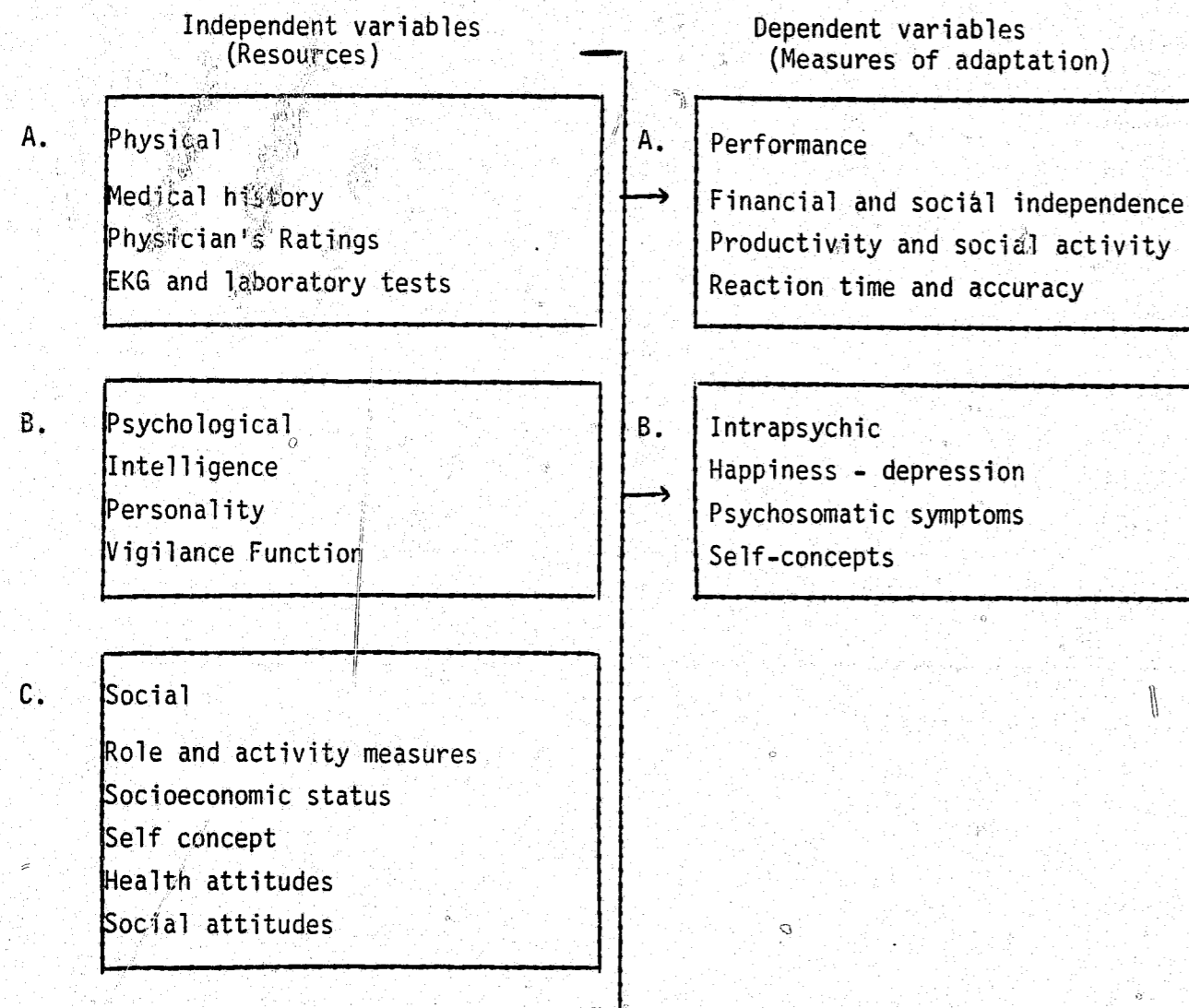
Table 3
Participants in Second Longitudinal Study by Round

Round #	Dates	Total Participants	Number/with:		
			Physicals	Psychologicals	Social Data
1	8/68-4/70	502	502	502	502
2	8/70-3/72	443	438	438	443
3	6/72-7/74	386	384	383	383
4	8/74-10/76	375	374	375	375

Table 4
Response Rate by Age and Sex

Cohort	Total contacted and eligible	Inter-viewed	Noninter-viewed	% inter-viewed
Males				
45-49	92	47	45	51
50-54	95	49	46	52
55-59	96	52	44	54
60-64	95	55	40	58
65-69	113	58	55	51
Females				
45-49	94	45	49	48
50-54	107	46	61	43
55-59	103	47	56	46
60-64	111	50	61	45
65-69	144	53	91	37
Total	1050	502	548	48

Figure 1.
Diagram of variables for the
Second Duke Longitudinal Study



References

- Palmore, Erdman (Ed.) Normal Aging: Reports from the Duke Longitudinal Study, 1955-1969. Durham: N. C.: Duke University Press, 1970.
- Palmore, Erdman (Ed.) Normal Aging 2: Reports from the Duke Longitudinal Studies, 1970-1973. Durham, N. C.: Duke University Press, 1974
- Palmore, Erdman. Social Factors in Normal Aging. Durham, N. C.: Duke University Press, in press.

Older Workers in Rural Areas*

Edward A. Powers

Willis J. Goudy

Patricia M. Keith

Iowa State University

*This research was supported in part by grant #10-P-57495/-01 from the Social Security Administration, Washington, D.C.

Older Workers in Rural Areas

Edward A. Powers

Willis J. Goudy

Patricia M. Keith

Iowa State University

Problems associated with adaptation to occupational withdrawal in later life are relatively new phenomena unique to industrial societies (Donahue, Orbach, & Pollak, 1960; Orbach, 1963; Friedmann & Orbach, 1974). Potentially, one of the more traumatic events for some older individuals is severance from the work force. Entry to retirement may have both positive and negative consequences, although the latter tend to be covered more frequently in the popular press.

To enhance the lives of older persons or provide opportunities for self-help, increased understanding of the processes occurring in later life is necessary. We were interested in following workers as they approached and passed those ages commonly thought to be most appropriate for retirement. To view work and withdrawal in later life, we examined older Iowa males initially studied in 1964. Reinterviewed in 1974, these respondents provided information on three categories of occupational involvement: full-time employment during the ten-year period, part-time employment (decreased involvement in the work force), and withdrawal from the labor force. The relationships between employment status, work and retirement attitudes, sociodemographic characteristics, and perceived adaptation were explored in this panel study. Although the importance of panel data cannot be overemphasized, two types of data are available from our study. Data reflecting patterns of occupational involvement and changes in selected social, social psychological, and community variables were obtained in both interviews. Secondly, cross-sectional data on variables included in the second interview but not in the first provided additional information. (This second category may seem inappropriate to those not familiar with what could be called the "longitudinal lament;" the problem of researchers never anticipating all the "right" questions in any phase of a longitudinal study.)

Our interest in older people as they face the decision to retire or continue working was precipitated, in part, by discussions of retirement policy and by the increasing numbers of people nearing the prospect of

retirement. The age structure of the population, in conjunction with changes in the technological, social, political, and economic segments of society, has resulted in retirement becoming a feasible option for most people. Indeed, there has been a decline in the average time spent in the labor force during the last 25 years (Fullerton & Byrne, 1976). Increases in longevity and more attractive pension plans have increased the number of potential retirees.

Study Design and Sample Characteristics, 1964 and 1974

The 1964 Study

In 1964 two sociologists at Iowa State University, Ward Bauder and Jon Doerflinger (1967), established benchmark data for a proposed longitudinal study of the patterns of withdrawal from occupational roles. They sampled older men living in or near nonmetropolitan communities. Although information was gathered on many aspects of the lives of the respondents, the major concern of that investigation was to establish trends in involvement in individual activities over the previous ten years and anticipated changes for the next five years as well as to study work and retirement attitudes. It was proposed that respondents would be restudied periodically, culminating in a ten-year restudy of each panel member.

Specifically, the study was designed to examine Iowa males, 50 years of age or older, living in or near towns with populations between 2,500 and 9,999. Respondents were employed full-time in one of five occupational categories--farmers, factory workers, owner-merchants, salaried professionals, and self-employed professionals. Therefore, the universe of this study was limited to 76 of Iowa's communities with populations between 2,500 and 9,999 according to the 1960 census. Three other communities in this size category were excluded because they were near metropolitan areas and, thus, were not considered comparable to the other 76. These towns were grouped into 12 strata according to geographical location and size. The state was divided into four quarters; within each section, towns were divided into groups (2,500-4,999; 5,000-7,499; 7,500-9,999). This insured selection of eligible participants from throughout the state.

Estimates of the population of men 50 years of age and older in each of the five occupational classes were made for each of the 76 cities. Sampling rates were obtained from these estimates that would yield approximately 500 interviews in each occupational class. For each occupation, the sampling rate within the sample communities in each stratum was determined so that (probability in town in sample) x (sample rate in town) equalled the desired sampling rate for this occupational group. Thirty-three communities were selected randomly from the 76 city universe for inclusion in the original sample and a uniform state-wide sampling rate was determined for each occupational group.

Farmers were sampled from the area immediately surrounding the sample cities. Generally, this consisted of two townships for the smallest cities, three for the medium-sized towns, and four for the largest category. The Master Sample of Agriculture Materials maintained by the Iowa State University

Statistical Laboratory was used to draw a cluster sample of farmers for each rural area at the appropriate within-community rate. The number of farmers interviewed near any one community varied from three to 14. Of the total interviewed, more than three-fourths were farm owners while the remaining one-quarter consisted of farm tenant-managers.

The group labeled factory workers was identified in the following manner:

The sample of factory workers was obtained by dividing the factories into two size groups, those employing 100 or more workers and those employing less than 100. All factories employing 100 or more persons were included in the sample . . . with a sample of interviewees selected from each at a rate equal to the appropriate within-town sampling rate. Factories employing less than 100 were sampled at a rate equal to the appropriate within-town sampling rate and all eligible persons in these factories were included. (Bauder & Doerflinger, 1967, p.7)

The actual occupations listed by respondents, however, suggest that the label given to this sample group is misleading. Examples from the craftsmen, foremen, and kindred census category are included (foremen in metal industries and others, typesetters and printing pressmen, glaziers, machinists and tool-and die makers, welders, assemblers and inspectors, packers, truck drivers); also included were operative (in lumber, food and dairy product manufacturing, farm and other machinery, fabricated metal, grain-mill production, for example), laborers (in farm and other machinery manufacturing, wholesale and retail trade, food, cement, for example), and even service (janitors, watchmen) and clerical (shipping clerks, stock clerks) employees. And, close examination of the initial sample suggest that most of these job holders were employed by small concerns rather than factories containing large numbers of workers performing essentially the same tasks. Thus, we changed the title of this category of sampled employees to blue-collar workers because it includes the diversity of jobs often listed as blue-collar occupations for males. Certainly, this portion of the sample is not restricted to the jobs normally called to mind when factory work is considered. Of course, few factories of the scale found in urban areas are located in small towns, especially in Iowa.

Samples of the two professional groups (salaried and self-employed) and the owner-merchants were drawn from lists compiled from city directories, phone books, and personal informants. The owner-merchants were businessmen involved in construction, manufacturing, food and dairy products, eating and drinking establishments, hardware and building materials, apparel, furniture, motor vehicles and farm machinery, service stations, insurance and real estate, and variety stores. Generally, these were proprietors--men who owned their businesses and who would be primary candidates for membership in local commercial clubs.

Estimates of professionals based on census figures were found to be too high shortly after the initial interviewing began. It became evident that 33 towns would not yield samples of the professional groups of sufficient size. Thus, nine additional cities were added at random, but only professionals were interviewed in these communities. Salaried professionals,

included accountants, clergymen, college professors and public school teachers, engineers (civil, mechanical, other), social workers, public administrators (federal, state, local), and some officials in manufacturing and food retailing. The sample of self-employed professionals contained physicians, osteopaths, chiropractors, dentists, pharmacists, morticians, lawyers, and veterinarians. In some instances, members of the same profession were listed under both categories; that is, although most lawyers were self-employed, some worked for a firm and were considered to be salaried employees. A decision was made for each respondent and he was listed in only one category for data analysis.

The original investigators reported that 1,922 interviews were completed in 1964. Further inspection by the 1974 research team showed, however, that 52 men had been erroneously screened by the interviewers. Therefore, these 52 subjects were deleted from the data set because they were not fully employed in 1964, a requirement of the original sampling design. Some had already retired and some had partially retired at the time of the 1964 interview, facts that were verified whenever possible in the 1974 interviews. Thus, the final sample size for the 1964 interviews was reduced from 1,922 to 1,870 respondents. Also, it should be noted that the occupational categorization of 28 respondents was changed after we reviewed the 1964 data on a case-by-case basis. Although a rationale for the previous occupational designations could have been constructed, we felt that additional evidence from the 1974 interviews coupled with the 1964 data forced us to revise some classifications. This occurred primarily in the owner-merchant and self-employed professional groups. For example, many pharmacists in small towns own the drug stores in which they work. Thus, should they be called self-employed professionals or owner-merchants? We elected to classify them as professional because of the training that was necessary to practice their profession. Also, some pharmacists actually had been employed by others, so we changed the designation to salaried professional for these. These examples indicate the difficulty in assigning job holders to general categories; our problems were compounded by the fact that this was a panel study in which the original investigators no longer were present.

A schedule was constructed for face-to-face interviews in 1964. Efforts were made to gather work histories, household mobility and living arrangements, and visiting patterns with friends and relatives. Also, in anticipation of later interviews, questions were asked about projected changes for five years in the future (1969). (Because the projected 1969 data collection was not completed, the data on anticipated change must be used with caution as the questions were limited specifically to change over a five-year period.) Also, in the 1964 interview, general sociodemographic questions were included, with additional attention given to income, net worth, and health, areas thought to affect the decision to work or retire in later life. Finally, a few attitude items were included on work satisfaction, retirement, morale, and central life interests. Table 1 displays categories of data collected in 1964 as well as in 1974.

Interviews were conducted by personnel hired, trained, and supervised by the Survey Division of Iowa State University's Statistical Laboratory. The interviews, conducted by women who were primarily middle-aged, averaged one hour in length. In the largest town, two or more women conducted the

interviews; in the smallest ones, one woman was assigned to complete all the interviews.

Our reconstruction of the 1964 study suggests a refusal rate of about eight percent, although there is some difficulty in calculating it. Substitution was made when originally selected individuals were no longer present in the community, were not fully employed, or were not 50 years of age or older. Thus, the actual refusal rate was probably lower, perhaps less than four percent.

Table 1:
Categories of Available Information and Scales Used in 1964 and 1974

	1964	1974
Categories of Data Available:		
Occupation	x	x
Work Status	x	
Work History	x	x
Characteristics of job	x	x
Attitudes toward job and work	x	x
Attitudes toward retirement	x	x
Residence patterns and change	x	x
Attitude toward community	x	x
Attitude toward neighborhood	x	x
Marital status	x	x
Household composition	x	x
Family interaction (children, grandchildren, sibs)	x	x
Visiting with friends and neighbors	x	x
Confidants		x
Household division of labor	x	x
Organizational involvement and change	x	x
Community involvement	x	x
Helping patterns and dependency		x
Health	x	x
Political attitudes		x
Leisure attitudes		x
Self perceptions		x
Anomia		x
Stereotypes about old age		x
Morale	x	x
Life satisfaction		x
Income and net worth	x	x
Widowhood experience		x
Perceptions of death		x
Religious activity		x
Specific Scales:		
Job autonomy	x	x
Job satisfaction	x	x
Retirement attitudes	x	x
Retirement adjustment		x
Stereotypes about old age		x
Political intervention		x
Political cynicism		x
Political participation		x
Anomia		x
Work ethic		x
Life satisfaction		x

The 1974 Restudy

In 1970, we were given the opportunity to work with the data gathered in 1964. The original project directors had left, one copy of a final report was located, and data and sample information were found. Location of study materials was not without some effort, however. For example, the 1964 survey instruments were located in a repair shed on one of ISU's farms. Needless to say, the paper had deteriorated somewhat, but we found everything we needed to reconstruct the original study. We were able to recreate and use the data set (Powers & Goudy, 1971; Powers, 1971). Then we entertained the idea of completing the interviews at the end of the ten years as originally planned by Bauder and Doerflinger.

A proposal to obtain additional information from survivors was drafted, revised, and accepted by the Social Security Administration in 1973. After receiving the grant, the first task was to locate survivors. An intensive search procedure was carried out using the 1964 addresses. In addition, 1966 addresses were available for most respondents from a mail survey of the original sample conducted but not analyzed by the initial researchers. This survey repeated some questions in the 1964 questionnaire and asked about changes over the two-year period. Analyses of these materials were never located, but we have published from them (Goudy, Powers, & Keith, 1975; Goudy, 1976). Because the questions were limited, data from the 1966 questionnaire will not be included in this report. First, names of the subjects were sent to local postmasters, who verified the address or supplied corrected information when available. Persons not located through this procedure were traced by using selected information from the previous interviews. The names of organizations to which respondents belonged, their church affiliations, and children's names were used in the search efforts. Also, calls to city clerks and other knowledgeable local people assisted in tracing some people. Finally, talking with former neighbors or others thought to have useful information, the interviewers attempted to track those few individuals we had trouble locating.

Efforts to identify the current addresses of living respondents were successful. Only 31 of the 1,870 potential respondents (1.7%) were never found, a figure lower than those reported in many other studies involving tracing efforts (Eckland, 1968). Persons reported to be deceased were listed and a check of death certificates was made at the appropriate state vital statistics offices. Reports were not considered verified until the death certificate was located or the death was confirmed by at least two knowledgeable individuals. These measures were taken because we found that some reports of death were "highly exaggerated;" that is, a few individuals were living even though some source suggested they had died years earlier. Because of our accounting system, then, some of the 31 individuals we never located probably were deceased by 1974 although they were not counted in that category.

The 1964 instrument was used as the base for a revised 1974 interview schedule. Most items of the 1964 instrument were retained to permit accurate assessment of the continuity or change during the decade in occupational involvement, physical condition, financial status, social networks, community involvement, attitudes toward work and retirement, morale, and central life interests. In an effort to incorporate more appropriate measures of the 1974 status of respondents on selected variables used in 1964 (e.g., morale,

physical condition, sources of income), items reflecting advancements in the fields of gerontology and research methods also were constructed (Back & Bourque, 1970; Rosencranz & Pihlblad, 1970; Smith & Lipman, 1972; Smith et al., 1969; Tissue, 1968). Finally, additional measures were included only in the second phase of the study (e.g., life satisfaction, anomia, disability, utilization of community support services, and retirement adaptation).

In the 1974 interview schedule, separate sections were constructed for four possible groups: fully employed, employed part-time, retired, and unemployed. Because only eight individuals classified themselves as unemployed, we have deleted them from analyses. Although most questions were asked of each group, verb tense varied from past to future. (See Table 1 for data categories.)

Again, female interviewers were hired, trained, and supervised by the Survey Division of the Iowa State University Statistical Laboratory. These interviewers provided another source of data. They were requested to complete a questionnaire which contained many of the attitude items they asked in their interviews; this was filled out before they knew anything about the sample. Those who conducted ten or more interviews completed a second questionnaire with the same attitude items. This will permit us to examine the relationship between respondent and interviewer attitudes and the effects of the perceived rapport. These research questions, among others, have not been addressed at this time. Administering the interview schedule averaged about an hour and forty minutes, although some continued for as long as four and one-half hours. 1,332 interviews were completed with the 1,870 original panel members. The largest group of those not interviewed was the deceased, who totaled about 22 percent of the initial panel. Refusals were received from 70 individuals or about five percent of those eligible for reinterviewing after the deceased were dropped from the calculations.

The research design used in our study is important because it provides panel data on a sample of person who have been overlooked in research on older workers and the retirement process (Goudy, Keith & Goudy, 1977; Goudy, Powers, & Keith, 1977). We view the facts that initial respondents were from small towns, reflect a rather large age span, and are members of selected occupational categories as advantages setting this study apart from others.

Comments

The sample included fully employed men who in 1964 were 50 or older and in one of five occupational groupings; these and other characteristics of this sample certainly necessitate cautious generalization to all rural men. The sample did consist of fairly large subsamples of men in five occupational categories that seldom are studied. The five subsamples were representative of the types of employment that affect late life transitions in rural areas--the self-employed and salaried or wage earner, the professional and non-professional, the nonagricultural and agricultural worker. The study, however, did not have many men whose life situation forced early retirement or who were in unskilled occupations.

Overall, for these rural men, life had been good during the previous decade. Occupation, more than either age or work status, affected life situations. Even for those who had retired, the transition had not been exceedingly difficult.

References

- Atchley, R.C., The sociology of retirement. Cambridge, Mass.: Schenkman, 1976.
- Back, K.W., & Bourque, L.B. Life graphics: Aging and cohort effect. Journal of Gerontology, 1970, 25, 249-255.
- Bauder, W.W., & Doerflinger, J. Patterns of Withdrawal from Occupational Roles among Older Men. Unpublished final report for the Cooperative Research and Demonstration Grants Program, Department of Health, Education and Welfare, Washington, D.C. (data approximate). 1967.
- Donahue, W., Orbach, H., & Pollak, O. Retirement: The emerging social pattern. In Clark Tibbitts (Ed.), Handbook of Social Gerontology. Chicago: University of Chicago Press. 1960, 330-406.
- Eckland, B.K. Retrieving mobile cases in longitudinal surveys. Public Opinion Quarterly, 1968, 32, 51-64.
- Friedman, D.A., & Orbach, H.L. Adjustment to retirement. In S. Arieti (Ed.), American Handbook of Psychiatry (vol. 1, second ed.) New York: Basic Books, 1974, 609-645.
- Fullerton, H.N., & Byrne, J.J. Length of working life for men and women, 1970. Monthly Labor Review, 1976, 99, 31-55.
- Goudy, W., Powers, E.A., & Keith, P.M. Work and retirement: a test of attitudinal relationships. Journal of Gerontology, 1975, 30, 193-198.
- Goudy, W. Nonresponse effects on relationships between variables. Public Opinion Quarterly, 1976, 40, 360-369.
- Goudy, W., Keith, P.M., & Powers, E.A. Older Workers in Small Towns: Occupational Involvement and Adaptation. Sociology Report 140, Ames, Iowa: Department of Sociology and Anthropology, Iowa State University, 1977.
- Lawton, M.P. The elderly in low density areas: Future Environmental research needs and methods. In R. Atchby & T. Byerts (Eds.), Rural Environments and Aging. Washington, D.C.: Gerontological Society, 1975.
- Lowenthal, M., Thurnber, M., & Chiriboga, D., Four Stages of Life. San Francisco: Jossey-Bass Publishers, 1975.
- Orbach, H.L., Social values and the institutionalization of retirement. In R.H. Williams, Tibbitts, & Donahue (Eds.), Processes of Aging (vol. II) New York: Atherton, 1963, 389-402.

Powers, E.A. The effect of the wife's employment on household tasks among postparental couples: a research note. Aging and Human Development, 1971, 2, 384-387.

Powers, E.A., Keith, P., & Goudy, W.J. Family relationships and friendships. In C. Atchley & T.O. Byerts (Ed.), Rural Environment and Aging. Washington, D.C.: Gerontological Society, 1976, 67-90.

Powers, E.A., Keith, P., & Goudy, W.J. Later Life Transitions: Males in Rural America. Sociology Report 139, Ames, Iowa: Department of Sociology and Anthropology, Iowa State University, 1977.

Powers, E.A., & Goudy, W.J. Examination of the meaning of work to older workers. Aging and Human Development, 1971, 2, 38-45.

Rosencranz, H.A., & Pihlblad, C.C. Measuring the health of the elderly. Journal of Gerontology, 1970, 25, 129-133.

Smith, K., & Lipman, A. Constraint and life satisfaction. Journal of Gerontology, 1972, 27, 77-82.

Smith, P.C., Kendall, L.M., & Hulin, C.L. The Measurement of Satisfaction in Work and Retirement: A Strategy for the Study of Attitudes. Chicago: Rand McNally, 1969.

Tissue, T.L. A Guttman scale of disengagement potential. Journal of Gerontology, 1968, 23, 513-516.

The Longitudinal Study of Transitions*

David A. Chiriboga, Ph.D.

Human Development and Aging Program
University of California, San Francisco

*This research is supported by the National Institute on Aging, Grant No. AG00002 (formerly by National Institute of Child Health and Human Development, Grants HD 03051 and HD05941).

The Longitudinal Study of Transitions*

David A. Chiriboga, Ph.D.

Human Development and Aging Program
University of California, San Francisco

Introduction

Over the past 21 years the Human Development and Aging Program of the University of California has conducted three longitudinal studies, two of which are still ongoing. The first was a comprehensive longitudinal study of 534 hospitalized mentally-ill aged individuals and 600 community-residing aged individuals. Initiated in 1958, the Geriatric Research Project (as it came to be called) found evidence that the individual's modes of adjustment to the traumas of later life are not unique to the aging process. They are influenced by "preferred" styles of coping, often adopted very early in life, and subsequently modified or elaborated upon to handle the stresses of later life (Lowenthal, 1964; Lowenthal, Berkman and Associates, 1967; Lowenthal and Zilli, 1969; Simon, Lowenthal and Epstein, 1970). The research also revealed an unexpectedly broad range of patterns of maturing and aging in the dominant culture, as well as in various subcultures of the urban setting, many of them both satisfying to the individual and socially productive (Clark and Anderson, 1967; Lowenthal, Berkman and Associates, 1967). The findings, as well as many unresolved issues, led to a study of transitions which is the focus of the present chapter. This study of transitions was designed to fill in the vast gaps in our knowledge of the adaptive and maladaptive processes of the early, middle and later adult years. A final study, begun in 1976, examined adult men and women in the process of divorcing their spouse. It incorporates much of the theoretical and empirical perspectives of the transitions study (see Chiriboga and Cutler, 1977; Chiriboga, Roberts and Stein, 1978; Chiriboga, Coho, Stein and Roberts, 1979; Chiriboga and Thurnher, in press; in 1980 a three year follow-up of the divorce sample was completed.

The Longitudinal Study of Transitions

The general objective of this longitudinal study has been to delineate processes of psychosocial change across the adult life course, and to assess the comparative influence of sociocultural, situational, and personal factors on adaptation at successive life stages. More specifically, the study seeks to determine how men and women within the dominant value system of our society cope with both the normally expectable and idiosyncratic transitions of adulthood, and to isolate factors, intrinsic or extrinsic to the individual, which contribute to continued personal growth, stagnation, regression, or mental illness.

A quasi-experimental research design forms the basis of the research effort. All subjects were selected prior to their passage through one of four major normative transitional periods; they included high school seniors facing departure from the family home, newlyweds facing the prospect of raising a family, middle aged parents facing the departure of children from the home, and men and women facing retirement. The timing and occurrence of the expected transition is one of the major "treatment" conditions of the research. Overall, the design approximates what Schaie (1977) has called a cross-sequential method with repeated measurement, or what Baltes et al. (1977) have called a longitudinal sequence. Two groups, the original high school seniors and the original middle-aged parents, will have passed through one transitional period and entered another of the researched transitions (marriage and anticipated birth of child for one, retirement for the other) over the ten years covered by the investigation. This design allows for consideration of historical change.

Whatever the label placed on the design (it might well be called a "transitions-sequential model," for example), the approach remains the same: an adult life-course study with repeated measurements. Respondents were initially interviewed in 1969-1970, with follow-ups conducted at 18 months, five years, and seven years. In 1980, another follow-up will be conducted which will allow for comparisons of short term and long term changes in psychosocial functioning. With this in mind, let us consider some of the methodological details of the study.

The Sample

Of the four transitions chosen for study, the two earlier ones were incremental, involving role gain, and were often voluntary. The high school seniors (N=52) and newlyweds (N=50) confronted role gains and new experiences and lifestyles. The two later transitions involved role loss and were often involuntary. The middle-aged parents (N=54) faced departure of the youngest child, and the oldest group (N=60) was anticipating either personal retirement or retirement of their spouse. All the subjects live in the central city of a metropolitan area on the west coast, are largely Caucasian, and were selected to be as representative of the middle

and lower-middle class as possible. A then very homogeneous senior high school was the central source for locating respondents for all four stages: the geographic boundaries of this school defined the district from which all four groups were drawn. The senior class of the high school constituted the universe for the youngest group. From school records, we determined which of these students were the youngest members of their families and thereby identified a second group, parents who were facing the "post-parental" or "empty-nest" stage of family life.

The group representing an intermediate stage between the above two consisted of newlyweds who had been married for less than twelve months and who had not yet started a family. They were selected primarily from marriage records from the county clerk's office. People in the fourth stage, made up of older persons planning to retire within the next two or three years, were drawn mostly from among friends, relatives, and co-workers of those already included in the sample, but a minority were located through local organizations which maintained records of the probable retirement dates of their employees. All persons so selected resided or had resided within the boundaries of the same middle class sampling district.

Table 1 presents some demographic characteristics of the total sample. The high school seniors were between 16 and 18 years old, and were the only cohort to include nonwhites (seven in all, including one Black, two Filipino Americans, two Chinese Americans, one Japanese American, and one Mexican American). The newlyweds ranged from 20-38 years of age, with an average of 24 (these were all first marriages). Combining men and women, the mean ages of the older groups were 50 for the middle-aged parents, and 60 for the preretirees (the N for the preretirees was enlarged to allow for morbidity and/or mortality in the longitudinal phases of the study).

Most of the newlyweds, middle aged parents and preretirees had some education beyond high school, either technical or general, but not usually to the completion of a bachelor's degree. They viewed education primarily as a necessary means of entering the work world, unlike the upper-middle class, where it is assumed that one will attain a degree independent of career considerations (Seeley, Sim, and Loosley, 1956; Keniston, 1968; Westley and Epstein, 1969). Most of our subjects (89 percent) were born in the United States and were raised and went to school in the neighborhood of their birth. Seventy-one percent of the high school seniors had been born in the city they currently lived in. For the older respondents, some geographic mobility was associated with the period following high school, but after that they tended to move very little.

For respondents of all stages, the average occupational prestige scores fall within the 40 to 49 range identified in national samples (Siegel, 1971) as signifying skilled or lower middle class occupations. In many ways the sample resembles those Coleman and Neugarten (1971) have called the "lower middle core," the majority of the men being

white collar clerical, sales, and technical workers, with the remainder divided between proprietors and contractors and skilled blue collar workers. They were home- and church-centered and lived in modest homes in a conservative, well-maintained part of town. Most took great pride in the neatness of their small homes and gardens.

Methods of Procedure

During the first contact (1969-1970) all respondents were interviewed for an average of six to eight hours, at a time and place of their convenience. The interview schedule consisted of a focused interview (Merton, Fiske and Kendall, 1956) designed to develop the respondent's own frame of reference and subjective perspectives on him or herself and respective "worlds." In addition, a number of structured instruments were included. Sections of the interview schedule at baseline were as follows:

Demographic and Sociostructural Data: Structured and semi-structured questions dealing with such standard variables as ethnic identity, religious affiliation, education, economic and occupational status, work history, geographic mobility, and basic sociodemographic data on families of origin and procreation.

Health History: (a) Structured questions designed to elicit information on subject's past and present physical and emotional status and the history of physical and emotional illness of immediate family members; (b) a Symptoms Checklist, a 42-item checklist of psychological symptoms developed collaboratively in consultation with Robert Butler, Alexander Simon, Leonard Micon, and the staff of the Human Development and Aging Program.

Behavioral Domain: (a) Open-ended questions designed to elicit the patterning of daily, weekly, and yearly activities and the problems, satisfactions, and meanings attached to these activities (based in part on the daily round questions from the Kansas City studies /Neugarten and Associates, 1964/, and the Cross-National Study of Aging /Bengtson, Chiriboga and Keller, 1969/); (b) Activities Checklist, a 33-item checklist designed to elicit the breadth and frequency of solitary and social activities, developed by staff during the pretest phase of the study.

Values and Goals Domain: (a) Open-ended questions dealing with the past, present, and future values and goals, determinants of goal choice, perceived supports and hindrances to goal attainment, and implementation of goals and outcome; (b) Goal Sort, a seven-item set of cards which the respondent is asked to rank in order of importance. This instrument, developed by staff during the pretest phase is especially useful in assessing change through time. It has been adopted by several other investigators (Gordon et al., 1973; Chiriboga, 1978b).

Family, Social Networks, and Social Perceptions: Open-ended questions focused on relationships with members of family of origin and procreation; dating experiences and events preceding marriage; attitudes toward sex and sexual experience; friendship patterns; organizational membership and participation; perceptions of neighborhood, social groups, and social horizons, including local and national problems and issues.

Evaluation of Life: (a) Open-ended questions dealing with the timing, nature, and evaluation of past life events and transitions, including experience with death and modes of coping with loss, anticipated future events and changes, and perceptions of the best and worst periods of life; (b) the Life Evaluation Chart, an instrument on which subjects rate each year of past, present, and anticipated future life on a nine-point scale of satisfaction/dissatisfaction. This chart is an elaboration of one developed by Jean MacFarlane of the Institute of Human Development, University of California, Berkeley. It has proved useful in assessing change through the follow-up studies.

Psychological Domain: Structured instruments and tests included (a) Bradburn Morale Scales (Bradburn and Caplovitz, 1965); (b) projective tests consisting of the Murray (1943) Thematic Apperception cards 1, 2, 10, 17BM, 19 and 20, and the Kansas City card (Neugarten and Gutmann, 1968); (c) WAIS Vocabulary and Block Design subtests (Wechsler, 1955); and (d) an Adjective Rating List of personality traits derived from Block's Modified Q-Sort for Nonprofessional Sorters (Block, 1961).

Interview Experience: (a) Subjects were asked their feelings about the interview experience; (b) interviews rated subjects on two scales designed to assess overall interview behavior and reaction to various sections of the interview; interviewers also rated subjects on the Adjective Rating List and wrote a narrative account of the interview and their impressions of and personal attitude toward subjects.

Instruments included at baseline and at each of the successive follow-up contacts are shown on Chart 1. Because of the scant knowledge of psychosocial change across the adult life course, the Longitudinal Study of Transitions was conceived in its first phase as descriptive and hypotheses generating. The baseline interview was designed to elicit comprehensive life history material, and emphasis was given to a combination of in-depth and semi-structured approaches. The intent in combining the qualitative and quantitative assessments was to provide both statistical data on change and an understanding of the intrapersonal dynamics of change. One of the major objectives of the baseline study was to provide an empirical base for the development of structured and pre-coded instruments which would provide for more economy in the follow-up studies. Consequently, in the 18-month, five and seven-year interviews we were able to cover many of the same study dimensions as in the baseline contacts, but in about half

of the interviewing time. The open-ended approach was used only to elicit information on the events, changes and stresses experienced by the respondent since the previous contact. In the five- and seven-year follow-up we paid special attention to the development, testing, and application of stress instruments, and to instruments for the evaluation of the perceived impact of social change.

The interview schedule for the ten year follow-up, now in preparation, will be essentially a shortened and more structured version of the baseline interview. In addition to the instruments described above, the following will be included:

Ways of Coping: A checklist of 68 items that describe a wide variety of thoughts and actions people use in dealing with stressful events. The theoretical framework underlying the checklist, which was developed by Lazarus (Lazarus, 1966; Lazarus et al., 1974; Lazarus and Launier, 1978), describes coping as a process that involves transactions between the person and the environment. It is currently being employed in an intensive panel study of 100 middle-aged men and women (Mediators Between Presumptive Stress and Adaptation, R. S. Lazarus, Principal Investigator).

Experience Inventory: A 48-item inventory of openness to experience (a construct which includes flexibility-rigidity). Originally developed by Coan (1972), the instrument has been modified for life span research by Costa and McCrae (1978), who have applied it to respondents from the Normative Study of Aging.

ENUF Scale: An instrument developed by Pierce (in press) in which 18 qualitative dimensions of leisure satisfaction (e.g., physical activity, achievement, novelty, excitement) are rated on a five-point scale ranging from "not enough" to "too much."

Beliefs and Attitudes Scale: A 21 item rating scale designed to assess conservatism/liberalism and traditionalism/modernity in selected areas (e.g., government, science, race relations, women's roles). This scale derives from the work of Smith and Inkeles (1966) and Suzman (1974).

Social Change Checklist: Respondents are asked to rate on a four-point scale (ranging from "none" to "strong") the effect that ten historical events and social changes (e.g., the Great Depression, Vietnam, new ways of doing things at work) have had on their lives.

Checklist of Hassles: A 12-item checklist (Chiriboga and Cutler, 1979) derived from a companion study (Mediators between Presumptive Stress and Adaptation; R. S. Lazarus, Principal Investigator). The checklist assesses how hassled or pressured respondents feel in 12 areas of life (work, marital, parental, etc.).

Sample Maintenance

A major problem in longitudinal studies is maintenance of the sample. In this longitudinal study we have adopted several strategies to maintain the cooperation of our respondents. They may be summarized as follows:

- (1) All respondents were asked during the initial interview whether or not they planned to remain in the same location and/or in the same area. They also provided the name, address and telephone number of a significant other who would know of a new address in the event of a move. Similar information is gathered at each subsequent contact.
- (2) Respondents are sent holiday greeting cards during the holiday season in December. Each card is signed personally by a staff member, usually one who has interviewed them.
- (3) Newsletters are sent out approximately every two years informing respondents of developments on the project and activities of staff.

Perhaps, as a result of these strategies, the attrition rate has been surprisingly low. Sample loss was 6.5 percent for the 18-month follow-up, 6.4 percent for the five-year follow-up, and 3.2 percent for the seven-year follow-up. The cumulative total percent attrition was 15.3 percent or a total of 33 persons. Of these, eight had died.

Respondents who dropped out have been compared with those who continued, on the basis of several baseline characteristics for each stage and sex group. Thus far, there is little selective attrition among those in the high school and retirement categories. It is slightly more pronounced among those initially newlyweds and empty-nest parents. Comparisons between "lost" respondents and those remaining in the study resemble findings from other longitudinal studies: dropouts had initially been somewhat more impaired. For example, they were more preoccupied with whatever stress they had experienced, reported more psychiatric symptoms, and were more pessimistic about the future. Group differences are small and do not notably change the distributions of the variables under consideration. Interestingly, one factor that often relates to dropouts in other studies, social class exhibited no differences between sample survivors and dropouts.

Theoretical Models

One of the greatest needs in the rapidly growing field of adult development and aging is for more efficient models than are provided by stage theories alone. Thus far, the study of adult transitions has provided solid grounding for the development of four analytic models which

show promise of elucidating both short and long term patterns of change. Descriptions of the models of commitment and transitions follow.

Commitment Model

One of the themes evident in both the self-reports of goals and values and the findings based on the dual model of adaptation was that those in the two younger groups (the original high school seniors and newlyweds) were in an expansive, growth-oriented phase, while those in the two later groups were more oriented to self-protection and ease and contentment. In the socioeconomic sector represented by the sample, the less challenged are better adapted as they grow older.

It was findings such as the above which led to the consideration of the concept of commitment, which subsumes the values and meanings invested in life (see Fiske, 1977; Fiske, 1978; Fiske, 1979; Fiske, in press). Commitments were defined as the patterning or configuration, sometimes hierarchical, of the fundamental concerns the person harbors at a particular time. This inner gestalt is drawn upon, consciously or otherwise, in organizing thoughts, which in turn allot priority to one kind of choice, decision, behavior or activity in preference to another. The commitment pattern thus influences one's disposition of time.

In-depth study of protocols from the Transitions Study has suggested the existence of four dimensions of commitment: (1) a cluster which includes curiosity, mastery and creativity; (2) commitment to other people; (3) commitment to a moral code which emanates from an awareness of the importance of nurturing integrity in one's self and others; and (4) commitment to self protectiveness. In preliminary analyses employing this model, changes in commitment to mastery among middle-aged men were examined; it was found that most of these men declined in mastery commitment during the first five years of the study. In contrast, the middle-aged women were more likely to have increased their self-assertive, mastery or creative commitments, which the cross-sectional analyses suggest were relatively low in the child-rearing phase (Fiske, 1979).

Among the middle-aged men, occupational peaking appears to be followed by many years of boredom on the job, if not sheer hatred of it. At the same time, in projective materials and in other measures thus far used to assess commitment, strong interpersonal needs emerge. This is true not only of the "middle majority" men in our Transitions Study, but has also been reported among the highly educated, self-appraising, and often very successful men in the Vaillant (1977) and Sears (1977) studies. In the Transitions Study, many of the men appear also to wish to be taken care of, indulged and supported (Fiske, 1979). Women in the study, as they seek more self-assertion, become less committed in the interpersonal domain. Few of the women were committed to their occupations, if they had one, and at middle age many were becoming aware of a need for more satisfying outlets.

The commitment framework is a promising one because it provides not only for development, expansion and growth, but also for stagnation and regression. It thus encompasses a far broader spectrum than do stage/developmental and self-realization paradigms. As we have found in our sample, middle-aged and older people who level off--or even regress--may be in better mental health than those who continue to struggle to grow.

Transitions Model

Although the concept of transitions in the adult life course is attracting considerable interest in the social sciences, and has formed the basis of popular and semi-popular books (Sheehy, 1976; Gould, 1978), much theoretical and empirical work remains to be done. One of the most informative treatments, in fact, comes not from the social sciences but from the philosophy of science. In his seminal work on the topic of scientific revolution, Kuhn (1962/1970) portrays change in the scientific community as a function of paradigm shift. A more recent work, by a humanistic philosopher, discusses stages and transitions within an overarching commitment framework (Norton, 1976).

Within the Transitions Study, a theoretical model of transitions is being developed drawing on the work of Kuhn (1962/1970), Van Gennep (1960), Cumming and Henry (1961) and Parkes (1972). A transition framework is currently defined (see Chiriboga, 1979; Chiriboga and Pierce, 1977) as follows: First, the individual's paradigm or assumptive world becomes increasingly disparate with new information, leading to a condition of precarious equilibrium (see Williams and Wirths, 1965). Associated with this precarious equilibrium may be an increasing tension or anxiety on the part of the individual, as she or he attempts to cope with the problem(s) in conventional ways that match with the increasingly mismatched assumptive world. Next, there may be a precipitating event. Levinson (1978) and others (Lundberg, 1939; Neugarten, 1977) have called these "marker" events or signposts. In themselves and by themselves, they may have little meaning. Taken in the context of other major changes, they may initiate the quest for a more satisfying paradigm. For example, the middle-aged worker whose career has peaked may begin to question her or his role in the world of work, and wonder if there is not something more rewarding in life. Whatever the new image or paradigm, the quest may take place in the context of heightening tension, as old coping strategies become less effective, but have not been supplanted by more effective ones. Finally, with the emergence of a more satisfying view of one's self in one's milieu comes re-equilibration. Efforts are now directed at consolidating a life style congruent with the new concept: "I am a family person, I am here to enjoy life and not work so hard and risk another heart attack. I am not so much a mother any more, but I am much more myself." The remainder of the new phase of life is then spent consolidating and refining the new paradigm, until it too shows signs of increasing discrepancy.

Research and model development in the area of transitions have proceeded simultaneously. One series of studies, cited above, compared the effects of the passage of time vs. the passage through a transition; the dependent variable was a set of self-concept dimensions (see Chiriboga and Pierce, 1977). Another analysis contrasted the effects of the normative transitions as experienced by the men and women in the Transitions Study, with the idiosyncratic transition of divorce; although both types of transitions influence adaptive processes, the results suggested that at least the short-range impact of divorce is greater (Chiriboga, in press).

Table 1: Selected Characteristics of the Sample at Baseline Contact (1969/70)

	High School		Newlywed		Middle-Aged		Preretirement	
	Men	Women	Men	Women	Men	Women	Men	Women
Sample N	25	27	25	25	27	27	30	30
Mean Age	17	17	25	23	52	48	61	58
Percent Some College	NA	NA	92	92	48	41	36	37
Percent Married	NA	NA	100	100	96	89	83	73
Religion (%)								
Protestant	36	26	28	24	41	44	60	40
Catholic	20	44	36	52	18	30	23	30
Jewish	12	4	4	12	18	22	7	10
Fundamentalist	12	4	8	12	11	4	3	17
Other/none	20	22	24	0	11	0	7	3
Percent Working Part Time	56	18	0	12	0	4	0	10
Percent Working Full Time	0	0	92	72	100	48	97	53
Socioeconomic Status (percent high) ^a	52	30	24	28	44	37	67	50
Mean Occupational Prestige ^b	38.9	44.2	45.7	44.8	45.6	43.1	47.7	44.4

^aBased on trichotomized scores.

^bOccupational Prestige scores computed for high school seniors from seven-year follow-up data (National Opinion Research Center, 1974).

Chart 1: Research Instruments

Instrument	Baseline	18-month	5-year	7-year	10-year (in process)
Life Events Questionnaire	(X)*		X	X	X
Bradburn Morale Scales	X	X	X	X	X
Life Evaluation Chart	X	X			
Adjective Checklist	X	X	X	X	X
Symptoms Checklist	X	X	X		X
Activities Checklist	X	X	X		X
Goal Sort	X	X	X		X
TAT	#1,2,10 17BM,19, 20, Kansas City Card	1,2,17Bm, 19	17BM		17BM, Kansas City Card
WAIS Vocabulary & Block Design	X				X
Rotter I-E Scale			X		
"Who am I" Test (Kuhn & McPartland)			X		
Friendship Card Sort			X		X
Sentence Completion Test (Loevinger)				X	X
Hassle Index				X	X
ENUF Scale				X	X

Chart 1: Research Instruments (Continued)

Instrument	Baseline	18-month	5-year	7-year	10-year (in process)
Social Attitudes				X	
Social Change Checklist				X	X
Ways of Coping					X
Experience Inventory					X

* preliminary version

References

- Baltes, P. B., Reese, H. W., & Nesselroade, J. R. Life-span developmental psychology: introduction to research methods. Monterey, Calif.: Brooks/Cole Publishing Co., 1977.
- Benson, D., & Lowenthal, M. F. Perceived stress across life course. In Lowenthal, M. F., Thurnher, M., Chiriboga, D., et al. Four stages of life: a comparative study of women and men facing transitions. San Francisco: Jossey-Bass, 1975, 163-175.
- Bengtson, V. L., Chiriboga, D. A. & Keller, A. C. Occupational differences in retirement: patterns of role activity and life-outlook among Chicago retired teachers and steelworkers. In Havighurst, R. J., Thomae, M., Neugarten, B. L. & Munnichs, J. K. L. (Eds.) Adjustment to retirement: a cross national study. Assen, Netherlands: Van Gorkum, 1969.
- Block, J. The q-sort method in personality assessment and psychiatric research. Springfield, Ill.: Thomas, 1961.
- Bradburn, N. M. The structures of psychological well-being. Chicago: Aldine, 1969.
- Bradburn, N. M. & Caplovitz, D. Reports on happiness: a pilot study of behavior related to mental health. Chicago: Aldine, 1965.
- Chiriboga, D. A. The adult self-concept: a longitudinal analysis. Presented at the 53rd Annual Meeting of the Western Psychological Association, Anaheim, April 11-14, 1973.
- Chiriboga, D. A. Perceptions of well-being. In Lowenthal, M. F., Thurnher, M., Chiriboga, D. A., et al. Four stages of life: a comparative study of women and men facing transitions. San Francisco: Jossey-Bass, 1975, 84-98.
- Chiriboga, D. A. Life event weighting systems: a comparative analysis. Journal of Psychosomatic Research, 1977, 21, 415-422.
- Chiriboga, D. A. Evaluated time: a life course perspective. Journal of Gerontology, 1978a, 33(3), 388-393.
- Chiriboga, D. A. The divorce experience in later life in California. Annual Meeting of the American Orthopsychiatric Association, San Francisco, March 1978b.
- Chiriboga, D. A. Conceptualizing adult transitions: a new look at an old subject. Generations, Issue on Life Transitions (David Chiriboga, Guest Editor), Summer 1979 IV(1), 4-6.

- Chiriboga, D. A. Marital separation and stress: a life course perspective. Alternative Lifestyles, Beverly Hills, Calif.: Sage, in press.
- Chiriboga, D. A., Coho, A., Stein, J. & Roberts, J. Divorce stress and social supports: a study in helpseeking. Journal of Divorce, 3(2), Winter 1979.
- Chiriboga, D. A. & Cutler, L. Stress responses among divorcing men and women. Journal of Divorce, Winter 1977, 1(2), 95-106.
- Chiriboga, D. A. & Cutler, L. Stress and adaptation: a life span study. Presented at the 97th Annual Convention of the American Psychological Association, New York City, September 1-5, 1979.
- Chiriboga, D. A. & Dean, H. Dimensions of stress: perspectives from a longitudinal study. Journal of Psychosomatic Research, 1978, 22, 47-55.
- Chiriboga, D. A. & Gigy, L. Perspectives on life course. In Lowenthal, M., Thurnher, M., Chiriboga, D., et al. Four stages of life: a comparative study of women and men facing transitions. San Francisco: Jossey-Bass, 1975, 122-145.
- Chiriboga, D. A. & Lowenthal, M. F. The correlates of social stress. Presented at the 52nd Annual Meeting of the Western Psychological Association, Portland, Oregon, April 26-29, 1972.
- Chiriboga, D. A. & Lowenthal, M. F. Complexities of adaptation. In Lowenthal, M. F., Thurnher, M., Chiriboga, D., et al. Four stages of life: a comparative study of women and men facing transitions. San Francisco: Jossey-Bass, 1975, 99-121.
- Chiriboga, D. A. & Pierce, R. C. The self in transition: a study in change. Presented at the 30th Annual Scientific Meeting of the Gerontological Society, San Francisco, Calif.: November 17-22, 1977.
- Chiriboga, D. A., Roberts, J. & Stein J. Psychological well-being during marital separation. Journal of Divorce, Fall 1978, 2(1) 21-36.
- Chiriboga, D. A. & Thurnher, M. Concept of self. In Lowenthal, M. F., Thurnher, M., Chiriboga, D., et al. Four stages of life: a comparative study of women and men facing transitions. San Francisco: Jossey-Bass, 1975, 62-83.
- Chiriboga, D. A. & Thurnher, M. Marital lifestyle and adjustment to separation. Journal of Divorce, in press.
- Clark, M. & Anderson, B. G. Culture and aging: an anthropological study of older Americans. Springfield, Ill.: Charles C. Thomas, 1967.

- Coan, R. W. Measurable components of openness to experience. Journal of Consulting and Clinical Psychology, 1972, 37, 346-352.
- Coleman, R. P. & Neugarten, B. L. Social status in the city. San Francisco: Jossey-Bass, 1971.
- Costa, P. T., Jr., & McCrae, R. R. Objective personality assessment. In Storandt, M., Siegler, I. C., & Elias, M. F. (Eds.) The clinical psychology of aging. New York: Plenum Press, 1978, 119-143.
- Cumming, E., and Henry, W. E. Growing old: the process of disengagement. New York: Basic Books, 1961.
- Fiske, M. Interpersonal relationships and adaptation in adulthood. In Ostwald, P. R. (Ed.), Communication and social interaction: clinical and therapeutic aspects of human behavior. New York: Grune and Stratton, 1977, 263-275.
- Fiske, M. Part II: Adult transitions: theory and research from a longitudinal perspective. Presented at the 31st Annual Scientific Meeting of the Gerontological Society, Dallas, Texas, November 1978.
- Fiske, M. Middle age: the prime of life? (Life Cycle Series, Lifetime Books, Ltd., London) London: Harper and Row, 1979.
- Fiske, M. Tasks and crises of the second half of life: resources and stumbling blocks. In Birren, J. E. & Sloane, R. B. (Eds.), Handbook of mental health and aging. Englewood Cliffs, N. J.: Prentice Hall, in press.
- Frenkel-Brunswik, E. Adjustment and reorientation in the course of the life span. In Neugarten, B. L. (Ed.) Middle age and aging. Chicago: University of Chicago Press, 1968, 77-84.
- Gordon, C., Gaitz, C. M. & Scott, J. Value priorities and leisure activities among middle aged and older Anglos. Diseases of the nervous system, 1973, 34(1), 13-26.
- Gould, R. Transformations: growth and change in adult life. New York: Simon and Schuster, 1978.
- Gutmann, D. L. An exploration of ego configurations in middle and later life. In Neugarten, B. L. et al. Personality in middle and late life. New York: Atherton, 1964, 114-148.
- Holmes, T. H. & Rahe, R. H. The social readjustment rating scale. Journal of Psychosomatic Research, August, 1967, 11, 213-218.

- Horowitz, M., Schaefer, C., Hiroto, D., Wilner, N., & Levin, B. Life events questionnaires for measuring presumptive stress. Psychosomatic Medicine, 1977, 39(6), 413-431.
- Keniston, K. Young radicals: notes on committed youth. New York: Harcourt Brace, 1968.
- Kuhlen, R. G. Developmental changes in motivation during the adult years. In B. L. Neugarten (Ed.) Middle age and aging. Chicago: University of Chicago Press, 1968, 115-136.
- Kuhn, T. S. The structure of scientific revolutions. Chicago: University of Chicago Press, (1962) 1970.
- Lazarus, R. S. Psychological stress and the coping process. New York: McGraw Hill, 1966.
- Lazarus, R. S., Averill, J. R., Opton, E. M., Jr. The psychology of coping: issue of research and assessment. In Coelho, G. V., Hamburg, D. A., & Adams, J. R. (Eds.), Coping and adaptation. New York: Basic Books, 1974, 249-315.
- Lazarus, R. S. & Launier, R. Stress related transactions between person and environment. In Pervin, L. A. & Lewis, M. (Eds.), Perspectives in interactional psychology. New York: Plenum, 1978, 287-327.
- Levinson, D. J. The seasons of a man's life. New York: Alfred A. Knopf, 1978.
- Lowenthal, M. F. Lives in distress: the paths of the elderly to the psychiatric ward. New York: Basic Books, 1964.
- Lowenthal, M. F. Some potentialities of a life cycle approach to the study of retirement. In Carp, F. (Ed.), Retirement. New York: Behavioral Publications, 1972, 307-336.
- Lowenthal, M. F., Berkman, P. L., et al. Aging and mental disorder in San Francisco: a social psychiatric study. Jossey-Bass, Inc., 1967.
- Lowenthal, M. F. & Chiriboga, D. A. Social stress and adaptation: toward a life course perspective. In Eisdorfer, C. & Lawton, M. P. (Eds.), The psychology of adult development and aging. Washington, D. C.: American Psychological Association, 1973, 281-310.
- Lowenthal, M. F. & Chiriboga, D. A. Responses to stress. In Lowenthal, M. F., Thurnher, M., Chiriboga, D. A., et al. Four stages of life: a comparative study of women and men facing transitions. San Francisco: Jossey-Bass, 1975, 146-162.

- Lowenthal, M. F. & Haven, C. Interaction and adaptation: intimacy as a critical variable. American Sociological Review, February 1968, 33, 20-30.
- Lowenthal, M. F. & Robinson, B. Social networks and isolation. In Binstock, R. H. & Shanas, E. (Eds.), Handbook of aging and the social sciences. New York: Van Nostrand Reinhold, 1976, 432-456.
- Lowenthal, M. F. & Weiss, L. Intimacy and crises in adulthood. In Whiteley, J. N. (Ed.), Counseling adults. Special issue of The Counseling Psychologist, 6(1), 1976, 10-15.
- Lowenthal, M. F. & Zilli, A. (Eds.), Interdisciplinary topics in gerontology: colloquium on health and aging of the population, Vol. 3. Basel, Switzerland, New York: S. Karger, 1969.
- Lundberg, G. A. Foundations of sociology. New York: MacMillan, 1939.
- Lurie, E. Sex and stage differences in perceptions of marital and family relationships. Journal of Marriage and the Family, May 1974, 36(2), 260-269.
- McKinney, J. P., Hotch, D. F. & Truhon, S. A. The organization of behavioral values during late adolescence: change and stability across two eras. Developmental Psychology, 1977, 13(1), 83-84.
- Merton, R. K., Fiske, M. & Kendall, P. L. The focused interview: a manual of problems and procedures. Glencoe, Ill.: The Free Press, 1956.
- Murray, H. A. Manual of thematic apperception test. Cambridge, Mass.: Harvard University Press, 1943.
- National Opinion Research Center. National data program for the social sciences. Chicago: University of Chicago Press, 1974.
- Neugarten, B. L. Personality and aging. In Birren, J. E. & Schaie, K. W. (Eds.), Handbook of the psychology of aging. New York: Van Nostrand Reinhold Company, 1977, 626-649.
- Neugarten, B. L., et al. Personality in middle and late life: empirical studies. New York: Atherton Press, 1964.
- Neugarten, B. L. & Gutmann, D. L. Age-sex roles and personality in middle age: a thematic apperception study. In Neugarten, B. L. (Ed.), Middle age and aging. Chicago: University of Chicago Press, 1968, 58-71.
- Norton, D. L. Personal destinies: a philosophy of ethical individualism. Princeton, N. J.: Princeton University Press, 1976.

- Parkes, C. M. Bereavement: studies of grief in adult life. New York: International Universities Press, Inc., 1972.
- Pierce, R. C. Dimensions of Leisure. 1: Satisfactions. Journal of Leisure Research, in press.
- Pierce, R. C. & Chiriboga, D. A. Dimensions of adult self-concept. Journal of Gerontology, 1979, 34(1), 80-85.
- Robinson, B. & Thurnher, M. Taking care of aged parents: a family-cycle transition. The Gerontologist, in press.
- Rokeach, M. Change and stability in American value systems, 1968-1971. Public Opinion Quarterly, 1974, 38(2), 222-238.
- Schaie, K. W. Quasi-experimental research designs in the psychology of aging. In Birren, J. E. & Schaie, K. W. (Eds.), Handbook of the psychology of aging. New York: Van Nostrand Reinhold, 1977, 39-58.
- Schaie, K. W. & Parham, I. A. Social responsibility in adulthood: ontogenetic and sociocultural change. Journal of Personality and Social Psychology, 1974, 30(4), 483-492.
- Sears, R. Sources of life satisfactions of the Terman gifted men. American Psychologist. February 1977, 32, 119-128.
- Seeley, J. R., Sim, R. A. & Loosley, E. W. Crestwood Heights: a study of the culture of suburban life. New York: Basic Books, 1956.
- Sheehy, G. Passages: predictable crises of adult life. New York: Dutton, 1976.
- Siegel, P. S. Prestige in American occupational structure. Unpublished Ph.D. Dissertation, Dept. of Sociology, University of Chicago, 1971.
- Simon, A., Lowenthal, M. F. & Epstein, L. J. Crisis and intervention: the fate of the elderly mental patient. San Francisco: Jossey-Bass, Inc., 1970.
- Smith, D. H. & Inkeles, A. The OM scale: a comparative socio-psychological measure of individual modernity. Sociometry, 1966, 29, 353-377.
- Spence, D. L. & Lonner, T. D. Career set: a resource through transitions and crises. International Journal of Aging and Human Development, 1978-1979, 9(1), 51-65.
- Spence, D. & Lurie, E. Style of life. In Lowenthal, M. F., Thurnher, M., Chiriboga, D., et al., Four stages of life: a comparative study of women and men facing transitions. San Francisco: Jossey-Bass, 1975, 1-25.

Suzman, R. M. Psychological modernity. International Journal of Comparative Sociology, 1974, 13(3-4), 273-287.

Suzman, R. Effects of employment and occupation on ego-level. Paper presented at the 30th Annual Meeting of the Gerontological Society, San Francisco, November 1977.

Thurnher, M. Goals, values and life evaluations at the preretirement stage. Journal of Gerontology, 1974, 29(1), 85-96.

Thurnher, M. Continuities and Discontinuities in Value Orientations. In Lowenthal, M. F., Thurnher, M., Chiriboga, D., et al. Four stages of life: a comparative study of women and men facing transitions. San Francisco: Jossey-Bass, 1975, 176-200.

Thurnher, M. Midlife marriages: sex differences in evaluations and perspectives. International Journal of Aging and Human Development, 1976, 17(2), 129-135.

Thurnher, M., Spence, D., & Lowenthal, M. F. Value confluence and behavioral conflict in intergenerational relations. Journal of Marriage and the Family, May 1974, 36(2), 308-319.

Vaillant, G. E. Adaptation to life. Boston: Little, Brown and Co., 1977.

Van Gennep, A. The rites of passage. London: Routledge and Kegan Paul, 1960.

Wechsler, D. Manual for the Wechsler adult intelligence scale. New York: Psychological Corporation, 1955.

Weiss, L. & Lowenthal, M. F. Life-course perspectives on friendship. In Lowenthal, M. F., Thurnher, M., Chiriboga, D., et al, Four stages of life: a comparative study of women and men facing transitions. San Francisco: Jossey-Bass, 1975, 48-61.

Westley, W. A., & Epstein, N. B. The silent majority: families of emotionally healthy college students. San Francisco: Jossey-Bass, 1969.

Williams, R. H. & Wirths, C. G. Lives through the years. New York: Atherton, 1965.

The Consumer Panels of Young Married Couples

Eileen Crimmins
Department of Sociology
Rutgers University - Camden

Linda Lannom
Survey Research Laboratory
University of Illinois

Phyllis A. Ewer
Department of Sociology
University of Illinois - Chicago Circle

The Consumer Panels are two samples of young married couples who have been interviewed regularly since the time of their marriages. One of the panels consists of couples married during the summer of 1968 in Peoria and Decatur, Illinois. The other panel was selected from young couples married in Chicago during the summer of 1972. Family economic behavior has been the primary focus of these panel studies, which are being conducted by the staff of the Survey Research Laboratory at the University of Illinois under the direction of Robert Ferber with the collaboration of Francesco Nicosia.

There are a number of features which make the Consumer Panels unique among studies of family economic behavior. The panel aspect is, of course, the primary of these. Another is that the family is not viewed as a single homogeneous unit. It is regarded as being composed of individuals with separate and possibly different backgrounds and personality characteristics which all influence family behavior to some extent. That couples are in the panels at the stage of family formation and growth is also unique to the consumer panels. This fact, along with the panel design, allowed the collection of data from families while their economic and consumption habits were being formed as well as when many major life cycle events such as birth of children, migration or divorce were taking place.

Criteria for inclusion

Because of the somewhat exploratory nature of the study, Ferber and Nicosia initially wanted to avoid the complexity of a large city and to interview subjects in "one or two small cities, as reasonably closed or self contained (economically and socially) as possible" (Ferber and Nicosia, 1972, p. 165). For this reason they selected Peoria and Decatur, Illinois, as sites of the initial panel. They describe Peoria and Decatur as "both between 100,000 and 250,000 (in population), industrial, and relatively isolated from other major urban centers" (Ferber and Nicosia, 1972, p. 165).

The Peoria/Decatur sample was selected from a list of couples who married in Peoria or Decatur between June 1, 1968, and September 30, 1968. This list of about 1,300 names was compiled from county marriage license data from the three counties of the Peoria and Decatur Standard Metropolitan Statistical Areas (Peoria, Tazewell, and Macon Counties), newspaper accounts of new marriages, and the records of clergy. A sample of 400 newly married couples was selected from this list, 150 in Decatur and 250 in Peoria. To obtain the newly married couple's address after their marriage their last names were matched with telephone directory listings or secondary sources (usually parents) were contacted. Only couples who lived in either the Peoria or Decatur Standard Metropolitan Statistical Areas after their marriages were included in the sample. In addition, the husband had to be 30 years of age or less at the time of the marriage. The panel consisted of all eligible couples who agreed to participate in the study. During the first wave of interviewing in September and October of 1968, 311 eligible couples completed the questionnaire.

When the Peoria/Decatur panel was four years old, the Chicago panel was formed. The sample for the Chicago Consumer Panel was selected from a list of couples who were married in Cook County between June 1, 1972, and August 31, 1972. This list was compiled from microfilm marriage records obtained from the Illinois Office of Vital Records. A sample of 1,003 couples was selected from the list. As in Peoria and Decatur, telephone directory listings or secondary sources were used to obtain the newly married couple's address after their marriage. Only couples who lived in the Chicago Standard Metropolitan Statistical Area after their marriage were eligible.

Four additional eligibility criteria were imposed in Chicago:

1. The husband should be 30 years of age or less at the time of the marriage;
2. The marriage should be the first one for both members of the couple;
3. The couple could not own a home or condominium at the time of the marriage;

4. The couple's income should exceed \$5,000 per year, unless one of them was a student.

Since it was difficult to find couples who met all of these criteria, the first criterion was relaxed to include couples with husbands who were 34 years of age or less at the time of their marriage. The Chicago panel was composed of 409 couples when it began in November, 1972.

Design

By June 1980 the Peoria/Decatur panel had been interviewed 17 times and the Chicago panel 11 times. The date of each of these interviews is shown in Table 1. During the first nine years of the Peoria/Decatur panel and the first five years of the Chicago panel the aim was to interview both panels every six months. After this time, the aim was to interview both panels once a year. Because of the vagaries of funding it was not always possible to accomplish these aims; thus, the actual time between interviews varied from 6 to 15 months.

The number of households completing interviews at each of the waves is also shown in Table 1. From the first through the sixteenth wave, the total number of households in Peoria and Decatur dropped from 311 to 210. The reduction in the number of cases was even more severe in the Chicago panel which dropped from 409 to 203 cases from the first to the tenth interview. Attrition of the original households occurred because people no longer wanted to participate in the panel or they had moved and could not be contacted. The composition of the panel was changed by addition as well as attrition. If the original marriage ended in divorce or death, the original household was said to have "died." A new household was born when one of the original members of a "dead" household continued in the panel. Thus, not every household interviewed after the first wave was one of the original panel households. Households "born" after wave one consisted of individuals from the original households, individuals from the original households plus new spouses, and in at least one case an original couple that divorced and then remarried. When a household was "born," the available data from earlier interviews with the "parent" household are copied onto the data for the new case as "secondhand" information in order to connect background information collected on previous waves to the new case. Among the 210 households completing the sixteenth interview in Peoria and Decatur, 146 were original couples and 64 were cases "born" after the beginning of the panel. Of the 203 households completing the tenth wave in Chicago, 165 were original couples and 38 were cases "born" after the beginning of the panel.

Until 1977 respondents continuing to reside in the metropolitan areas were interviewed face to face by professional interviewers with the exception of one mail survey in 1972. Respondents who moved away from the metropolitan areas were, however, interviewed regularly by mail. For all

respondents, the last three waves of interviewing were conducted over the telephone.

In keeping with the idea that the family is a unit composed of separate individuals, each interview has consisted of at least three parts--a general household questionnaire, a husband's questionnaire and a wife's questionnaire. These separate spouses' questionnaires were self-administered until 1977. After that time separate interviews with the husband and wife were conducted by telephone. At four of the waves the panel respondents received a fourth questionnaire, called the "Family Financial Survey," requesting detailed financial information. With the exception of the most recent wave, respondents were asked to fill out this questionnaire after the interview and return it by mail. On the most recent wave, respondents were asked to complete this questionnaire in advance of the telephone interview and to read the appropriate code numbers over the phone.

For many of the waves, the husband's and wife's questionnaires were identical so that husbands and wives often responded individually to the same questions. In addition, the questionnaires for both panels are usually identical at each date. Some attempt has been made to repeat questions so that both panels answer the same questions after similar durations of marriage.

Characteristics of Samples

Peoria/Decatur. Not all couples in the Peoria/Decatur panel were marrying for the first time. Of the original 311 cases, only 266 were first marriages. In keeping with the character of the two communities, this is a racially homogeneous sample. Only six of the 252 couples for whom race is known are black. Median age at the first interview was 22 for male panel members and 20 for female panel members. In response to a question on educational attainment, both husbands (52 percent) and wives (65 percent) in the Peoria/Decatur sample were most likely to answer that they had completed high school at the time of marriage. Median years of school completed at this point were 12.2 by men and 12.1 by women. At the first interview in Peoria/Decatur, 96 percent of the husbands and 70 percent of the wives were employed. While the men are represented in all occupational categories, the women are heavily concentrated in clerical and sales occupations. In the Spring of 1970 (at the fourth interview) Peoria/Decatur respondents reported that their median family income for 1969 (which was their first full year of marriage) was \$10,500.

Chicago. Reflecting the complexity of a large city, the Chicago panel is more heterogeneous than the Peoria/Decatur panel. Eighteen percent of the couples for whom data on race are available are black. Educational attainment is also more varied and higher within the Chicago panel than in the Peoria/Decatur panel. Males were most likely to have some college at the time of marriage but only 30 percent were in this modal response category. As in Peoria and Decatur, the females were most likely to respond that they

had only a high school degree but less than half as many (30 percent) Chicago panel wives as Peoria/Decatur wives were in this category. Median years of schooling completed at the first interview was 13.9 for men and 13.5 for women in Chicago. Both husbands and wives in this panel were a year older at marriage than the Peoria/Decatur panel. Median age at first interview was 23 for the men and 21 for the women. The older age of the Chicago panel may account in part for the higher level of education attained by this group. At the same date 91 percent of the husbands and 70 percent of the wives were employed. Twenty-eight percent of both the employed husbands and wives in the Chicago panel have professional or managerial occupations. This percentage is higher than the 20 percent for females in the Peoria/Decatur sample, indicating that the Chicago couples were more concentrated in higher status jobs at the time of marriage, which is to be expected given the educational differences between the panels. At the first interview, the members of the Chicago panel reported their expected combined income for the year of 1972. The median value was \$11,600.

Topics Included in the Surveys

Economic Behavior. The major purpose of these panels is to collect data on the consumption habits of these young married couples. For this reason each time couples are interviewed they respond to questions on their purchases of durable goods during the time between interviews and their intentions to purchase durable goods during some specified time period in the future. At regular intervals, information on their stock of goods is updated. For many of these goods, data are available on such characteristics as the price paid, the brand name, on whose advice the purchase was made, and who made the decision to purchase. Extensive information is also collected on the couple's financial portfolios, saving and budgetary practices, use of credit, and ownership of life insurance. Many of the questions in these areas are geared toward determining the role of each of the marital partners in the decision-making process, forces outside the household influencing decision-making, and the actual division of labor within the marriage surrounding the task of money management.

Data relevant to the economic personalities of the two spouses are also collected periodically. Both husbands and wives respond to questions by which they can be characterized as economically minded, price conscious or bargain seeking. Subjective assessments of the country's current and future economic situation including inflation along with ideas about government spending have also been gathered from the respondents. Intermittently the subjects have assessed their own economic situations relative to that of their parents, their friends, their own pasts and their own futures. Data on actual household income have normally been collected once a year.

Labor Force Participation. Data on additional economic characteristics relevant to labor force participation such as occupation, industry, part or full time employment, and involvement in a second job were collected at each interview. The surveys have also included an extensive number of questions relevant to the study of female labor force participation (most of these after wave nine in Peoria/Decatur and wave three in Chicago). These included questions of nonworking wives on their reasons for not working as well as their own assessments of their income and occupational prospects if they were to seek a job. These women were also asked to estimate the costs of child-care arrangements they would incur by working. Working women reported on the costs that they actually incurred while working as well as their reasons for working. All women responded to questions on their work plans after having children and in the more general future.

Background Characteristics. A series of questions on the respondents' background characteristics was included in the first wave for each of the panels. These included birthplace, occupation of father, number of siblings, parents' marital status, and whether parents were currently living. Current religion and religious background were assessed on the tenth wave in Peoria/Decatur and the fourth wave in Chicago.

Aspirations and Goals. Periodic attempts were made to measure respondents' aspirations and goals for the future. These included their educational, occupational and familial aspirations as well as their financial aspirations. One wave also included questions which measured the respondents' aspirations for their children.

Children. Other aspects of childbearing were a major topic in these surveys. At each wave, respondents indicated whether they had had a birth during the interval. On the tenth wave in Peoria/Decatur and the fourth wave in Chicago, the exact dates of all births were obtained. This provided a check on earlier answers and filled in missing data for a few cases. Respondents also reported at each wave whether they were pregnant or in the process of adoption. After wave six in Peoria/Decatur and beginning in the first wave in Chicago, respondents answered questions periodically on their desired and expected family sizes. Respondents were asked at the seventh interview in Peoria/Decatur whether they were presently using birth control and on the fourteenth wave in Peoria/Decatur (eighth in Chicago) whether each of their children was conceived while the couple was practicing birth control. On the same wave (fourteenth in Peoria/Decatur, eighth in Chicago) respondents also reported their opinions about abortion. On various waves an attempt was made to reveal the decision-making process surrounding fertility. Respondents indicated what factors influenced the decision to have children, who made the decision, and the effect of an additional child on various facets of their lives.

General Attitudes and Personality Characteristics. Throughout the surveys attempts were made to measure general attitudes toward life and personality characteristics of the respondents. Questions regarding attitudes toward life in general and satisfaction with one's own life and marriage were included repeatedly in the surveys. Among other characteristics, measures of how experiment prone, timid, conservative, and self-assured a person was were constructed from responses to individual and groups of questions.

Sex-Role Orientation. Both husbands and wives have responded periodically to questions concerning at what age and under what circumstances a mother/wife should work, the importance of job advancement, factors important in job selection, decision-making concerning employment of family members, and work histories of respondents' mothers. Husbands and wives also responded to questions on the division of labor by sex within their own homes as well as those of their families of origin. These questions can be used to measure couples' attitudes toward male/female roles, background characteristics that might have influenced these attitudes, and the actual division of labor and decision-making within the home according to sex.

Time Use. Use of time for household tasks and leisure activities was the focus of a number of questions on waves eight, nine, twelve and fifteen in the Peoria/Decatur study and two, three, six and nine in the Chicago study. A series of questions on vacation plans and habits was asked of the Peoria/Decatur respondents on the eighth wave.

Social and Political Attitudes. Some topics covered in the surveys reflect the political and social climate of the times in which they were asked. For instance, respondents' attitudes toward electric cars, reusable containers, and organically grown food were assessed. Questions on transportation were asked during the energy crisis of 1973-1974. Attitudes toward protest by blacks, students, and truckdrivers were measured at the time truckers were blocking the highways. At this same time questions concerning violence on television and police violence were also asked. During earlier waves questions on media preferences (newspaper, magazine, television) and exposure were asked.

This list represents a fraction of measures taken during the study and should not be regarded as exhaustive. Many of the topics mentioned above were covered on the spouses' separate questionnaires so that answers are available from both members of a couple. In addition, many of the topics listed are included repeatedly in the survey so answers are available to the same questions at different points in time and after different durations of marriage. Table 2 summarizes selected data which are currently available for the consumer panels.

Data Analysis

Because the consumer panels are ongoing projects, analysis to date has been limited. Most of it has been done using data from the Peoria/Decatur sample since the Chicago marriages are still of relatively short duration. The analysis which has been done has concentrated on two major areas. First, Ferber and his associates have studied various facets of family financial behavior. Second, Ewer and Crimmins have used the panel data for work on fertility and female labor force participation.

Conclusions

Because of their uniqueness, the consumer panels provide exceptionally rich data sets. For instance, analyses using these data sets have shown that marriage cohorts are meaningful groups for the study of many family behaviors and processes. In much research on family behavior based on other data sets, it is necessary to control for the duration of marriage or divide families into life cycle stages. Analysis of data from the consumer panels can be done without using these somewhat cumbersome techniques. In addition, the fact that data are available for both members of the couple gives an added dimension to these data sets in a time when the roles men and women play are being questioned and changed. Analyses of these data have shown the importance of considering the characteristics of both the husband and wife in studying family behavior. Although comparisons between the two panels have been very limited to date, it is expected that the different locations of the two panels and the different economic times in which the two groups were married will also prove to be valuable dimensions of the panels.

While the consumer panels provide a wealth of data, their use is not without problems. The major problem arises from sample attrition. In order to reap the full benefits of the panel dimension of these data, material from the largest possible time span should be used. Doing this, however, results in a sample size that is significantly smaller than the original sample size. The panel dimension itself can provide problems for certain analyses if the households "born" after the beginning of the panel are included and assumed to be similar to the original households. Another problem arises from the selectivity of the sample. Because the panels were not chosen as representative samples of other groups, the results cannot be assumed to be true for other groups. This is a problem to users of the consumer panels because there is a tendency in the behavioral sciences to dismiss the results from such samples as unreliable rather than seeing them as based on selected groups that, in some sense, already control for some of the usual confounding influences. The last problem is one that is common to all survey research and arises from the fact that hindsight is clearer than foresight. There are other problems in using the consumer panel data related to the way questions were worded, the timing of certain questions, and the omission of other questions, but these problems are no greater than those to be expected from such a large undertaking.

References

- Anderson, Dan R., & Nevin, John R. Determinants of Young Marrieds' Life Insurance Purchasing Behavior: An Empirical Investigation. Journal of Risk and Insurance, 1975, 42, 375-387.
- Crimmins, Eileen, & Ewer, Phyllis A. The Prediction of Young Mothers' Labor Force Participation and Future Participation Expectations through Use of a Household Decision-Making Model. Paper presented at the meetings of the Population Association of America, 1978.
- Crimmins-Gardner, Eileen, & Ewer, Phyllis A. Relative Status and Fertility. Research in Population Economics, 1978, 1, 247-259.
- Ewer, Phyllis A., & Crimmins-Gardner, Eileen. Tempo of First Births and Goods Acquisition in a Marriage Cohort. Paper presented at the meetings of the Population Association of America, 1977.
- Ewer, Phyllis A., & Crimmins-Gardner, Eileen. Income in the Income and Fertility Relationship. Journal of Marriage and the Family, 1978, 291-299.
- Ewer, Phyllis A., & Crimmins, Eileen. Age at Marriage, Age at First Birth and Sex Role Attitudes as Predictors of Family Size and Working. Paper presented at the meetings of the American Sociological Association, 1978.
- Ewer, Phyllis A., Crimmins, Eileen, & Oliver, Richard. An Analysis of the Relationship between Husband's Income, Family Size and Wife's Employment in the Early Stages of Marriage. Journal of Marriage and the Family, 1979, 41, 40-51.
- Ferber, Robert, & Lee, Lucy Chao. Increasing the Effectiveness of Buying Plans Variables in Economic Models. Working paper No. 119, College of Commerce and Business Administration, University of Illinois, 1973.
- Ferber, Robert, & Lee, Lucy Chao. Husband-Wife Influence in Family Purchasing Behavior. Journal of Consumer Research, 1974a, 1, 43-50.
- Ferber, Robert, & Lee, Lucy Chao. The Role of Life Style in Studying Family Behavior. Working paper No. 226, College of Commerce and Business Administration, University of Illinois, 1974b.
- Ferber, Robert, & Lee, Lucy Chao. Accumulation of Durable Goods by Young Marrieds. In Modelling for Government and Business: Essays in Honor of Prof. Dr. P.J. Verdoorn, edited by C.A. Van Bochove et al., Leiden: Martinus Nijhoff Social Sciences Division, 1977, 227-246.

Ferber, Robert, & Nicosia, Francesco. Newly Married Couples and their Asset Accumulation Decisions. In Human Behavior in Economic Affairs: Essays in Honor of George Katona, edited by B. Strumpel et al., San Francisco: Jossey Bass, Inc., 1972, 161-187.

Lee, Lucy Chao, & Ferber, Robert. Asset Accumulation in Early Married Life. Working paper No. 403, College of Commerce and Business Administration, University of Illinois, 1977.

Liebman, Ellen, & Ferber, Robert. The Use of Configurations of Purchase Likelihoods to Predict Auto Purchases. Working paper No. 385, College of Commerce and Business Administration, University of Illinois, 1977.

Nicosia, Francesco, & Ferber, Robert. The Use of Configuration of Buying Intentions for Predicting Auto Purchases: Some Explorations. In Attitude Research at Bay, edited by D. Johnson and W. Wells. Chicago: American Marketing Association, 1976, 1-30.

Table 1. Interview Date, Collection Methodology, and Number of Households Completing Interviews at Each Wave:
Consumer Panels of Young Married Couples

Date	Method of Administering the Main Questionnaire ¹	Peoria-Decatur				Chicago			
		Wave	Total	Original Households	Households formed after Wave 1	Wave	Total	Original Households	Households formed after Wave 1
Fall, 1968	Face - to - Face	1	311	311	0				
Spring, 1969	Face - to - Face	2	286	286	0				
Fall, 1969	Face - to - Face	3	259	258	1				
Spring, 1970	Face - to - Face	4	253	248	5				
Fall, 1970	Face - to - Face	5	237	231	6				
Spring, 1972	Mail with telephone	6	224	220	4				
Fall, 1972	Face - to - Face	7	224	212	12	1	409	409	0
Summer, 1973	Face - to - Face	8	227	213	14	2	334	325	9
Winter, 1974	Face - to - Face	9	225	203	22	3	281	268	13
Spring, 1975	Face - to - Face	10	225	191	33	4	257	238	19
Fall, 1975	Face - to - Face	11	220	185	35	5	253	219	24
Spring, 1976	Face - to - Face	12	221	184	37	6	235	215	20
Fall, 1976	Face - to - Face	13	215	180	35	7	212	194	18
Spring, 1977	Face - to - Face	14	207	165	42	8	199	176	23
Fall, 1977	Telephone	15	201	160	41	9	205	173	32
Winter, 1979	Telephone	16	210	146	64	10	203	165	38
Spring, 1980	Telephone	17	198*	N.A.	N.A.	11	177*	N.A.	N.A.

*-These are not final figures; followup efforts are still in process.
N.A. = Not Available

1. Panel members who have moved outside of the Peoria, Decatur, or Chicago SMSA since the initial interview are sent mail questionnaires. In addition for Waves 1 - 14 of the Peoria/Decatur Panel and for Waves 1 - 8 of the Chicago Panel, the male and female heads of the household each received separate self-administered questionnaires.

CONTINUED

7 OF 10

TABLE 2: SELECTED DATA CURRENTLY AVAILABLE -- CONSUMER PANELS OF YOUNG MARRIED COUPLES

TOPIC	Wave: Peoria/Decatur Chicago	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
								1	2	3	4	5	6	7	8	9	10	11
ECONOMIC BEHAVIOR																		
Durable goods purchases		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Durable purchase decisions		x	x					x			x	x			x	x		x
Relevance of factors in purchase decisions		x				x		x								x		
Durable goods purchase plans		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Durables purchase priority rank		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Home ownership				x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Income data				x				x	x	x	x	x	x		x	x	x	x
Income from job		x						x	x	x	x	x	x		x	x	x	x
Income relative to expenses				x											x		x	x
Awareness of price increases									x		x	x					x	x
Financial assistance and debts		x																
Finances and money management		x		x				x		x				x		x		x
Financial situation expectations, attitudes, etc.		x									x	x						x
Investments in variable dollar form (Ownership of, satisfaction with, financial services)				x				x			x						x	
Factors affecting investment decisions														x				
Shopping attitudes		x	x						x			x			x			
Credit cards		x		x		x			x				x				x	x
Family balance sheet				x				x		x				x				x
Attitude toward, decisions concerning savings		x			x	x		x			x	x			x			x
Saving rate					x	x					x	x				x		x
Use and ownership of savings accounts			x															
Awareness of ads on savings						x												
Ownership of life insurance		x		x						x							x	
Insurance other than life insurance		x			x						x							
LABOR FORCE PARTICIPATION AND SEX ROLE ORIENTATION																		
Employment (unemployment)		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Women and work		x																
Comparison of couples roles with parents																		
Discussion of decisions				x							x							x
BACKGROUND CHARACTERISTICS																		
Miscellaneous		x																
Age		x									x	x						x
Race																		
Draft status			x	x	x	x	x											
Religion																		
Education		x								x	x	x	x	x	x	x	x	x
GOALS AND ASPIRATIONS																		
Education		x		x														
Employment		x		x														
Income				x														
Homeownership		x																
CHILDREN																		
Fertility plans							x	x	x	x	x	x	x	x	x	x	x	x
Changes in family circumstances			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
Parents and children																		
GENERAL ATTITUDES AND PERSONALITY CHARACTERISTICS																		
Self-determination		x	x															
Self-assessment																		
TIME USE																		
Use of leisure time																		
Vacation plans and habits																		
Media preferences and exposure		x																
SOCIAL AND POLITICAL ATTITUDES																		

First and Second Marriages
The First Three Years of Married Life

Elizabeth Benson-von der Ohe, Ph.D.
University of Maryland (European Division)
Munich Campus
APO New York 09407

First and Second Marriages
The First Three Years of Married Life

Elizabeth Benson-von der Ohe, Ph.D.
University of Maryland (European Division)
Munich Campus
APO New York 09407

This study differs in several ways from most others described in this volume. It was not supported by a grant and consequently combined less-expensive data-gathering techniques (telephone interviews, mail questionnaires) with more-expensive ones (face-to-face interviews). And it followed pairs of people, rather than individuals, through time, a situation which intensified response-rate problems because an entire pair is "lost" whenever one member declines to participate.

The study also differs from most marriage research in that it is longitudinal and also included both spouses at each point in time. Given the now-well-known finding that "his marriage" and "her marriage" are not the same thing (Bernard, 1972; Safilios-Rothschild, 1969.) -- that is, that spouses give different answers even to simple questions--it is widely recommended that husbands as well as wives be contacted. This advice, like the advice that marriage research ought to be longitudinal, is given much lip-service. In both cases, there are clear reasons why it has seldom been done. Husbands are reluctant respondents, especially in the beginning; it is simply easier to limit one's research to wives. More important, to double the sample size (by including husbands) plus to double or triple the number of contacts per respondent (by making the study longitudinal) costs at least four times as much time and money. And, for

statistical purposes, this increase in sample size does not increase the "N" for most tables; the pair becomes the unit of analysis or, when not, husbands' and wives' data must be handled separately, because they are not independent of each other and because of possible gender response-bias.

Purpose of the Study

The purpose of the study can best be described in terms of the general questions it was intended to answer. The questions in Table 1 are those which ran through all three phases of the study plus a few which--because researchers as well as respondents change their interests over time--were added during the later stages.

Longitudinal Design

The decision in favor of longitudinal research was made because certain predictor variables (e.g., pre-marital communication) preceded the outcome variables in time and because of the conviction that accuracy of recall about early communication would deteriorate rapidly over time.

Table 2 introduces the design of data collection waves for the entire study (although the last phase was not a part of the original design-plans). There were three data-gathering stages (Times 1, 2 and 3) plus a preliminary postcard questionnaire for use in double-checking whether sampling criteria were met. Couples were interviewed as soon as possible after the wedding, again after about half a year of marriage, and finally--via mail questionnaire--between their third and fourth wedding anniversaries.

The problem of respondents' recalling information about the past has already been mentioned. Such recall data is especially likely to be biased when it deals with an emotionally laden subject like person's own previous marriage. Therefore it was decided to examine differences between first and second marriages primarily, but not entirely, by comparing second marriages with first marriages of other (but similar) couples. Maximizing this similarity meant selecting only certain kinds of couples and was achieved via the specification of eligibility or sampling criteria.

Sampling Criteria

It was clearly necessary to include first-marriage couples (as a control group) in addition to second-marriage ones. Third, fourth, fifth, and sixth marriages were excluded because of the increased chance of getting respondents with emotional problems.

Basically, the goal was to obtain as homogeneous a sample as possible. Hence the sample included only persons of U.S. or Canadian cultural background and those from a limited range of social-class levels. Attempts to make additional restrictions--e.g., about religion or education or to include the wife's as well as husband's occupational level as an indicator of social class--were abandoned because the number of eligible couples would have been reduced too drastically. Although race was not a selection criterion, the above-average social class requirement resulted in a racially homogeneous (almost entirely Caucasian) sample. Rural-urban variation was controlled not by sampling criteria per se but by choosing only urban counties in which to carry out the research.

The second purpose of the sampling criteria was to increase the probability of finding explicit marital negotiation behavior to study.

The very young and the less educated might not be inclined to question traditional marital roles and hence might not engage in the extensive communication often required to work out new roles. It seemed desirable to exclude them. This was done by limiting age-at-marriage for both spouses to a minimum of 22 (with an upper limit of 40). Education was not an explicit sampling criterion, but because it is highly correlated with occupational status, a high average educational level was nevertheless obtained by limiting the sample to higher occupational levels. (See Table 3.)

Of course, a good proportion of second marriages within the sample was indispensable. These were restricted to persons whose first marriages had ended in divorce, eliminating the widowed and the multiple divorced--of whom there are few in this age group anyway--in an effort to make the second-marriage sample more homogeneous with respect to past marital experience. Such second marriages are not rare in this age group but proved to be a fairly small proportion among upper-middle-class couples. Given the limited funding and hence limited sample size, it was clear that sufficient numbers of second marriages could not be obtained from a simple random sample; stratification was necessary.

Stratification of the Sample

The sampling pool (all couples meeting the above eligibility criteria) was subdivided into four categories, as shown in Table 4, with the aim of drawing equal numbers of couples from each category: (a) both spouses in first marriages, (b) husband in second marriage (but wife in first marriage), (c) wife in second marriage and husband in first, and (d) both in second marriages. This design results in equal numbers of individuals of each gender who have and have not been previously married. The "mixed cases" (b and c) were included because they provide the best test of certain hypotheses and because there would not have been sufficient second marriages of type d alone (both in second marriages, which was the least frequent of all four "marriage types").

This sample stratification, in combination with the sample criteria, was an attempt to create a quasi-experimental design, that is, to control for as many extraneous variables as possible such that the effect of theoretically relevant variables can be identified with more confidence. This is a particularly fruitful strategy for a study where funding precludes a very large sample size; with smaller samples, only one or two extraneous variables can be simultaneously "controlled for" later by statistical procedures. In a true experiment, selection criteria may also be used to control for extraneous variables. The difference is that varying "states" of the independent variables are created in a true experiment, while in a quasi-experiment one must take steps to insure finding such variations (e.g., stratifying the sample to insure finding enough second marriages, and limiting eligibility to higher status levels to maximize changes of finding sufficient cases of high marital communication).

All of this has the consequence that results cannot be generalized to any population not sharing this particular combination of characteristics. Judging the representativeness of one's sample also becomes more problematic, since published census descriptions of similar populations are seldom available.

Representativeness and Characteristics of the Sample

In lieu of census data as a comparison base for assessing sample representativeness, the maximum possible information about all couples in the sampling pool was obtained from marriage licenses and return-postcard questions). This allowed the comparison of respondents against non-respondents, to at least determine how similar the sample was to the pool of couples from which it was drawn (see Table 5). This was especially

important to ascertain, given the method by which couples were "drawn" from the sampling pool.

Under ideal circumstances, one would first have a complete list of all eligible respondents, subdivide them according to the stratification categories, and then draw a random sample from each category such that each couple had an equal chance of being chosen. But in order to have a sufficiently large sampling frame prior to the start of interviewing--when it is necessary to interview people soon after marriage--a much larger population base would have been required than the five million total population of the four standard metropolitan statistical areas included in the present research. During any one month, there simply were not enough eligible newlyweds to fill up the sampling frame.

The next best thing to a random sample was what might be called a saturation sample; one takes the list of eligible newly married couples and comes as close as possible to interviewing all of them, meanwhile continuing to gather names as more people get married and again interviewing every possible couple. Of course, the budget-determined quotas for the most frequently occurring marriage types were filled first. The gathering of new names among the less-frequent types continued until their quotas were also filled. Hence the first-marriage couples had weddings in February and March while second-marriage couples included April and early May dates in addition.

Among the four marriage types, the goal of saturation sampling was more nearly achieved for second marriages (especially the rarer wife-second and both-second ones) than for both-first couples. Reluctant second-marriage couples were "encouraged" to participate more than were first marriages, simply because such encouragement was absolutely necessary to get enough second-marriage respondents. In retrospect, it would have been possible and advisable to have done some sort of random sampling among both-first couples, drawing from that month's names about the same number of couples as were being interviewed for other marriage types.

Table 5 summarizes the characteristics of the sample, in total as well as broken down by marriage type, and compares these to the characteristics of the total sampling pool from which these couples came. The characteristics listed in the top half of Table 5 were taken from marriage licenses, while information about the remaining characteristics came from the return-postcard obtained for over half of the sampling pool, or telephone follow-ups conducted. The sample resembled quite closely the characteristics of the sampling frame (asterisk mark, in Table 5, those few instances where the fit was less than ideal and these deviations are also summarized in Table 6). Please note that all characteristics are as of the time of marriage.

It is also clear that the four marriage types differed somewhat from each other, despite the homogeneity built in via sampling criteria, in such characteristics as age and occupational status.

The processes of questionnaire construction and pretesting did not differ substantially from those in non-longitudinal research, and they will not be described here. The measures utilized and questions asked will be taken up later, in conjunction with the findings. The actual carrying out of the data gathering and analysis will be discussed next, for it was here that the bulk of the work (and problems) lay.

Locating and Contacting Respondents

Several decisions made in drawing up the research design had repercussions upon the process of locating eligible couples. The restrictive sampling criteria (in particular, the age and social-class limitations) required that the study cover a large geographic area in order to find enough respondents. At the same time, the decision that the first interviews should take place soon after the wedding meant that marriage-license applications were not useful if more than a few weeks old. As there was no published source containing complete names, and as some counties were so slow in forwarding marriage records to the state records bureau as to make this source unusable, periodic trips to the County Clerk's office in each of the seven counties were required. Access to marriage records was restricted, in practice although not by law, and had to be negotiated with each County Clerk. Only by intervention from the state records bureau was access to two of the most populous counties' records obtained.

Letters, with enclosed return-cards, were sent to couples immediately. Follow-up reminder letters were sent to the pre-marriage address of the other spouse and, finally, telephone contact was attempted. When initial contacts were made by telephone, it proved extremely helpful--as a means of convincing people that the study was "legitimate"--to begin by asking "just a few quick questions" (those factual ones which also appeared on the return-card) before mentioning the study itself. Secondly, because husbands were usually the more suspicious and less easily convinced, acceptance rates benefited dramatically from the employment of a caller who was both female and very charming.

The cost of such telephone follow-ups (and of the telephone interviews themselves) was made bearable by the availability of toll-free university tie-lines to most of the major urban areas covered by the research.

Telephone Interviews (Time 1)

For the actual telephone interview, each spouse was interviewed separately by an interviewer of the same gender. Experience in the pre-test interviews had suggested that women would be franker with female interviewing and men with male interviewing. In order to keep the gender of the interviewer, it was necessary to hire male interviewers, which are more difficult to find than are female ones.

An ever-present problem inherent in "linked-pair" research is that of obtaining individual, rather than jointly formulated, answers. For the Time-1 telephone interviews, respondents were directly asked not to discuss the interview with their spouses until both had been interviewed (They were told: "We want each person's opinions as an individual"). A similar statement was included in the Time-3 mail questionnaires, along with a request for a frank report about how much discussion had actually taken place ("it's important that we know"). The problem was handled in the Time-2 personal interviews by questioning spouses simultaneously in separate rooms.

A vital detail for the early phases of longitudinal research is to ascertain how people can be located later, in the event that they--like most newlyweds--move to a different address. In retrospect, the obtaining of the employer's address (in addition to that of a parent or friend) would have been advisable. Addresses of parents, taken from the marriage license records, also proved helpful later. Certain interviewers at Time-1 failed to get permanent addresses for couples who weren't planning to move, and the request for a permanent address at Time-2 was included in a pencil-and-paper section rather than in the interviewer-administered part of the questionnaire; both of these mistakes were very costly in terms of time and effort required to locate respondents later.

The questions asked in the first interview were deliberately non-threatening and fairly impersonal (in retrospect, perhaps unnecessarily so, even for a telephone situation). Whether for this or for other reasons, the telephone interviews must have been a good experience for participants because--when letters announcing the Phase 2 face-to-face interviews were sent out--over 70 percent accepted the invitation for a second interview before any follow-ups were made. This suggests that preliminary telephone interviews might be a useful tactic simply as a way to get cooperation for any personal-interview study.

Face-to-Face Interviews (Time 2)

The second wave of interviews was done by appointment in the couples' homes, by a male-female team. The major part of the 1½ hours was spent with each spouse separately. In a concluding segment of the interview, some open-ended questions were asked of the couple together which provided a chance to observe their interaction. Each interviewer separately rated his or her observations immediately afterwards.

Mail Questionnaires (Time 3)

The only reasonable technique for the last wave of data collection was via mail, because neither the research director nor many of the couples had remained in the original study-area. Phase 3 began with the sending of a long-promised report of selected findings. This report was of a descriptive nature which did not reveal the hypotheses being tested. The accompanying letter solicited agreement to participate further, in return for the receipt of additional findings on topics of the respondent's choice. Depending on the reply received, couples were then sent either the regular questionnaire or a modified version containing additional questions for those couples who were not divorced or separated. Each spouse received a separate questionnaire and separate return-envelope, along with the request not to discuss their answers (which most claimed to have carried out).

Although announced as the "Third Anniversary Study," this phase was not completed until some couples had reached their fourth wedding anniversaries due to the time required to locate people and to induce them to return their questionnaires. In retrospect, Dillman's (1978) idea of sending the last reminder by certified mail--for which he reports excellent returns--might have been a useful tactic.

Because a number of Time-2 respondents could not be located, participation was also solicited from persons who had taken part in Time 1 but not also in Time 2. This doubled the number of returns but proved to be only moderately useful. Key variables involving change (or other variables constructed with data from two time-points) often required Time-2 data; Time-1-only respondents were superfluous. For other analyses, of course, this procedure did increase the "N".

Most difficult of all to locate were those now-divorced persons not living at their original addresses. Retaining all possible information from the small number of divorced participants required a deviation from

the previous practice of eliminating all data from "incomplete couples." A number of non-respondents were also coded as being divorced, on the basis of reports by spouses and third parties, for use in comparing the early months of marriage of now-divorced versus non-divorced couples.

Measures Used

Knowing that spouses will often differ in their perceptions of the marriage and hence that their answers will differ, one interviews both of them. As expected, spouses do indeed give different answers to the same questions, sometimes radically different ones. But, after gathering this information from both, what does one do with it (beyond simply describing how spouses differ)? For instance, when it comes to hypothesis testing, how can one settle on a value for a given variable without ignoring the fact that values differ between wife and husband?

There are a number of ways of handling such answers from linked-pairs of respondents, most of which were tried in the present research. The simplest, but in some ways least informative approach is to construct each table twice, once with wives' data and once with husbands'. One can then see whether a given relationship holds for both genders, for neither, or only for one. The problem inherent in this approach is how to interpret findings when a relationship holds only for one gender but not for the other. For example, the hypothesis that husband-dominant couples would engage in less premarital communication was supported when using husbands' reports of communication but not for wives'. How, then, should one interpret such a finding?

The other basic alternative is to construct pair-variables, combining the answers of husband and wife into a single score. With longitudinal data, it is even possible--though not simple--to create change-over-time scores for pairs. Table 7 illustrates several ways in which pair-codes can be obtained, using data from one or from two time-points.

For measures and questions for which pair-scores were calculated, plus those used only for individual-scores, are summarized in Table 8. The exact wording of questions can be provided upon request.

Preliminary Findings

As of the time of this writing, the analysis of findings was still underway. A few preliminary findings are provided in Table 9.

In general, findings raise very serious doubts about the meaningfulness of simple generalizations about how first-versus second-marriage individuals differ. A number of previously reported generalizations of this type (e.g., that divorced men marry women much younger than themselves) held for some types of second marriages but not at all for others. And the variation among the several second-marriage types was often greater than that between first and second marriages. For the deriving of meaningful statements about how the nature and quality of married life are related to spouses' previous marital history, analysis in terms of the couple's joint marital history ("marriage type") appears far superior to that employing only the marital history (previously married versus not) of individuals.

References Cited

- Bailyn, L. Career and family orientations of husbands and wives in relation to marital happiness. Human Relations, 1970, 23, 97-113. Also in A. Theodore (Ed.), The Professional Woman. Cambridge, Mass.: Schenkman, 1979.
- Benson, E.E. Marital power: A theoretical explication and empirical exploring of three process models. Unpublished doctoral dissertation, Michigan State University, 1976.
- Benson-von der Ohe, E. First and second marriages: a description for beginning longitudinal researchers of one research project. Unpublished manuscript, 1979.
- Bernard, J. The future of marriage. New York: World, 1972.
- Blood, R.O., Jr., & Wolfe, D.M. Husbands and wives: the dynamics of married living. New York: Free Press, 1960.
- Dillamn, D.A. "Mail and telephone surveys" the total design approach. New York: Wiley, 1978.
- Duberman, L. The reconstituted family: a study of remarried couples and their children. Chicago: Nelson-Hall, 1975.
- Emerson, R.M. Power-dependence relations. American Sociological Review, 1962, 17, 31-40.
- Hoffman, L.W. Effects of the employment of mothers on parental power relations and the division of household tasks. Marriage and Family Living, 1960, 22, 27-35.
- Hill, R. Family development in three generations: a longitudinal study of changing family patterns in planning and achievement. Cambridge, Mass.: Schenkman, 1970.
- Navran, L. Communication and adjustment in marriage. Family Process, 1967, 6, 173-184.
- Orden, S.R. & Bradburn, N.M. Dimensions of marital happiness. American Journal of Sociology, 1968, 73, 715-731.
- Rodman, H. Marital power in France, Greece, Yugoslavia and the United States: a cross-national discussion. Journal of Marriage and the Family, 1967, 29, 345-352.

References (Continued)

Safilios-Rothschild, C. Family sociology or wives' family sociology? a cross-cultural examination of decision-making. Journal of Marriage and the Family, 1969a, 31, 290-301.

Safilios-Rothschild, C. Patterns of familial power and influence. Sociological Focus, 1969b, 2 (Spring): 7-19.

Warner, W.L., Meeker, M., & Eels, K. Social class in America, Chicago: Science Research Associates, 1949.

Table 1

Aims of the Study
on First and Second Marriages

General Questions

- A. Do people do things differently the second time around (as compared with similar first-marriage people)?
- For instance, do they choose different kinds of people to marry? Do they know each other better and "talk things over" more before they marry? Do they handle conflicts more effectively?
- B. Marital equality: if one wants an equal marriage, what does it take to get one? To keep one, after babies start arriving?
- Also, do second marriage people want (and/or get) more equal marriages? Is there more or different communications in equal marriages than in traditional ones? Are equal marriages any happier than traditional ones? If so, for whom? Under what conditions do marriages move toward more equality or slip away from equality over time?
- C. Marital happiness: could and have predicted who would be happy and unhappy (early warning signs)?
- Also, are second marriages any happier than first marriages? Is premarital communication related to later happiness? Do older (30+) people have the same chances for happiness as do younger ones? Are couples who cohabited before marriage any happier?
- D. Dual-careers: many want to have both highly-satisfying marriages and highly-satisfying careers, but only a few manage to achieve this. What did these few do differently?

More Theoretical Questions

- E. What are the conditions necessary for extensive communication, especially for early communication about desired marital arrangements? In particular, does it take a certain degree of power-equality between spouses before such communication will take place?
- F. Alternative theories of marital power: which theory provides a better fit with the data from this study?

Table I (Continued)

Aims of the Study on First and Second Marriages

Major Variables

Marriage type
(1st & 2nd
marriages)

Marital power/
influence
Sex-role attitudes
and marital-role
preferences
Marital communica-
tion

OUTCOMES
Marital happiness
Marital equality

Table 2

Design of Data Collection Waves

TIME 0 1-2 mos.	TIME 1 4-5 mos.	TIME 2 2½-3½ yrs.	TIME 3
Application for marriage license N = 566 couples	1/2 hour TELEPHONE INTERVIEWS with each spouse	1-1/2 hour PERSONAL INTERVIEWS with both spouses (separately & together)	Mail QUESTIONNAIRES from both spouses
POSTCARD QUESTIONNAIRE N = 316 couples	N = 156 couples	N = 90 couples	N = 67 couples

Additional data from divorced or separated individuals, whose mates did not respond, is also available for Time 3. About half of the Time 3 couples had participated in all previous phases, while others had participated only through Time 1.

Table 3

Criteria for the Sampling Pool

Eligibility criteria:

- 1) Both spouses between ages 22 and 40;
- 2) Husband employed in a middle-middle, upper-middle, or upper status occupation (the top three categories of the seven-point Revised Scale for Rating Occupations, Warner et al., 1949);
- 3) Both spouses either:
 - a) entering a first marriage, or,
 - b) entering a second marriage preceded by divorce (rather than widowhood);
- 4) Both spouses born in the United States or Canada.

The sampling frame consisted of all couples meeting the above criteria who applied for marriage licenses between February 1 and May 15, 1974 (for both-first marriages, only until April 1) in seven urban counties in Michigan (encompassing the greater Detroit area, Ann Arbor, Grand Rapids, and Lansing-East Lansing).

Table 4

Design of the Stratified Sample

The sampling pool was stratified according to couples' "marriage type," such that roughly equal numbers of couples were obtained in each of four categories:

HUSBAND IN:	1st marriage	a (both 1st	c (wife 2nd)
	2nd marriage	b (husband 2nd)	d(both 2nd)

Table 6

"Representativeness" of the Sample

As can be seen from Table 5 the respondents were quite similar to urban couples in this age group, from this Midwestern region, and of this social class level (roughly, upper-middle class). There were a few instances, however, where the sample was less representative:

- | | |
|------------------------|---|
| Total Sample: | some overrepresentation of couples who brought a child (from the previous marriage) into the new marriage;

otherwise, no real differences (despite the fact that the sample differed greatly from the sampling-pool comparison base in the proportions of various marriage types, due to sampling stratification). |
| Both-1st marriages: | considerable overrepresentation of Protestants (and of conservative religions in general);

Underrepresentation of couples where husband's occupational status exceeded that of the wife (equal-status couples were overrepresented). |
| Husband-2nd marriages: | some overrepresentation of younger husbands and of husbands in professional and technical jobs (the highest status level). |
| Wife-2nd marriages: | no real differences from the sampling pool. |
| Both-2nd marriages: | considerable overrepresentation of wives in the highest occupational-status level (and hence, fewer-than-expected couples where the husband's occupational status exceeded that of the wife);

some overrepresentation of wives employed full time. |

Table 7

Types of Pair-Variables

constructed by combining answers from linked-pairs of respondents

Below are some approaches which were tried. For simplicity, it is shown how each could be applied to the same response-item, a question about marital happiness (a 9-point scale ranging from not at all happy to ideally happy).

A. Using Data from One Time-Point Only

1. Discrepancy scores
 - *a) equal happiness, small discrepancy, large discrepancy (or: small, medium, large discrepancies),
 - **b) absolute difference between spouses' rating (e.g., if her rating minus his rating = "-3", code as "3").
2. Directional-discrepancy scores
 - *a) husband happier, about equal, wife happier (or: husband much happier, some happier, etc.)
 - **b) husband's score minus wife's score (resulting pair-value ranges from -8 to +8; add 9 to eliminate signs, such that range is from 1 to 17).
3. Anchored directional-discrepancy scores
 - a) combined answers of spouses (e.g., wife very happy and husband very happy, wife very happy and husband fairly happy, wife very happy and husband not very happy; wife fairly happy and husband very happy, etc.).
 - b) every possible combination of answers given a separate code (e.g., wife 1 and husband 1=1, wife 1 and husband 2=2; wife 1 and husband 3=3; etc.)

NOTE: Anchored directional-discrepancy scores preserve the most information but must always be machine-recorded for use. From the same pair-score, one can recode to generate any of the other types of pair-variables.
4. Averages
 - a) wife's score plus husband's score, divided by 2 (content-wise, averages are usually of questionable validity; two "pretty happy" spouses would receive the same score as a couple where he was "very happy" and she "not very happy").
5. Evaluative outcome-codes (usually for dependent variables)
 - ***a) couples dichotomized into both-positive-outcome vs. all other combination possibilities (e.g., both-very-happy vs. other).

NOTE: The purpose of dichotomizing is to allow the use of control variables on samples of limited size; if the independent variable also has few categories, one can use a control variable and still maintain some minimal cell-size on tables.

B. Using Data from Two Time-Points

1. Change measures (same question asked both times)

Table 7 Continued

Types of Pair-Variables

constructed by combining answers from linked-pairs of respondents

- a) pair-scores first calculated separately for each time-point, then coded for change (e.g., for type-a discrepancy scores, increase in agreement on happiness, no change, decrease in agreement).

NOTE: This only works where it's possible to obtain a meaningful and simple pair-value at individual time points.

- b) Compute change-measure for each individual separately, then combine (e.g., both decrease in happiness, husband only decreases, wife only decreases, neither decreases).

2. Non-change measures

- a) assessing relative achievement of spouses of their respective earlier wishes (I have no happiness example, but can give one from my pair-indicator of marital power. On a five-point scale from "husband almost always" to "wife almost always," one asks division-of-labor preferences at Time 1 and actual division at Time 2. Suppose that he prefers a "wife almost always" division, she prefers a "both about the same" division, and both report actually having a "wife almost always" division. His preference-actual difference is/5-5/=0. Her preference-actual difference is/3-5/=2. The pair score (a directional discrepancy is 0-2+2 where "-2" indicates a disadvantage to the wife, i.e., his preferences won out over hers. (For details, see Benson, 1976, pp 40-44.)

NOTE: The rationale for asking wishes (preferences, goals, etc.) at a somewhat earlier time point than asking about achievement of these, is simply that--if wishes were asked later--respondents might be reluctant to admit having wished for something they didn't get.

- * This can also be coded after the collapsing of individual answers into fewer categories, e.g. very happy, pretty well, and not very happy.
- ** This measure, more often calculated by computer, can be later collapsed into fewer categories.
- *** In some cases, three (rather than two) categories are the minimum that make any sense, e.g., for marital powers: husband dominant, equal, wife dominant.

Table 8

Summary Of Measures Available For Major Variables

Marriage Type (See Table 4)

- 1) Individual: first marriage vs. second marriage (marital history).
- 2) Couple: both-first marriage, his second, her second, both-second marriage (joint marital history).

Marital Equality:

Time 1: -Attitudes (including traditional sex role ideology, Hoffman, 1960).
-Preferences: for 10-item division of labor.

- Time 2: *-Blood & Wolfe, 1960, decision-making index (of marital power, very slightly reworded).
-Division of tasks and decisions (18 items, including birth-control decision, taking sexual initiative, child discipline, voting, various household tasks, financial items).
-Opinion about women working (Bailyn, 1970): of self and perceptions for spouse and circle-of-friends.
-Relative personality traits of spouses (including initiative, giving in, etc.).
-10-item division of labor: a) current division b) satisfaction with division c) preferred division d) importance of task or decision (plus: how and why arrived at current division for each item).
-Interviewer ratings (adapted from Hill, 1970): including who talks more, shows more initiative, who has more say in joint-decisions.
-Influence techniques (adapted from Safilios-Rothschild, 1969b).
Also: who usually starts quarrels and *who usually gives in.

- Time 3: -10 item division of labor: as at Time 2 but without how and why questions. Added: items about household tasks in general, child-care in general, *giving in when there is a disagreement.
*-Blood & Wolfe decision-making index.
-Perceived fairness of division of all work (total of housework, child-care, paid employment, etc.): from much more than my fair share" to "much less."

Table 8, (continued)

Major Indicators I Used:

- a) Attitudes: "Hoffman traditional sex role ideology (Time 1)
Opinion about women working (Time 2)
- b) Reported outcomes of marital power:
 - 1) traditionalism of marital roles (from T3, division of household tasks in general): Traditional = wife does more than half.
Non-traditional = equal division or husband does more.
 - **2) constructed influence scale, based on whose division-of-labor preferences won out in actuality (for the 8 non-child items from the 10).
 - a. Time 2 actual vs. Time 1 preferences
 - b. Time 3 actual vs. Time 2 preferences
 - c. Time 3 actual vs. Time 3 preferencescollapsed into: husband dominant, equal, wife dominant at each time point.

Marital Communication:

- Time 1: -Quantity of communication (hours during previous 6 months) about 10 general decision-topics.
-Perceived quality of communication (for same 10 topics): similarity of spouses' ideas, progress toward reaching decision or understanding, sureness about where spouse stands, touchiness of topic.
- Time 2: -Quantity and quality of communication since Time 1 (similar to Time 1 but only for 2 topics chosen by respondent -- one where spouses' ideas are the most different and one where ideas are somewhat different).
-Personal communication style (openness of expressing dissatisfactions to mate).
-Quality of communication: Primary Communication Inventory (Navran, 1967).
-Relative personality traits of spouses (including who talks more, who's more likely to say so if dissatisfied, etc.).

Table 8, (continued)

- Communication about division of labor (among answer possibilities for how arrived at division for specific items).
- Interviewer ratings: parallel vs. coordinated conversation, who interrupts more, who talks more (HILL, 1970).
- How much, and what, tell spouse about your day and your work? How much does spouse tell you? Wish there were more or less of such communication? etc.

Time 3: -Changes over last 3 years (much more to much less), in talking about: future plans, problems or disagreements, what happened during the day, who will do which household tasks.

Major Indicators Used:

- a) Quantity of communication (Time 1): average across all non-child items.
- b) Quality of communication:
 - 1) perceived similarity, sureness, touchiness (Time 1): average across all non-child items.
 - 2) Navran Primary Communication Inventory (Time 2).

Marital Happiness:

- Time 2: -Satisfaction with 13 aspects of married life (plus importance of each for making a marriage work out and how different from original expectations).
-Interviewers' ratings: seem happy (yes or no) and how compare with other couples interviewed (above average, average, below).
- Time 3: -Changes over last 3 years: in satisfaction with marriage, with life in general, in interest in marital sex, etc.
-Orden-Bradburn (1968) scales of marital satisfaction and marital dissatisfaction.
-Marital happiness (9-point scale) for self, perception of spouse, and -- if these differ - why.
-Report on major crises, long-standing issues faced.
-Relative marital happiness scale (a great deal more to a great deal less).
 - 1st marriages: how much different than originally expected (and why).
 - 2nd marriages: how much different from 1st marriage (and why).

Table 8 (continued)

- Advice would give to friend, about how to get new marriage off to a good start.
- Divorce questionnaire: listed 30 possible reasons for divorce, rated for how much each bothered respondent and how much apparently bothered spouse.

Major Indicators I Used:

- *-Marital happiness (9-pt. scale) of self, collapsed into very happy, pretty happy, and not very happy (for many purposes the pair-version was dichotomized into both-very-happy vs. other) (Time 3).
- Orden-Bradburn satisfaction and dissatisfaction scales (Time 3).

Additional Aspects of Dual-Career Situations:

- Time 1:
 - Communication questions included decisions about jobs of both spouses.
 - Women: projected work plans, by age of youngest child.
 - Facts about current work situation, occupation of self and parents.
- Time 2:
 - Work history (years worked part and full time).
 - Many questions about present work situation -- how much work, how tiring, bringing work home, how difficult to find time for various other things, talking about work with spouse, type of help exchanged, etc.
 - Rank-order of top-3 sources of satisfaction (adapted from Bailyn, 1970): from list of 9 items, including job, relation with spouse, being a parent, etc.
 - Income; also, who earns more, would it bother you or spouse if wife earned more, banking arrangements (joint or separate accounts, etc.).
 - Career ambition.
 - Satisfaction with work.
- Time 3:
 - Changes since last interview: in employers, hours, pay, etc.
 - Facts about current job.
 - Satisfaction with work.
 - Labor-saving appliances; help received for child care or house-and-yard tasks.
 - Comparison of spouses' jobs (rated as husband more, same, wife more): hours worked, job prestige, pay, physical fatigue, emotional energy invested, chances of advancement.

Table 8 (continued)

Major Indicators I Used:

- *-Combined marital and job satisfaction measure (dual-career couples only). The pair-version, for many purposes, was dichotomized into both-spouses-high (i.e., both reported very high marital satisfaction and both reported being very or extremely satisfied with their jobs) vs. other (Time 3).

Miscellaneous:

- Time 1:
 - Perception of chances of men vs. women to find good spouse.
 - Mate-selection criteria.
 - Cohabitation: interviewer reports on any indications.
 - Education, religious activity, number of children want, ages of present children, etc.
- Time 2:
 - Dating history, info. about length of 1st marriage, etc.
 - *-Cohabitation (and length).
- Time 3:
 - For each child: age, sex, whose child it is, plus perceived closeness with respondent and respondent's spouse (both earlier and at present) (partly adapted from Duberman, 1975).
 - Number of different houses/apartments lived in during this marriage.

-
- * Pair-measures also constructed, from answers of both spouses.
 - ** Solely a pair-measure.

Table 9

Selected Findings

to date

The analysis is currently in progress. A few of the findings so far are organized according to the general questions from Table 1.

DO PEOPLE DO THINGS DIFFERENTLY THE SECOND TIME AROUND?

1. Do they choose different kinds of people to marry?
 - compared with the mate-selection criteria of 1st-marriage people: 2nd-marriage men* (especially if there were children in the first marriage) much more often chose a wife because of her "ability to be romantic"; 2nd-marriage women more often chose a husband for his "dependability" and "ability to be a good parent."
2. Do they know each other better and "talk things over more"?
 - compared with 1st-marriage people: 2nd-marriages couples** had shorter courtships but a much higher frequency of premarital cohabitation; 2nd-marriage men reported being more sure of their wives' opinions on a variety of decision-issues.
 - compared with their own first marriages; 2nd-marriage people reported having achieved via "special effort" a different sort of 2nd-marriage w.r.t. the ease of talking over problems, how disputes are settled, and the maturity of both partners. In addition, women reported achieving differences in the division of decision-making responsibility, and men in how practical or romantic their own approach to marriage was.
 - both-2nd couples, followed by both-1st, reported the greatest quantity of pre-marital "talking things over."
 - both-2nd women did an outstanding job in predicting what their mates and marriages would be like at Time 2.

MARITAL EQUALITY

1. Do 2nd-marriage people want more-equal marriages?
 - no, they don't. In fact, 2nd-marriage man--while themselves no more traditional in attitudes than other men--more often chose wives having traditional attitudes.
2. Do they get more-equal marriages?
 - with respect to the division of labor at Time 3: no, not according to wives' reports; according to husbands, wife-2nd marriages are the most traditional and both-2nd marriages the most equalitarian.
 - with respect to my pair-indicator of relative influence: wife-2nd couples had the least equality (most husband-dominance) at Time 2.
 - with respect to a number of marital power indicators, there is preliminary evidence that 2nd-marriage husbands tend to be more dominant than 1st-marriage husbands.

Table 9 Continued

Selected Findings

to date

3. Is there more or different communication in equal marriages?
 - for husbands' communication reports: the greater the husband's influence-dominance, the less the communication quantity; the less traditional the attitudes, the greater the communication quantity (especially high among the liberal-minded who also achieve a non-traditional division of labor at Time 3).

MARITAL HAPPINESS

1. Are 2nd-marriages any happier than 1st-marriages?
 - no, differences at Time 3 are not significant (though wife-2nd spouses tend to be less happy).
 - 2nd-marriage wives report fewer dissatisfactions but also fewer satisfactions, than do 1st-marriage wives.
 - 2nd-marriage husbands also report fewer dissatisfactions. For men, dissatisfactions were much more strongly related to global happiness than were satisfactions.
2. Do older (30+) people have the same chances for happiness?
 - yes, for men, the chances are even slightly better.
 - age-discrepancy between spouses made no difference, except that couples where the husband was much older than the wife were happier than average.
3. Are couples who cohabited before marriage any happier?
 - it depends; long-cohabitation was generally associated with lower happiness; for couples with younger husbands, short-cohabitation was the best; with older husbands, no cohabitation couples had the highest happiness.

DUAL-CAREERS (no findings yet)

CONDITIONS NECESSARY FOR EXTENSIVE COMMUNICATION (see B-3 above)

ALTERNATIVE THEORIES OF MARITAL POWER

1. Which micro-theory of marital power provided the best "fit" with this data? (see also Benson, 1976).
 - in test-situations where alternative theories would predict different outcomes, power-dependence theory (Emerson, 1962) received somewhat more support than did the resource theory of marital power (Blood & Wolfe, 1960), regarding the theoretical "mechanisms" by which resources get translated into power.
 - within power-dependence theory, hypotheses about "alternatives" receive more support than those about "motivational investment".

*2nd-marriage men (or women) are those individuals who have been divorced-and-remarried. (In "wife-2nd" couples, for example, the wife but not the husband would be a 2nd-marriage individual.)

**"2nd-marriage couples" are all couples in which at least one spouse has been married before (wife-2nd, husband-2nd, and both-2nd).

The Consumer Panels of Young Married Couples

Eileen Crimmins
Department of Sociology
Rutgers University - Camden

Linda Lannom
Survey Research Laboratory
University of Illinois

Phyllis A. Ewer
Department of Sociology
University of Illinois - Chicago Circle

Patterns of Child Rearing (PCR)

Robert R. Sears
Stanford University

Patterns of Child Rearing (PCR)

Robert R. Sears

Stanford University

Overview

Rather unexpectedly, the research reported in Patterns of Child Rearing (Sears, Macoby & Levin, 1957) became the basis for a continuing longitudinal study of the relation between children's early rearing experiences and their later personality development. The initial study was planned and executed by the staff of the Laboratory of Human Development of Harvard's Graduate School of Education during the Winter and Spring of 1951-52. It was designed simply to test some hypotheses concerning the child rearing determinants of identification in 5 year old children. The independent variables (rearing) were measured by coding the transcripts of two-hour taped interviews with 379 mothers of children enrolled in public school kindergartens in Newton and Watertown, Massachusetts. The dependent variable was measured by three sessions of doll play, a method which, at the time, was presumed to provide a good measure of children's feelings of identification with their parents.

A content analysis of the interviews provided measures of 167 variables describing: (1) the mother's self-reported actions and feelings toward her child; (2) her perception of the father's behavior and feelings; and (3) role relationships between husband and wife. In addition, there were 21 variables measuring the mother's perception of her child's behavior. These 188 variables provided the data reported in Patterns, and are the only data collected directly from the mothers.

The breadth and depth of coverage of relevant variables, however, together with the large number of cases for whom comparable measures were available, offered a challenging opportunity for further research.

Hence, over the years a number of investigators have been inspired to search out subsamples of the children in order to discover possible antecedent-consequent relations between early childhood experiences and later measures of personality development.

There have been several waves of such child measurement between 1951 and 1979. In addition to the original doll play measures, which were secured from all 379 children, there have been numerous other measures obtained from subsamples of 20% to 50% of the original group. These measures have related to school achievement in third grade, aggressive feelings, adult role attitudes, self-concepts and gender-role typing at age 12, level of aspiration, psychopathological symptoms and hypnotizability in later adolescence and a large number of personality qualities, including interview as well as standardized projective and questionnaire test measures at the subjects' ages 30 to 32. Most of the measures obtained in these follow ups have been preserved in computer-useable form. The study continues under the supervision of a committee appointed by the Henry A. Murray Research Center of Radcliffe College.

Sample Selection

The purposes of the original 1951-52 study determined the nature of the sample selected. These purposes were partly normative -- to secure a broad sample of self-reported child rearing behaviors and attitudes -- and partly hypothesis-testing with respect to the antecedents of children's identification with their parents. The former purpose, which was fulfilled ultimately by the publication of Patterns, led to the selection of a stratified sample by sex and ordinal position of the child and by socio-economic status of the family. The second purpose required that all the children be approximately five years old, and that a few infrequently occurring child or family deviations be avoided. These exclusions were intended to control-by-elimination some possibly major sources of idiosyncratic variation in child rearing which could not, within the size of the projected sample, be controlled by subsample comparisons.

With the assistance of the Newton and Watertown school administrations, eight public schools were chosen as the sources from which the final sample of children and their mothers were selected. Six were in Newton (a rather strongly middle-class community) and two from Watertown (with a greater working class representation). The neighborhood locations

of the schools were such that the staff expected to find a reasonable representation of lower-and middle-class families. School census data indicated that the 640 children enrolled in the eight kindergartens constituted 84.5% of the children in those districts. The other 14.5% were enrolled in parochial schools, which were not sampled in the research. Many Catholic children attended the public kindergartens, however, because not all parochial schools had such facilities.

The sample selection process is reviewed in Table 1, which lists the bases for the final exclusion of 41% of the public school children potentially available. The upper half of the table is self-explanatory except for the two "disproportionate" items. The aim was to secure 50 children of each sex at each of four ordinal positions, with half from each SES group. In fact, only children were seriously underrepresented in the source population and the middle-class families were more numerous than working class. The "disproportionate" exclusions were the extra cases for the other three ordinal positions and for middle class status who were not needed to fill their respective cells in the stratified design.

The final sample contained 379 mothers and their kindergarten-aged children, one child per mother. The stratification is shown in Table 2. Some other demographic characteristics are listed in Table 3 (see also Patterns, pp. 24-27).

Insert Table 1, 2, 3 about here

The Measures of Child Rearing

Method

A standardized interview with open-ended questions and pre-constructed probes was used for securing the mothers' reports of their attitudes, feelings and behaviors relating to child rearing. All questions were asked of every mother, while the probes were used only as needed. Ten mature and well-trained women interviewers conducted the recorded interviews in the mothers' homes.

Typists later transcribed these recordings, converting personal names to case identification numbers. These transcriptions were then

coded independently by two coders, using a set of rating scales that served to define the dimensions of child rearing to be measured. This double coding permitted the determination of coding reliability. It also markedly improved the accuracy of the final measures, since when the two coders disagreed by only one point on a 5-point scale, their ratings were averaged, but when greater disagreement occurred, the coders re-examined the transcript together and reached agreement on a pooled rating. Full details of the interview construction, procedure and coding are given in Patterns (pp. 12-24); the interview itself is published, with all probes, as Appendix A (pp. 491-501); the main rating scales used in coding the interviews, together with their respective reliabilities, are listed in Appendix B (pp. 501-510).

Content

The theoretical orientation that determined which dimensions of child rearing were measured stemmed from three sources, a behavioral psychology of learning and action, psychoanalysis and cultural anthropology. This theoretical orientation was responsible for the emphasis on:

1. learning process variables such as rewards, punishments, disciplinary methods, permissiveness, restrictiveness and reinforcement contingencies;
2. the social setting provided by the family as represented by expressed standards of conduct, values, achievement goals, inter-parental feelings and attitudes, pervasive feelings of parents toward their child such as warmth and rejection and influences on the parents of the child's sex and ordinal position;
3. major motivational systems of two kinds: (a) those that were organized around basic biological drives which required control (hunger, elimination, genital sexuality), and (b) those that involved the development of strongly dyadic elements (aggression, dependency, achievement, sex, conscience and self-control).

Thus, the interview was designed to secure information about the stimulus conditions provided by the parents and their methods of control and training with respect to the child's major motivational and action systems. Some of the information referred to actions and feelings covering limited periods of time (e.g., weaning, toilet training, infant care-taking, feelings of parents when the pregnancy was discovered; or later, demands for chores, opportunity to earn money, school anticipations). Other information, however, was not time-limited and the ratings

of control or training methods summarized qualities of parental behavior covering the whole first five years of the child's life, e.g., affectional warmth and various punitiveness and permissiveness scales, as for aggression to parents, masturbation, social sex play and dependency supplications; and customary methods of discipline and the use of various kinds of rewards or incentives. These longer-time-inclusive measures were intended as indicators of certain stable and enduring qualities of the parents' behavior which the investigators expected to be predictive of such stable outcomes in the child as self-control and conscience.

In 1975, Prof. Martha Kent used a different theory to construct additional dimensions of parental qualities. Using the same coding methods, she recoded all the interviews with the new scales. These were based on a linguistic theory of affect-expression and represent an important supplement to the original measures.

Follow Up Studies

First: 1955

Most of the children were in third grade in the Spring of 1955 and Harland Abbott (1960) gave them Stanford Achievement Tests. The purpose was to search for the effects of different early child rearing experiences on school performance. These test scores later became useful measures of early competence as antecedent to self-esteem (Sears, 1970).

Second: 1958

The Spring of 1958 was the last time at which the members of this cohort, now age 12, would be together in their original schools, in the sixth grade. With most of them just approaching adolescence, it seemed to the Laboratory staff an opportune time to begin follow up studies. Two of these utilized dependent variables providing additional tests of the original identification theory. A third focused on a variety of other personality dimensions for which behavior theory suggested some testable hypotheses concerning the effects of early rearing experiences. These studies were performed with 160 children who were a quite exact subsample of the total original sample (cf. Sears, 1961).

For one of the identification theory tests, Grinder (1962) designed an assessment situation for measuring a child's resistance to temptation. A "ray gun" marksmanship task was presented to the subject and was self-scored in the absence of the experimenter, giving the child a chance to

cheat in reporting his score. Resistance to this temptation was one presumed outcome of high identification. Heinicke's earlier finding of a correlation between maternal warmth and expressions of guilt at the fantasy level was not replicated in the comparison of warmth with real-life resistance, but Heinicke's other findings with respect to the effects of high parental demands and restrictions were -- in boys; high standards for neatness, orderliness and obedience in the early years were clearly associated with resistance to temptation at age 12. For girls, stringent weaning, bowel training and sex control were the corresponding correlates. (The theory did not account for this sex difference.) Children whose mothers judged them to show signs of strong conscience at age 5 more frequently resisted temptation in this school-world task at age 12.

The third set of measures included in this 1958 follow up included five questionnaire scales measuring aspects of aggressive feelings (anti-social, prosocial, projective, self-aggression and aggression anxiety), four scales measuring aspects of self-concept (self-criticism, ideas of reference or hypersensitivity to others and two self-evaluation inventories), and masculinity-femininity. Comparisons of these "feelings" measures with the early child rearing measures supported the following hypotheses:

1. the development origins of aggression anxiety are substantially different for boys and girls; for boys such feelings stem from rapid socialization of aggression in a love-oriented atmosphere, while for girls they derive from a masculinization process induced by high infusion of the father in rearing and caretaking (Sears, 1961);
2. self-aggression in boys is produced by early, punitive, successful control of aggression in the family;
3. high prosocial and self-aggression both develop in children who are originally highly aggressive in the home, but whose behavior is eventually controlled by punishment;
4. high self-esteem at age 12 derives from (a) competence in school work, (b) early ordinal position, (c) in a small family, (d) where at least one of the parents is warm and supportive, and (e) in boys, the father is not overly dominant toward the mother (Sears, 1970).

Nothing from early child rearing predicted masculinity-femininity at age 12.

Third follow up: 1964

With the names and original addresses of the subjects available, David Nowlis (1969) worked through school records and located a subsample of 82 subjects who were high school seniors in the Winter of 1964. His purpose was to test a developmental theory of hypnotizability earlier proposed by the Hilgards (1962). The latter had found, with retrospective child rearing experience reports from college-age hypnotic subjects, that firm parental discipline in childhood enhanced later hypnotizability and was associated with trance-like experiences.

Nowlis secured measures of both these qualities, from the Patterns subjects, and was then able to test the Hilgards' theory by appeal to data actually collected from the mothers at the relevant age of the children. All the significant correlations obtained in this comparison supported the hypothesis. These included pressure for conformity to table standards, frequency of spanking, use of scheduled feeding in infancy and punishment for aggression. There were also significant correlations with the 1958 aggression anxiety measure and adult-child role choice, both of which variables were also associated with early socialization pressures.

Fourth follow up: 1964-65

The following autumn another group of investigators sought out a subsample composed mainly of subjects still living at their original addresses. Most of them were now graduated from high school. Crowne, Conn, Marlowe and Edwards (1969) located and secured the cooperation of 83, who proved to be a good representative subsample of the original 379. This follow up was a search for developmental antecedents of various patterns of level of aspiration behavior. They used the Rotter Level of Aspiration Board and scored a number of response patterns for comparison with the child rearing scales. When the several patterns were reduced to the customary three D-scores often used in level of aspiration research (negative, low and high positive), clear developmental relationships emerged. Defensive high goal setting, known to be associated with failure, was significantly associated with earlier maternal punitiveness toward aggression and dependency, high maternal anxiety and a reported prevalence of obedience problems with the child. The failure-avoidant response pattern (negative D-scores) seemed clearly to have emerged from maternal warmth and protectiveness, possibly a kind of infantilization. These findings, coupled with those of p. 5 Sears (1940, 1941), suggest that the anlage for realistic-unrealistic approach to achievement-arousing life situations is established quite early in life, and is displayed from school age on into late adolescence and young adulthood.

Fifth follow up: 1968

Four years later, 64 of the original subjects agreed to participate in a study of psychopathology and of styles of social interaction. These 27 males and 37 females, while a small subsample, were found to be representative of the original participants.

The Minnesota Multiphasic Personality Inventory (MMPI) was administered as a measure of psychopathology. Results suggested that the interaction of different parental practices related to particular psychopathological dimensions and that these dynamics were different for male and female children (Rousell & Edwards, 1971). Among males, there tended to be an association between an excessively warm-nonpunitive home atmosphere in childhood and the development of anxiety and psychotic disturbances. Phobic and psychotic symptoms tended to have emerged from a cold-punitive home atmosphere. For females, a warm-punitive or permissive-punitive home atmosphere appeared to be related to neurotic disturbances or paranoid reactions.

Participants' preferred styles of social interaction were measured by the Situational Preference Inventory (Edwards, 1973). Different interactive styles were found to be associated with very different developmental antecedents. Cooperational interaction was found to be characteristic of young adults who, as children, had close relationships with the parent of the opposite sex. In addition, parental discipline and withdrawal of love placed contingencies upon those relationships and served to reinforce cooperational behavior. Instrumental styles which emphasize authority and adherence to tradition were associated with strict child rearing and high standards. Aggression was more likely to be punished in females and reinforced in instrumental males. Analytic interactive styles characterized by exploration of other than existing or normative courses of action were found associated with minimal restrictions and greater permissiveness and by parental behavior which suggested that parents provided a measure of security to accompany their children's greater freedom.

Sixth follow up: 1977-79

The following decade produced much mobility among the group and by 1976 it was obvious that if these subjects were to become part of a longer term life-cycle study, serious effort would have to be expended on locating their current whereabouts. With support from the Foundation for Research in Child Development, Prof. George Goethals directed such a search and was able to locate addresses for about 95% of the subjects. In the following year, Prof. David C. McClelland began a large and intensive follow up study with a subsample nearly as large as that of the first (1958) follow up. The subjects were now 30-32 years of age.

The dependent variables for McClelland's research include:

1. TAT measures of n Achievement, n Affiliation and n Power;
2. Rosenthal's PONS, an ambiguous voice test for measuring non-verbal sensitivity;
3. moral development tests; and
4. interview measures of a host of adult behavioral variables are still in process; preliminary analyses suggest that scheduling of feeding and severity of toilet training are related to adult n Achievement, permissiveness about sex and aggression to adult n Power and mother's affection and lack of punitiveness to adult measures of maturity.

The Data Bank

A computer tape containing all previous measurements of the Patterns group, both mothers and children, has been deposited in the Henry A. Murray Research Center of Radcliffe College. The intent is to make these data available, in coded anonymous form, to any serious researcher who wishes to make further analyses of them. The transcripts of the mother interviews themselves are presently maintained at Stanford but they, too, will eventually be transferred to the Center. The correspondence file relating to the actual writing of Patterns of Child Rearing is on permanent deposit in the Archives of the Stanford University Library.

The Patterns subjects are a resource of no small value. A Committee for supervising further research contacts with them has been appointed by the Murray Center.

References

Data Reports from the Study

- Abbott, H.C. A study of the relationship between child rearing practices and school achievement. Unpublished Ed.D. dissertation, Harvard Graduate School of Education, 1960.
- Crowne, D.P., Conn, K.K., Marlowe, D. & Edwards, C.N. Some developmental antecedents of level of aspiration. Journal of Personality, 1969, 37, 73-92.
- Edwards, C.N. Interactive styles and social adaptation. Genetic Psychology Monographs, 1973, 87, 123-174.
- Grinder, R.E. Parental childrearing practices, conscience, and resistance to temptation of sixth grade children. Child Development, 1962, 33, 803-820.
- Havighurst, R.J. & Davis, A. A comparison of the Chicago and Harvard studies of social class differences in child rearing. American Sociological Review, 1955, 20, 438-442.
- Heinicke, C.M. Some antecedents and correlates of guilt-fear in young boys. Unpublished doctoral dissertation, Harvard University, 1953.
- Levin, H. & Sears, R.R. Identification with parents as a determinant of doll play aggression. Child Development, 1956, 27, 135-153.
- Maccoby, E.E. The taking of adult roles in middle childhood. Journal of Abnormal & Social Psychology, 1961, 63, 493-503.
- Maccoby, E.E. & Gibbs, P.K. Methods of child rearing in two social classes. In W.E. Martin and C.B. Stendler (Eds.), Readings in Child Development. New York: Harcourt, Brace, 1954.
- Milton, G.A. A factor analytic study of child-rearing behaviors. Child Development, 1958, 29, 381-392.
- Nowlis, D.P. The child-rearing antecedents of hypnotic susceptibility and of naturally occurring hypnotic-like experience. International Journal of Clinical and Experimental Hypnosis, 1969, 17, 109-120.
- Rousell, C.H. & Edwards, C.N. Some developmental antecedents of psychopathology. Journal of Personality, 1971, 39, 362-377.

Sears, P.S. Child-rearing factors as related to playing of sex-typed roles. American Psychologist, 1953, 8, 431 (abstract).

Sears, R.R. Relation of early socialization experiences to aggression in middle childhood. Journal of Abnormal & Social Psychology, 1961, 63, 466-492.

Sears, R.R. Relation of early socialization experiences to self-concepts and gender role in middle childhood. Child Development, 1970, 41, 267-289.

Sears, R.R., Maccoby, E.E. & Levin, H. Patterns of child rearing. New York: Harper & Row, 1957. Reprinted by Stanford University Press, Stanford, Ca., 1976.

Other References

Dollard, J., Doob, L.W., Miller, N.D., Mowrer, O.H. & Sears, R.R. Frustration and aggression. New Haven: Yale University Press, 1939.

Dollard, J. & Miller, N.D. Personality and psychotherapy. New York: McGraw-Hill, 1950.

Hilgard, J.R. & Hilgard, E.R. Developmental-interactive aspects of hypnosis: Some illustrative cases. Genetic Psychology Monographs, 1962, 66, 143-178.

Sears, P.S. Level of aspiration in academically successful and unsuccessful children. Journal of Abnormal and Social Psychology, 1940, 35, 498-536.

Sears, P.S. Level of aspiration in relation to some variables of personality: Clinical studies. Journal of Social Psychology, 1941, 14, 311-336.

Sears, R.R. A theoretical framework for personality and social behavior. American Psychologist, 1951, 6, 476-483.

Sears, R.R., Whiting, J.W.M., Nowlis, V. & Sears, P.S. Some child-rearing antecedents of aggression and dependency in young children. Genetic Psychology Monographs, 1953, 47, 135-234.

Whiting, J.W.M. & Child, I.L. Child training and personality. New Haven: Yale University Press, 1953.

Table 1

Selection of the Sample

Total number of kindergarten children in eight schools	640
Shrinkage due to sample requirements:	
1. Either or both parents foreign born	38
2. Child not living with natural parents	13
3. Parents separated, divorced, dead	15
4. Twins	8
5. Disproportionate ordinal position	41
6. Disproportionate socio-economic status	38
7. Other (child deaf, retarded, etc.)	6
Total shrinkage for sample requirements	159
Total number of kindergarten children after initial selection	481
Shrinkage due to cases becoming unobtainable:	
1. Mother refused to participate	38
2. Mother did not have time to be interviewed	24
3. Sickness in family	7
4. Family moved	19
5. Child would not cooperate in doll play	7
6. Interview recording defective	3
7. Unknown reasons	4
Total shrinkage for dropouts	102
Total sample	379

Source: Sears, Maccoby and Levin, 1957, p. 511.

Table 2

Composition of Sample: Frequency in Terms of Sex and Ordinal Position of Child and Socio-economic Status of Family*

	Ordinal Position			
	Only	Oldest	Middle	Youngest
Boys:				
Middle class	16	28	33	35
Working class	17	30	15	24
Girls:				
Middle class	4	23	27	32
Working class	15	25	29	19

*The total obtained sample was 379 mothers, but 7 of them have been omitted from this table because no estimate of their SES could be made. Of the omitted cases, 4 were boys' mothers and 3 were girls'; 3 of them were mothers of oldest children, 1 was the mother of a middle child and 3 were mothers of youngest children.

Source: Sears, Maccoby and Levin, 1957, p. 512.

Table 3

Other Demographic Characteristics of the Sample

	Mother	Father
Median age	33.6 years	36.5 years
Education		
Less than high school	14%	16%
High school or some college	64%	48%
Finished college	22%	36%
Family income		
Median	\$7,150	
Interquartile range	\$3,700-8,275	
Ethnic background of mothers	Middle class	Working class
Italian	7	36
Irish	40	30
Jewish	64	15
Old American	33	19
Contemporary British	22	32
Mixed	6	10
Not ascertained	<u>26</u>	<u>32</u>
Number of cases	198	174

The Hyperkinetic Child Grows Up:
Predictors of Symptoms, Delinquency, and Achievement at Follow-Up*

Jan Loney, Ph.D.
John Kramer, Ph.D.
Richard Milich, Ph.D.
Department of Psychiatry
University of Iowa

*Supported by grant MH-22659 from the National Institute of Mental Health, by grants from the Iowa Mental Health Research Fund, and by grant DA-02134 from the National Institute on Drug Abuse.

The Hyperkinetic Child Grows Up:
Predictors of Symptoms, Delinquency, and Achievement at Follow-Up

Jan Loney, Ph.D.

John Kramer, Ph.D.

Richard Milich, Ph.D.

Department of Psychiatry

University of Iowa

Over the past two decades, activity by researchers has produced hundreds of scientific papers about a condition variously called childhood hyperkinesis, minimal brain dysfunction, or attention deficit disorder. As a result, we now know a good deal about what is generally acknowledged to be the most common behavior problem of childhood. We also know that numerous important questions remain unanswered. For example, the short-term effectiveness of treatment with CNS stimulant medication in increasing attention and compliance was rapidly and conclusively shown, but the accumulation and verification of facts about long-term outcome have been slower in coming. For both parents and professionals, knowledge about the adolescent and adult outcome of childhood hyperactivity is important for informed decisions about management.

Our own attention has been focused upon potential predictors of outcome. Specifically, we have used multivariate statistical techniques (a) to identify those variables from the referral and early treatment periods that predict variation in adolescent behavior and (b) to estimate the relative importance of the identified variables in accounting for that variation. To date, both child symptom dimensions (e.g., aggression) and socioecological environment measures (e.g., parenting styles) have proven to be significant predictors of adolescent symptoms (Loney, Langhorne, Paternite, Whaley-Klahn, Broeker, & Hacker, 1976) and of delinquent behavior (Kramer & Loney, 1978). The goals of this paper are to summarize these previous studies and to extend our analyses to the prediction of academic achievement.

The data for all analyses to be reported here were drawn from several multivariate computer archives which have been accumulated and stored as part of an ongoing longitudinal study of the hyperkinetic/minimal brain dysfunction (HK/MBD) syndrome. The main study group consists of 135 nonretarded four- to twelve-year-old boys - all of whom displayed HK/MBD symptoms and were referred for psychiatric evaluation at the University of Iowa Child Psychiatry Clinic between January 1, 1967, and September 1, 1972; all of whom began treatment with CNS stimulant drugs at the University of Iowa within six weeks after their outpatient evaluation; and 92% of whom (N = 124) were followed up an average of five years later, when 79% of them were between 12 and 18 years old (mean age = 13.9 years). Independent variables for the present analyses were drawn from the referral and treatment archives; dependent variables were drawn from the follow-up archive. A more complete account of the procedures that were followed and the data that have been accumulated is presented elsewhere (Loney, et al., 1976).

Methods

Subjects

Of all four- to twelve-year-old boys seen for outpatient diagnostic evaluation by the University of Iowa Child Psychiatry Service between January 1, 1967, and September 1, 1972, 482 had not been drug-treated within the previous six months. Sixty-two percent of these boys (N = 300) were recommended by a senior psychiatrist for a trial of cerebral stimulants because their history and symptom picture suggested minimal brain dysfunction (Clements, 1966). Of these 300 boys, 59 were excluded from the sample either because their tested IQs were less than 70 or because they were suffering from peripheral sensory loss, major psychosis, epilepsy, cerebral palsy, or unequivocal brain damage.

Because our original aim was to identify predictors of initial clinical response to Ritalin and Dexedrine, the sample was limited to outpatient boys whose treatment with these drugs was initiated and followed by staff psychiatrists and/or psychiatric residents of the University of Iowa Child Psychiatry Service. It was further required that drug treatment had been initiated within six weeks of the day of the outpatient evaluation and that at least one treatment progress note was placed in the child's medical record before termination of drug treatment, addition of another drug or therapy, or referral to a local physician for drug follow-up.

Application of these exclusionary criteria resulted in a total clinical sample of 135 boys with a mean age of 8.2 years. The group includes boys from both rural and more urban areas of Iowa and several surrounding states and, like the Iowa general population, is 98% white. Referrals of children in our sample originated with local health professionals (40%), school personnel (15%), the child's family (12%), social agencies or juvenile courts (12%), or multiple sources (21%). A wide range of socioeconomic backgrounds is represented. Thirty-five percent of our sample were seen as private (full-fee) patients, 35% chose to be seen as clinic (reduced-fee) patients, and 30% were seen as indigent (no-fee) patients. Expressed in terms of Hollingshead and Redlich's socioeconomic classes (1958), 21% of our families fell into Classes I and II; 33% into Class III; 26%, Class IV; and 20%, Class V. Average family income was \$9615, compared with an Iowa 1970 census average of \$9018. The mean IQ for the 97 boys who were administered the WISC at or prior to the outpatient evaluation was 100.43 (S.D. = 14.64; range = 72 to 139).

Diagnostic and Treatment Procedures

During the period of time between the child's original referral and the outpatient evaluation, the parent(s) completed a 163-item Child Intake Form and 29-item Mother and Father Intake Forms. During the same time period, the child's teacher completed an 81-item School Intake Form.

Outpatient diagnostic evaluations were scheduled one per morning and supervised by whichever one of two senior psychiatrists was assigned to the day in question. The psychiatric resident, social worker, educational consultant, and psychologist saw the child and his family during the morning. In the afternoon the outpatient staff presented verbal accounts of their findings at a staff meeting with the senior psychiatrist, who was responsible for making the psychiatric diagnosis, prescribing a course of treatment, and dictating a lengthy staff summary. Each of the outpatient staff members also dictated a detailed narrative report of his/her interactions with and impressions of the child and/or parents; these were also typed and placed in the child's chart as were the medical-psychiatric summary, the social summary, the psychological report, and the educational report.

After the staff evaluation, about one-third of our sample was treated by one of the senior psychiatrists. The remaining cases were assigned for drug regulation and follow-up to a psychiatric resident, usually the one who had done the initial outpatient diagnostic workup. Although there was some general supervision of these treatment cases by the senior psychiatrists, the resident had primary responsibility, and it was he or she who wrote prescriptions, conducted follow-up interviews with child and parent(s), coordinated the efforts of other staff involved with the family, and dictated progress notes summarizing the course of the child's treatment.

Follow-Up Study

Between September 1, 1973, and January 1, 1976, we carried out a long-term follow-up study of the 135 boys whose records had been employed for the referral study. As part of these efforts, we repeated most of the original home and school intake forms; obtained a variety of behavioral ratings and questionnaires from parents, teachers, and the boys themselves; conducted parallel forms of a structured interview with each boy and each of his parents; and administered an extensive battery of cognitive, personality, and achievement tests. An average of five years elapsed between referral and follow-up. Ninety-two percent of our referral sample (124 boys and their parents) were seen at follow-up.

Selection of Samples of Children

A more homogeneous sample of 84 boys was obtained from the 124 seen at follow-up by limiting analyses to data from boys who were: (a) between 6 and 12 years of age at the time of the outpatient evaluation; (b) living with both a mother and a father figure at the time of referral; (c) placed on Ritalin within six weeks of the outpatient evaluation; and (d) between 12 and 18 years of age at the time of follow-up. Such a strategy may reduce the size of the intercorrelations by restricting the range on some variables, but it allows generalization to a more precisely-defined population.

The analyses that were carried out to predict adolescent symptoms were first reported in 1976 at a meeting of the Society for Life History Research in Psychopathology (Loney, et al.). The analyses that were carried out to predict teen-age delinquency were first reported in 1978 as part of a symposium on juvenile delinquency at a meeting of the American Psychological Association (Kramer & Loney). The analyses that were carried out to predict adolescent achievement were first reported in 1978 at an NIMH Workshop on childhood hyperactivity. Because we added several familial and achievement predictors to the delinquency and achievement analyses and chose to eliminate all cases with incomplete data on any predictor or criterion, those analyses involve a residual subsample of 30 cases. For many of the discarded cases, initial School Intake Forms were missing, and therefore teacher ratings of achievement at referral could not be determined.

Selection of Independent Variables

Potential predictor variables (Table 1) were selected to sample data across three environmental categories (ecological, familial, and treatment), two child categories (symptoms and achievement), and one miscellaneous category. These potential predictors were included in the multiple regression analyses as independent variables. The single-starred variables were

Table 1

Independent Variables Included in Multiple Regression Analyses as Potential Predictors

Environmental Variables

Ecological

Socioeconomic status (SES)
Rural or urban residence*

Familial

Intactness of biological family*
Length of time that biological triad was intact*
Number of intervening changes in living situation*
Number of children in family*
Maternal love-to-hostility rating
Paternal love-to-hostility rating
Maternal autonomy-to-control rating
Paternal autonomy-to-control rating

Treatment Variables

Response to CNS stimulant treatment (methylphenidate)
Length of CNS stimulant treatment*

Child Variables

Symptoms

Aggression factor score
Hyperactivity factor score

Achievement

Past academic performance rating*
Reading rating*
Arithmetic rating*

Miscellaneous Variables

Age at onset of symptoms**
Age at psychiatric evaluation
Neurological signs rating
Perinatal complications rating

*Variable not included in analyses to predict symptoms at follow-up

**Variable not included in analyses to predict delinquency and achievement at follow-up

not included in the original analyses to predict symptoms at follow-up (Loney, et al., 1976); the double-starred variable was not included in the later analyses to predict delinquency (Kramer & Loney, 1978) and achievement.

Ecological Variables

Socioeconomic status (SES) is a weighted composite of the scores assigned to the father's educational level and to his occupational class (Hollingshead & Redlich, 1958; Lesser, Fifer, & Clark, 1965). Rural or urban residence separates boys living on rural mail routes (27% of sample) from those with street addresses in towns or cities (73%).

Familial Variables

Three variables reflect changes in the children's living situation. Living situation was defined in terms of the specific adults (e.g., biological father, stepfather, grandparent, houseparent) with whom the child resided, and a living situation was considered to have ended whenever it was interrupted for longer than one month. One of these variables assesses the intactness of the biological family at the time of the psychiatric evaluation by separating boys into those who were living with both biological parents (60% of sample) and those who were not (40%). A second variable specifies the number of months that all members of the biological triad (father, mother, and child) lived continuously together (length of time biological triad intact). A third variable enumerates the number of intervening changes in living situation between the one involving the original biological triad and the one in which the child was living at the time of his psychiatric evaluation.

Assessment of the number of children in the family (mean = 3.4) is based on individuals below the age of 18 who lived in the same home with the child. Full, half-, and step-siblings who shared the child's living situation are included in the assessment; those who lived elsewhere are not included.

Four measures of parenting style are included. Maternal love-to-hostility was rated on a 7-point scale (1 = much love, 7 = much hostility, and 4 = a midpoint representing equal amounts of each). The two graduate students who independently rated this variable were provided with definitions and examples from previous questionnaires and studies (Schaefer, 1965), and they based their rating on the staff, medical, and social summaries and on the psychological and educational reports contained in each child's psychiatric chart. Paternal love-to-hostility, maternal autonomy-to-control, and paternal autonomy-to-control were similarly rated, and the numbers assigned by the two raters were averaged to produce a composite rating for each of the four measures of parenting style. Effective rater reliabilities (Rosenthal, 1974) ranged from .51 to .83 for these measures.

Treatment Variables

The medication progress notes in the charts were of several kinds: written summaries by a clinic staff person of information from a parent, written summaries of observations of and interviews with the child during a return visit, letters reporting on the child's condition, repeated educational and/or psychological test reports, etc. A secretary retyped each progress note, removing any reference to the child's name, age, or grade in school; the names of all staff persons involved in the case; and any information indicative of socioeconomic class, parent-child relationship, or other variables hypothesized to be related to improvement.

The 800 notes were then placed in random order, and a child psychiatrist and a child psychologist independently rated each one on a 5-point scale according to the response to medication (bad to dramatically positive) that it was considered to represent. Factor analysis reduced several potential drug-response measures to a single dimension that accounted for most of the variation in the set of measures. The measures that loaded on this response to treatment factor were the average ratings given the child's first response to medication, first response to the highest dose of medication, highest-rated response, last response to medication, overall response, and mean response over the first three months. Effective rater reliabilities for the various drug response measures ranged from .86 to .90. An additional child treatment variable was one indicating the number of months that the child remained on medication (length of treatment).

Progress notes detailed an average of seven months of treatment at the University of Iowa clinic. During that time, the sample as a whole received an average daily dosage of 34 mgs. of Ritalin. Sixty-eight percent of the sample were considered by one or both raters to have shown a generally or totally positive response to this initial CNS stimulant drug treatment. Subsequently, some boys terminated treatment, others changed to or added other drugs or therapies, and still others were transferred and followed by their local physician. As a group, they remained on medication for an average of 29 more months.

Child Symptom Variables

Two advanced graduate students in school psychology independently reviewed each child's staff, medical, and social summaries and psychological and educational test reports and then rated the severity/frequency of a set of operationally-defined symptoms (Loney, et al., 1976). By means of factor analysis, two relatively independent dimensions were identified in the judges' chart ratings of thirteen of these symptoms at referral (Loney, Langhorne, & Paternite, 1978). The first dimension accounted for 45% of the factor variance and consisted of negative affect (e.g., excitability, irritability), aggressive interpersonal behavior (e.g., fights, won't mind),

and control deficits (e.g., delinquent acts, evasion of rules). The second dimension accounted for 23% of the factor variance and consisted of hyperactivity (e.g., restlessness, running around), inattention (e.g., distractibility, forgetfulness), and judgment deficits (e.g., impulsivity, immaturity). These two dimensions were labelled Aggression and Hyperactivity, respectively, and they are used in the present investigation as the symptom measures at referral. Effective rater reliabilities for the symptoms that formed the basis for the Aggression and Hyperactivity factor scores ranged from .74 to .82. Each of the symptoms was seen by one or both raters as frequent or constant at referral for at least half of the boys.

Child Achievement Variables

Assessments of the child's academic achievement were included in the 81-item School Intake Form that the child's teacher completed before his evaluation by the outpatient staff. Past academic functioning was rated on a 5-point scale ranging from poor to excellent; current reading achievement and arithmetic achievement were rated as satisfactory (or A, B, and C grades), unsatisfactory but improving, or unsatisfactory (D and F grades). Teachers considered 62% of the children to have been functioning at a borderline or poor level prior to the psychiatric evaluation, while only 8% were considered to have been functioning at a good or excellent level. Forty-seven percent were rated as unsatisfactory (failing) in current reading and in current arithmetic achievement.

Miscellaneous Variables

Age at onset of referral problem (mean = 5.8 years) was determined from the answer to a question ("How old was the child at the time that the problem was first noticed?") on the Child Intake Form completed by the parent(s) before the outpatient diagnostic evaluation. One-fourth of the parents said they had noticed the problem for which the child was referred when he was less than four years old. Age at evaluation (mean = 8.2 years) was a straightforward computation based on the child's birthdate and the date of the outpatient evaluation.

The psychiatric resident's neurological findings (e.g., observations of fine and gross motor movement and tests of synkinesia, graphesthesia, tandem walking, cranial nerves, postural reflexes, etc.) were recorded on a physical examination form and summarized in the medical-psychiatric report. Synopses of these findings were presented to a pediatric neurologist and a pediatric psychiatrist, who rated them independently for soft signs of neurological impairment on a 5-point scale. Similar ratings of perinatal complications were made by the same pediatric psychiatrist and an obstetrician, based on material reported by the mother and summarized in the medical-psychiatric report (e.g., reproductive history, details of

pregnancy and delivery, child's condition at birth, etc.). The effective rater reliabilities for these two sets of synopses were .89 and .83 respectively. The two neurological ratings for each boy were averaged to produce a composite rating of neurological signs, and the two perinatal ratings were averaged to produce another composite rating of perinatal complications. Moderate to strong indications of neurological signs were seen by both raters in 16% of the cases and of perinatal complications in 14% of the cases.

Selection of Dependent Variables

Dependent or criterion variables (Table 2) were chosen to provide information about three areas of adolescent outcome: adolescent symptoms, adolescent delinquency, and adolescent achievement.

Adolescent Symptom Variables

Measures of adolescent symptomatology were derived from the mother's description of her boy on a symptom checklist administered at follow-up. The number of symptom descriptors checked by a mother within a specific symptom category served as his severity score on that symptom. More than two-thirds of the boys were considered by their mothers to have some degree of problem within each of the five symptom categories presented here. Analyses of two kinds of aggressive symptoms (negative affect and aggressive interpersonal behavior) and three kinds of hyperactive symptoms (hyperactivity, inattention, and judgment deficits) are shown for illustrative purposes. A more detailed treatment of these analyses and of additional analyses of adolescent symptom ratings made by psychologists at follow-up is given elsewhere (Loney, et al., 1976).

Adolescent Delinquency Variables

Measures of adolescent delinquency were obtained from independent structured interviews with the boy, the mother, and the father at follow-up. Each informant's view of the boy's involvement in several categories of illegal or norm-violating behavior was rated on a 4-point scale (no involvement to chronic involvement) by a trained rater. For the present analyses, a composite involvement score within each category was derived by considering the maximum involvement reported by any informant to be valid. For example, if the boy reported occasional drinking and the parents reported none, the boy was scored as drinking occasionally. Conversely, if the parents reported that the boy drank occasionally and the boy said he didn't drink at all, the boy was again scored as drinking occasionally.

Table 2

Dependent Variables Included in Multiple Regression

Analyses as Adolescent Outcome Criteria

Symptom Variables

As Rated by Mother at Follow-Up

Negative Affect
Aggressive Interpersonal Behavior
Hyperactivity
Inattention
Judgment Deficits

Delinquency Variables

As Reported by Boy and Parents at Follow-Up

Offenses against Persons
Offenses against Property
Involvement with Alcohol
Involvement with Illegal Drugs

Achievement Variables

From Wide Range Achievement Test at Follow-Up

Score on WRAT Reading Subtest
Score on WRAT Arithmetic Subtest
Score on WRAT Spelling Subtest

Analyses of four broad a priori categories of involvement in illegal behavior are presented here. The category, offenses against persons, included a variety of aggressive and sexual acts (e.g., assault, intimidation, rape, peeping). Most of the 21% of our subjects who obtained scores in this category were involved in physical fights. The category, offenses against property, included such acts as shoplifting, vandalism, firesetting, and breaking and entering where material damage or theft was involved. Twenty-nine percent of our subjects were reported to have had some involvement in such offenses, usually shoplifting or minor vandalism. The category, involvement with alcohol, included any possession and use while a minor and drunkenness at any age. Some involvement with alcohol was reported for 28% of our sample, as compared with 37% of a group of Iowa sixth-graders and 84% of a group of Iowa twelfth-graders surveyed by Hays (1976). The category, involvement with drugs, included possessing, furnishing, selling, or using one or more of a variety of illegal substances (e.g., marijuana, stimulants, sedatives, hallucinogens). Involvement with illegal drugs was reported for 19% of our sample; sixteen percent of the boys themselves said they had smoked marijuana. By way of comparison, a survey of unselected Iowa sixth-, eighth-, tenth-, and twelfth-graders (Hays, 1976) produced 2%, 11%, 22%, and 33% respectively, who had used marijuana at least once and from 6% to 17% who had tried other drugs.

Adolescent Achievement Variables

Measures of adolescent achievement were derived from the boys' performance at follow-up on the Wide Range Achievement Test. Scores on subtests assessing reading (word recognition), arithmetic, and spelling serve as the dependent variables in analyses to predict adolescent academic functioning. On each of the WRAT subtests, fully two-thirds of our sample scored two or more years below the level appropriate for their grade placement at follow-up. Thirty percent were below the 15th percentile on the WRAT Reading subtest, 54% were below the 15th percentile on the Arithmetic subtest; and 56% were below the 15th percentile on the Spelling subtest.

A Model for the Prediction of Outcome in Childhood Hyperactivity

A generalized summary of the multivariate relationship between child and environmental variables on the one hand and adolescent outcome variables on the other hand is offered in Figure 1. This model begins with the hypothesis that ecological variables and child aggression, which are correlated at referral (Loney, et al., 1976), contribute to adolescent symptoms and delinquency. It is further hypothesized that treatment variables, childhood achievement, and childhood hyperactivity, which are also intercorrelated (Loney, Prinz, Mishalow, & Joad, 1978), contribute to adolescent achievement. However, only family variables are associated with all four categories of outcome. Childhood aggression and childhood hyperactivity, specifically, are assumed to have different correlates both at referral and at follow-up.

PREDICTORS OF ADOLESCENT OUTCOME

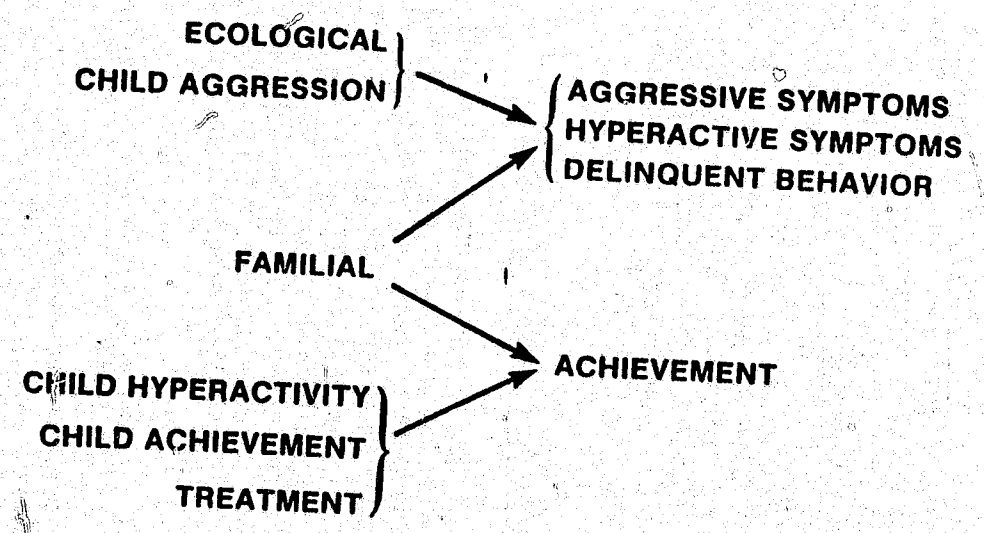


Figure 1

This model, if valid, would explain why drug treatment does not lead to improved adolescent behavior and reduced delinquency: Although drug treatment affects childhood hyperactivity, behavioral outcome and subsequent delinquency are determined instead by childhood aggression and by its ecological antecedents--which are not affected by drug treatment (Milich & Loney, in press). This model would suggest that, in order to improve eventual adolescent behavior and reduce delinquency among HK/MBD children, one would want to reduce their aggression and/or improve their environment. In short, it is postulated that drug treatment is ineffective in reducing ultimate problem behavior because child hyperactivity is not the first link in a chain that leads to teen-age symptoms and delinquency.

References

- Barkley, R. A., & Cunningham, C. E. Do stimulant drugs improve the academic performance of hyperkinetic children? A review of outcome studies. Clinical Pediatrics, 1978, 17, 85-92.
- Borland, B. L., & Heckman H. K. Hyperactive boys and their brothers. Archives of General Psychiatry, 1976, 33, 669-675.
- Clements, S. D. Minimal brain dysfunction in children--terminology and identification (NINDS Monograph No. 3, U.S. Public Health Service Publication 1415). Washington, D.C.: U.S. Government Printing Office, 1966.
- Cunningham, C. E., & Barkley, R. A. The role of academic failure in hyperactive behavior. Journal of Learning Disabilities, 1978, 11, 15-21.
- Dykman, R. A., Peters, J. E., & Ackerman, P. T. Experimental approaches to the study of minimal brain dysfunction: A follow-up study. Annals of the New York Academy of Sciences, 1973, 205, 93-108.
- Gittelman-Klein, R., & Klein, D. F. Methylphenidate effects in learning disabilities. Archives of General Psychiatry, 1976, 33, 655-664.
- Hays, M. Report of the Iowa study of drug behaviors among youth. Des Moines: State of Iowa Department of Public Instruction, 1976.
- Hechtman, L., Weiss, G., Finkelstein, M. A., Werner, A., & Benn, R. Hyperactives as young adults: A preliminary report. Canadian Medical Association Journal, 1976, 115, 625-630.
- Hollingshead, A., & Redlich, F. Social class and mental illness. New York: Wiley, 1958.
- Huessy, H. R., Metoyer, M., & Townsend, M. Eight-ten year follow-up of 84 children treated for behavioral disorder in rural Vermont. Acta Paedopsychiatrica, 1974, 40, 230-235.
- Kramer, J., & Loney, J. Predicting adolescent antisocial behavior among hyperactive boys. Paper presented at the annual meeting of the American Psychological Association, Toronto, 1978.
- Langhorne, J. E., Jr., & Loney, J. A four-fold model for subgrouping the hyperkinetic/MBD syndrome. Child Psychiatry and Human Development, 1979, 9, 153-159.

- Laufer, M. W. Long-term management and some follow-up findings on the use of drugs with minimal cerebral syndromes. Journal of Learning Disabilities, 1971, 4, 519-522.
- Lesser, G., Fifer, G., & Clark, D. Mental abilities of children from different social class and cultural groups. Monograph of the Society for Research in Child Development, 1965, 30: Serial No. 102.
- Loney, J. The intellectual functioning of hyperactive elementary school boys: A cross-sectional investigation. American Journal of Orthopsychiatry, 1974, 44, 754-762.
- Loney, J., Langhorne, J. E., Jr., & Paternite, C. E. An empirical basis for subgrouping the hyperkinetic/minimal brain dysfunction syndrome. Journal of Abnormal Psychology, 1978, 87, 431-441.
- Loney, J., Langhorne, J. E., Jr., Paternite, C. E., Whaley-Klahn, M. A., Broker, C., & Hacker, M. The Iowa HABIT: Hyperkinetic/aggressive boys in treatment. Paper presented at the meeting of the Society for Life History Research in Psychopathology, Ft. Worth, Texas, 1976.
- Loney, J., Prinz, R. J., Mishalow, J., & Joad, J. Hyperkinetic/aggressive boys in treatment: Predictors of clinical response to methylphenidate. American Journal of Psychiatry, 1978, 135, 1487-1491.
- Mendelson, W., Johnson, N., & Stewart, M. A. Hyperactive children as teen-agers: A follow-up study. Journal of Nervous and Mental Diseases, 1971, 39, 393-399.
- Milich, R., & Loney, J. The role of hyperactive and aggressive symptomatology in predicting adolescent outcome among hyperactive children. Journal of Pediatric Psychology, in press.
- Paternite, C. E., & Loney, J. Childhood hyperkinesis: Relationships between symptomatology and home environment. Chapter in Hyperactive Children: The Social Ecology of Identification and Treatment. Whalen, C., & Henker, B. (Eds.), Academic Press, in press.
- Rie, H. E., Rie, E. D., Stewart, S., & Ambuel, J. P. Effects of methylphenidate on underachieving children. Journal of Consulting and Clinical Psychology, 1976(a), 44, 250-260.
- Rie, H. E., Rie, E. D., Stewart, S., & Ambuel, J. P. Effects of ritalin on underachieving children: A replication. American Journal of Orthopsychiatry, 1976(b), 46, 313-322.

Rosenthal, R. Estimating effective reliabilities in studies that employ judges' ratings. Journal of Clinical Psychology, 1974, 30, 342-345.

Schaefer, E. Children's reports of parental behavior: An inventory. Child Development, 1965, 36, 413-424.

Sprague, R., & Berger, B. D. Drug effects on learning performance: Relevance of animal research to pediatric psychopharmacology. NATO Conference on Learning Disabilities, Cantraken, Canada, 1978.

Sprague, R., & Sleator, E. K. Methylphenidate in hyperkinetic children: Differences in dose effects on learning and social behavior. Science, 1977, 198, 1274-1276.

Stewart, M. A., Mednelson, W. B., & Johnson, N. E. Hyperactive children as adolescents: How they describe themselves. Child Psychiatry and Human Development, 1973, 4, 3-11.

Weiss, G., Kruger, F., Danielson, U., & Elman, M. Effect of long-term treatment of hyperactive children with methylphenidate. Canadian Medical Association Journal, 1975, 112, 159-165.

Weiss, G., Minde, K., Werry, J. S., Douglas, V., & Nemeth, E. Studies on the hyperactive child: VIII. Five-year follow-up. Archives of General Psychiatry, 1971, 24, 409-414.

Whalen, C. K., & Henker, B. Psychostimulants and children: A review and analysis. Psychological Bulletin, 1976, 83, 1113-1130.

Life Assessment For a Cohort
of Chronic Schizophrenics
Discharged Twenty Years Ago

Courtenay M. Harding, R.N.
Research Associate and Project Director
Department of Psychiatry
University of Vermont, Burlington, Vermont

and

George W. Brooks, M.D.
Superintendent, Vermont State Hospital
Waterbury, Vermont
Professor of Clinical Psychiatry
University of Vermont

Presented in part, at the 21st Annual Meeting of the Group-Without-A-Name International Psychiatric Research Society, Burlington, Vermont, April 27-29, 1979.

Research funded by NIMH Grant MG-29575.

The authors wish to thank numerous consultants, especially Robert Shapiro, M.D., Brendan Maher, Ph.D., Jane Murphy, Ph.D., and Bonnie Spring, Ph.D. at Harvard; Joseph Fleiss, Ph.D. at Columbia; Lee N. Robins, Ph.D. at Washington University School of Medicine; Jon Rolf, Ph.D. and Takamaru Ashikaga Ph.D. at the University of Vermont for their assistance on the design of this project.

Life Assessment For a Cohort
of Chronic Schizophrenics
Discharged Twenty Years Ago

Courtenay M. Harding, R.N.
and
George W. Brooks, M.D.

This chapter describes our experience in the formulation of a comprehensive design assessing both the cross-sectional status and the longitudinal study of 253 long-term, profoundly disabled, former patients released to the community from Vermont State Hospital twenty years ago. We will also report our efforts devoted to insure replicability of the study and to resolve some methodological loopholes found in earlier long-term follow-up studies.

Subjects

Unlike most research samples, this cohort from THE VERMONT STORY (Chittick et al, 1961) was chosen specifically for its chronicity (see Table 1). A more thorough description of the group can be found in the literature. (Chittick et al, 1961; Brooks and Deane, 1965; Deane and Brooks, 1967; and Daum et al, 1977). The cohort was described in the mid 1950's as "quite characteristic of the schizophrenic population group" as outlined by Hollingshead and Redlich (1958). They were middle-aged, poorly educated, lower class individuals further impoverished by repeated and prolonged hospitalizations, (Chittick et al, 1961). They were enrolled in an innovative program jointly sponsored by the Department of Mental Health and the

Department of Vocational Rehabilitation. All but three members of the original cohort were discharged.

With THE VERMONT STORY cohort we are able to conduct such a long-term follow-up study because the Senior Investigator has maintained personal contact with many members of the cohort over the intervening years. Eighty-nine percent of the original group have been located and accounted for. All but 13 live in the State of Vermont.

The stability of Vermont's population has made the research atmosphere similar to that of Scandinavia. Nearly all of the original cohort members, their children, other family members, original investigators, caseworkers, and general practitioners are still living within state borders. We do have the rare opportunity to conduct both a cross-sectional assessment and a longitudinal retrospective analysis of patterns, shifts and trends demonstrated during the last twenty years by this group of chronic patients struggling to reintegrate into the community.

Our Design

Within the 20 Year Follow-up Study we have designated four sections: (1) The Life History of Individuals, (2) The Course and Outcome Study for the Cohort, (3) The Prediction of Outcome, and (4) The Control Study (Table 2). A narrative summary of each section follows:

1. The Life History for Individuals

This section of the 20 Year Follow-up Study will focus upon the uniqueness of each individual rather than upon how he or she fits into groups of other former patients. In the course and Outcome Study, we will have dissected, scrutinized and quantified various aspects of group outcome. In this study, we propose to take that entire mass of data and reintegrate it into rich descriptive biographies. By focusing on "buffers" and strengths and not just on critical events we hope to discover some of the "invulnerabilities" (Zubin and Spring, 1977) of this group of very vulnerable former patients. We will use a modified Meyerian-Leighton Life Chart to assist in the history documentation (Leighton and Leighton, 1949).

2. The Course and Outcome Study for the Cohort

This major area of work is conceived as divided into four segments: (1) a cross-sectional appraisal at discharge from the rehabilitation program; (2) a cross-sectional assessment at current interview; (3) the interval between -- (a longitudinal scan of

Table 1
Basic Description of The Vermont Story Cohort

253 Subjects Referred from January, 1955 to April, 1959	
1. Average age at onset of mental illness.....	27 years, 9 months
Average age at onset of disability.....	30 years, 3 months
Average length of time from disability to first admission.....	1 year, 2 months
Average length of time from disability to last admission	4 years,
Average length of time from disability to in-hospital referral to rehabilitation service..	9 years, 11 months
Average length of time from in-hospital referral to discharge.....	7 months
2. Time Between First Admission and In-Hospital Referral	
Less than 1 year for	37
From 1 to 2 years for	16
From 2 to 5 years for	45
From 5 to 10 years for	72
More than 10 years for	83
	Median - 7 years, 1 month
3. Length of Last Hospitalization Preceding In-Hospital Referral	
Less than 1 year for	82
From 1 to 2 years for	40
From 2 to 5 years for	16
From 5 to 10 years for	45
More than 10 years for	71
	Median - 6 years, 10 months
4. Age Range	
Men:	16 years to 60 years
Women:	16 years to 65 years
Average Age for Men:	38 years
Average Age for Women:	42 years
5. Education	
Range:	0 to 16 years
Average:	9 years, 6 months
Median:	9 years
	High school graduates: 57
	College graduates: 15
6. Marital Status	
Married:	44
Single, separated, divorced, widowed:	209

twenty years); (4) a comparison, between Index Discharge and 20th anniversary. Particular emphasis will be placed upon the quality of life achieved by these individuals over the years. There will be an assessment of functioning in such broad areas as psychopathology and psychosocial adjustment functioning will be further delineated by a detailed look at the following variables:

Diagnosis	Living Status-Degree of Independence
Symptomatology	Psychiatric Treatment
Occupational Functioning	Use of Community Support Systems
Interpersonal Relationships	Readmissions, episodes, and intervals
Psychosexual Adjustment	Physical Health
Quality of Life	Global Assessment of Adjustments
Recreational Activities	
Self-Care Skills	
Life Events	

2a. Diagnosis and Symptomatology

Special attention will be directed toward the problem of diagnosis. We have devised the following method to describe more carefully our population and to make our study more replicable. To take advantage of the newly developed criteria in the DSM III manual and the Research Diagnostic Criteria (RDC), (Spitzer et al, 1978a; Spitzer et al, 1978b) we plan to rediagnose, retrospectively, all proband hospital record diagnoses to establish how they would be viewed by today's standards. Shifts and trends between categories will be noted. Comparisons will be made between the rediagnosis and the proband's current mental status also established by the RDC and DSM III.

Thus we will have 3 early diagnoses (DSM II, DSM III, and RDC) and 3 current diagnoses (Clinical Judgment, DSM III and RDC) each varying in degree of strictness of criteria. Prognosis, course, and outcome will be assessed across the various diagnostic classifications. A lifetime RDC diagnosis will also be generated utilizing both field and record information.

2b. Psychiatric Treatment

In addition, careful documentation will be made of the use of psychotropic drugs, major side effects and interpersonal treatments over the course of the twenty years of follow-up.

2c. Intervening Variables

Assessment will also be made of intervening variables such as life events and presence/absence of support systems in the course and outcome.

Table 2

THE VERMONT LONGITUDINAL RESEARCH PROJECT

PHASE I	The Vermont Story 1961	
PHASE II	5 Year Follow-up 1967	
PHASE III	THE CHRONIC MENTAL PATIENT IN THE COMMUNITY FOR 20 YEARS The 20 Year Follow-up Study 1980	Life Histories of Individuals Course and Outcome for Cohort <ol style="list-style-type: none">1. Cross-sectional assessment at index discharge from program2. Cross-sectional assessment 20 years after Index discharge3. Longitudinal assessment over the course of the 20 year follow-up period Prediction of Outcome prognosis correlated with actual long-term outcome The Control Study to provide normative data for comparative analysis
PHASE IV	Children-at-Risk Study	Course and Outcome/Buffers and Folklore Matched Controls from 6th Grade Cases Prospective Follow-Up

2d. Physical Health

There are 46 probands known to be deceased. They, in addition to any others documented who may also be deceased, will be followed to the time of death. Their outcome will be included until the time of their deaths. A look at suicide, accident, and natural death rates will be compared to the Vermont rates. Type of natural death will be delineated whenever possible.

3. The Prediction of Outcome Study

With twenty years of post-discharge outcomes already known from the Course and Outcome Study, we have an excellent opportunity to investigate possible early prognostic predictors of outcomes of THE VERMONT STORY cohort.

Individual prediction of outcome will be achieved by blindly rating the records which have been purged of all outcome since discharge from the program. Strauss and Carpenter's 1974 Prognostic Scale will be used. Their four item Outcome Scale (1972) will be used to give a retrospective assessment of outcome at years 5, 10, and 15 post-discharge. Their expanded nine item Levels of Functioning - Outcome Scale (Hawk et al, 1975) will be utilized in our assessment of current status of outcome at the 20th year. We will scrutinize the life histories of those probands who defy the predictions in an effort to elicit possible reasons. The group scores will enable us to test the hypotheses already generated by those earlier studies which used these scales.

Taking the outcome measures outlined in the previous section, we propose to look at the mass of acquired data to seek correlations and strength of relationships between these outcome measures and the set of predictors selected from the literature (Jablensky et al, 1979).

Utilizing multiple regression techniques, we will evolve a set of the most powerful predictors for the entire cohort. In measuring the impact of mediating events between predictor and outcome, we will look at sample trends in outcome variances using one degree of curvilinearity.

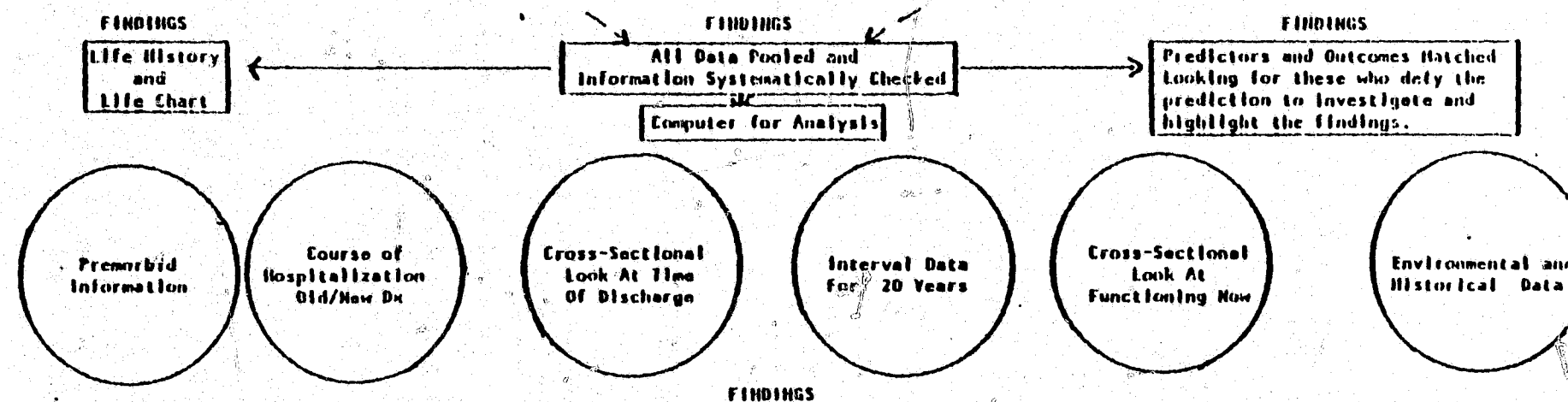
4. The Community Control Study

It is our aim to gather normative data from a community sampling on such items as work and residence histories, health, use of community support systems, interpersonal relationships, use of leisure time, and other assorted demographic variables.

This information will help us assess the outcome of our probands. A greater degree of certainty will be obtained by separating out what is cultural, environmental or normative, from what may be due to

Table 3

Overview of the Vermont Design



-655-

Diagnosis		X	X	X	X
Symptomatology	X	X	X	X	X
Readmission		X		X	X
Medication and Side Effects		X	X	X	X
Residence	X	X	X	X	X
Levels of Independence					
Self-Supporting	X		X	X	X
Basic Needs	X	X	X	X	X
Household	X		X	X	X
I.P.A.	X	X	X	X	X
Work	X	X	X	X	X
Health	X	X	X	X	X
Quality of Life	X		X	X	X
Global Assessment	X	Family System	X	X	X
Community Support System			X	X	X
Death	X	X	X	X	X

reduced post-discharge functioning or continued symptomatology. We will be able to convert our scores into more meaningful data concerning the differences between the two populations.

The community control study component will require its own project and budget at a later date. However, we have incorporated into the current study the following environmental and historical markers to place as a matrix over the course and outcome period.

- a. Catchment area items taken from:
 - 1) W.H.O. descriptors
 - a) Vermont Data Bank, which will include 1960, 1970, 1980 census.
- b. Changes in legislation or philosophy of health care delivery or social welfare systems
- c. Natural events (e.g., floods, blizzards)
- d. Economic Events (e.g., closing of the only mill in town)

We will then be able to construct community profiles for comparison against our proband's individual and group profiles. In this manner, we can make a gross measurement of success and opportunities. It is the least expensive way for us to achieve a sense of normative data of Vermonters.

The Vermont Community Questionnaire (VCQ)

The creation of an instrument to assess both current and longitudinal psychosocial course and outcome functioning for our 20 year follow-up study of the cohort is the pivotal factor in making the design operational.

On one hand, we wish to assess the usual variables of current and lifetime psychopathology, behavioral deviance, impact on families, and costs to the community in manpower, money and services. On the other hand, we also wish to focus on areas of competence, social support systems, buffers and strengths, quality of and satisfaction with life and contributions to the community and family within this group of highly vulnerable people (see Table 3).

The Vermont Community Questionnaire, referred to hereafter as the V.C.Q., is not a new instrument. In fact, it is not one instrument at all but a unique blending of several known instruments in addition to new experimental items created by us to fill the longitudinal requirements of the follow-up.

The V.C.Q. battery includes the following types of measures: (a) standard instruments used as a unified whole; (b) instruments used in their entirety but interspersed throughout the interview; (c) selected items from other instruments; and (d) experimental items on longitudinal variables

because none are available in existing instruments (see Table 4).

We have uniformly changed all cross-sectional time assessments to assessments of the past six months and we have taken the liberty to rearrange the sequences to fit our naturalistic interview design. We have had to compromise pure research for pure reality. The interspersing of items out of their original sequence may give rise to some differences in answers, but this problem is more than balanced by smooth progression, coherence, and natural rapport with the interviewee. Statistically, the ability to recapture subscales and subsets is a fairly simple matter. Some new summary methods will be used too. In this manner we will be able to share data, instrument by instrument, category by category, and/or item by item with other researchers.

There will be two one-hour interviews held in close temporal proximity to one another, but not consecutively, to eliminate fatigue factor for both rater and the proband. The first 2 segments comprised the current and longitudinal psychosocial assessment.

The major outcome areas covered by the V.C.Q. were the following:

- a. Interview I
 - i. Residence
 - ii. Work and Finances and Basic Self-Care Skills
- b. Interview II
 - i. Interpersonal Relationships
 - (I) Affectional and Psychosexual
 - (II) Children
 - (III) Family or Origin
 - (IV) Social Contacts
 - (A) Use of Leisure Time
 - (B) Use of Community Resources
 - (C) Anti-Social Behavior
 - ii. Use of Social Services, CMHC's and Treatment Received
 - iii. Global Overview

Each of these outcome areas were anchored with instrumental tasks (current, best, and 20 years); quality of life by respondent and rater; support systems; life events; and affective assessments such as degrees of independence, adaptability, function, expressed feelings, worry.

Table 4 - Internal Components of V.C.Q.

The battery consisted of the four following types:

- a. Complete standard instruments used in their entirety as is:
 - i. GADS - Schedule for Affective Disorders and Schizophrenia - by Spitzer and Endicott (3rd Edition, 1978)
 - ii. GAS - Global Assessment Scale - Endicott, et al (1975)
 - iii. Admission Medical History - Clayton and Neu (1978)
 - iv. AIMS - Abnormal Involuntary Movement Scale by NIMH's Division of Psychopharmacology
- b. Complete standard or relatively new instruments used in their entirety, but whose items are interspersed throughout the interview formats where they would logically occur in a flowing naturalistic sequence.
 - i. SAS-II, the Social Adjustment Scale for Schizophrenia Weissman, et al (1975)
 - ii. PERI-M, Psychiatric Epidemiology Research Interview as modified by Hirschfeld and Clayton (1977)
 - iii. Outcome Scales I and II by J. Strauss, et al (1972 and 1975)
- c. Standard instruments from which a variety of special items were selected (with the author's permission) and interspersed in logical sequence within the interview format.
 - i. SSIAM, Structured and Scaled Interview to Assess Maladjustment by Land, et al (1972)
 - ii. CCS, Community Care Schedule by Schwartz, et al (1977)
 - iii. PRI, Personal Resource Inventory by Clayton and Hirschfeld (1978)
 - iv. HH, the Horizon House Community Care Survey by Rutman and Baron (1975)
 - v. LIFE, Longitudinal Interval Follow-up Evaluation, VII, by Shapiro and Keller (1978)
- d. New items devised by us to complete the rigorous demands of the design and for which there were no existing instruments. These items specifically reflect the following types of areas to be assessed and are also interspersed with others.
 - i. Longitudinal functioning and worksheet
 - ii. Comparisons of now with "best" of 20 years
 - iii. Desire for change or needs to be filled
 - iv. Buffers and contributions

communication, desire for change and satisfaction.

Life Chart - We added a Meyerian-Leighton Life Chart (Leighton and Leighton, 1949) to provide a longitudinal work sheet with increased parameters, a workable method to deal with 20 years of memories, and an easier way to visualize patterns, trends, and unusual occurrences. The completion of this chart will be a joint effort between the interviewer and the proband working together.

Some Problems with Retrospective Longitudinal Studies

We have overcome some of the primary difficulties inherent in long-term follow-up studies. We have found 89% of the cohort. Our study will profit from centralized record keeping, and a very stable population in which to locate all those significant others with whom to verify information (e.g. family, friends, general practitioners, vocational rehabilitation, and aftercare workers plus clinicians who have known these probands for over two years.)

The next set of problems evolved from our efforts to make the study as relevant and replicable as possible. These difficulties include such areas as: (1) operational definition of schizophrenia; (2) what to do with early hospital DSM I and II diagnoses; (3) the question of a selection bias; (4) the problem of control for treatment; and (5) how to handle life events. In the following section, we will discuss each of our solutions.

Operational Definition of Schizophrenia

The first challenge was to redefine our cohort by today's standards. The original title of the study in 1956 was THE VERMONT STORY: Rehabilitated Chronic Schizophrenics.

The elusiveness of the definition of schizophrenia has persisted since the term was coined by Bleuler. The inclusion/exclusion criteria created narrow or broad selections from hospital samples. Not only is there lack of agreement between clinicians, but interpretations of schools of criteria are individualized by each clinician and even modified within the same clinician over time, all within the framework of cultural bias.

Our only recourse was to select a specific set of criteria which are operationally defined. We selected the Research Diagnostic Criteria and the DSM III for our assessment of psychopathology. Our rationale is the following: RDC/DSM III Systems are apt to be used in the next decades

both for clinical and research work. We do not claim those symptoms to be more valid nor do we necessarily agree with all their tenets. We need criteria which are carefully defined and, therefore, replicable from a research standpoint. We expect to lose about 12% from our schizophrenic categories reducing the number of those diagnosed with schizophrenia from 78% to 66% within the cohort.

What to do with Early Hospital DSM I and II Diagnoses?

The question of whether one should retrospectively re-diagnose patients in a study such as the current one is a complex issue. On the one hand, one must be cautious about changing a diagnostic impression based on the observations of someone who actually saw the patient at the time in question.

Even though the basis on which the diagnoses were made may not be entirely clear from the records, the diagnosticians may have been picking up important information, which was not recorded in the hospital record. On the other hand, since diagnostic practices vary from time to time, and from era to era, clinician to clinician, it is difficult to know if diagnoses are comparable, if they are not made on basis of important information, which was not recorded in the hospital record. On the other hand, since diagnostic practices vary from time to time and from era to era, clinician to clinician, it is difficult to know if diagnoses are comparable, if they are not made on the basis of clearly defined criteria.

In the current study, we have chosen to re-diagnose the patients according to clearly operationalized definitions, because we have reason to believe that the basis on which the diagnoses were made at the time of index hospitalization were considerably influenced by factors such as the availability of programs for people with specific diagnoses, and that diagnostic practices at that time were considerably different from what they are today.

Was There a Selection Bias?

Since one of the purposes of our study is to reveal what ultimately happens to patients that are discharged into the community, it is important to know to what extent these patients are representative of all patients with severe mental disorders. Further, in view of the Pilot Study results which indicate that this patient population did better over the course of 20 years than might have been expected, it is particularly important to know whether this group of patients was likely to do well.

Our cohort evolved from the pool of chronic patients, which typically existed in state hospitals during the custodial era. These "lifetime" residents performed all the menial tasks. With the advent of phenothiazines, this group became more manageable (easier to get along with) and less apprehensive about release into an alien world. The patients, who responded quickly to the new drug therapy, were discharged. Thus, members of our cohort had been on drugs for 2½ to 5 years without sufficient improvement to warrant release. A rehabilitation program to increase living and vocational skills was instituted. Staff retraining was necessary in order to replace therapeutic pessimism with optimism and higher expectations. Staff were taught that the program was something positive for patients, not just a dumping ground. Our enthusiastic anecdotes belie the fact that we are dealing with a group where selection bias, if any, leans toward those with a very poor prognosis. Therefore, our findings will reflect the conservative spectrum.

In order to quantify the degree of selection bias, we aim to assess the parametric characteristic of the entire group before our cohort were selected by a random sampling of 100 records for the population residing at Vermont State Hospital in 1956. From this effort we can ascertain how close to the mean our group of patients was at that time. This will allow us to account for bias and its amount of variance in the findings.

The Problem of the Control for Treatment

A number of studies have called attention to the importance of long-term medication for the purpose of relapse prevention in the severely mentally ill. The work of Birley and co-workers (1970) suggest that such medication is particularly important in preventing relapse in patients that go back to live in families where a high degree of negative feeling toward the patient exists.

Clearly, the current study cannot shed light on the effectiveness of medication. That is not a purpose of this study. To answer this question, a prospective double blind random assignment design would be necessary. On the other hand, because of the excellent records, relative geographic stability of the patient population, and cooperation of general practitioners and psychiatrists who have treated the patients, we will probably be able to chart out a relatively accurate picture not only of the types of drug and interpersonal treatments received by the patients, but also dosage ranges and duration of time on particular drugs. This information will allow us to make at least rough estimates of whether variations in course and outcome which seem due to other factors are, in fact, the result of variations in treatments received. Most retrospective follow-up studies have not been able to compile such treatment data, and this added dimension to the current study is an important one.

Retrospective Analysis of Life Events

In every other section of the 20 Year Follow-up Study, this cohort of patients has been grouped together to achieve cross-sectional and longitudinal assessments in a variety of areas and measures. With the same information available to us from these studies, we will also be able to spotlight the individual. The general aim of this substudy is to provide a rich, descriptive narrative account of each proband's life, with particular emphasis on critical events and "buffers" that may have occurred.

Brown and Birley continue to document life events and their inter-relationships with psychopathology. These investigators have emphasized a new awareness of many intervening variables which have tempered the inquiry. They have pointed out the difficulty in determining whether an illness was precipitated by an event or the event brought on the illness. Other variables influencing the situation concern anticipation (was the event expected?); control (how much mastery over the event did the person feel?); and desirability (was the event perceived to be desirable at the time it occurred?). These researchers question the validity of memories older than 3 months. We, too, recognize how severely limited we are in retrospective analysis of 20 plus years but we can propose the following solution.

At the very least, we will be able to provide a longitudinal scan quantifying number and kind which can then be compared to the data from a control sample. We can establish some operational criteria to rate the degree of independence of these events from the proband's illnesses. Mediating variables can be ferreted out by the use of appropriate probes from an instrument such as the PERI-M, otherwise known as the Modified Life Events Section of the Psychiatric Epidemiology Research Interview (original by the Dohrenwends and modified by Herschfeld et al, for use in Collaborative Program on the Psychobiology of Depression). These probes look at those mediating variables of anticipation, control, desirability, and impact upon self-esteem. Anecdotal descriptors of individual behavioral response to these events would be helpful in determining the usual patterns of response to stressors and unusual deviations (both adaptive and maladaptive). We acknowledge that longitudinal work in these areas is one of the most difficult efforts to defend from strict research standards but one of the most interesting to explore.

This is a unique opportunity to look at a gross overview of life. Even though retrospective analysis will miss some of the items that a solid prospective study would have picked up, we do have 20 years of living accomplished already with the outcome known.

All of the above types of data will fit into the scheme of the Life Chart as used by the Leightons in their studies of Navaho Indian Tribe and in particular, Gregorio, The Hand-Trembler (Leighton and Leighton, 1949).

Categorical Imperative

As the policy of deinstitutionalization has been vigorously pursued in this country in recent years, there has been considerable speculation about what happens to formerly chronic mental patients who are placed in the community. There is little doubt that this policy change has resulted in moving patients from one place (hospitals) to another (the community, or some other institution such as a nursing home). It is not at all clear in what sense ex-patients are better off after this move. There has been surprisingly little follow-up research of patients that have been "deinstitutionalized" into the community.

At the twentieth anniversary of this cohort's index discharge from the rehabilitation program, we are beginning to accumulate data which reveals that chronic patients may be capable of rising above the legacy of too much nature and too little nurture. Such patients demonstrate more potential for a better outcome than was heretofore envisioned. It seems imperative that we find out what kind of community career these people have pursued.

References

- Bellak, L.: *Dementia Praecox*, New York: Grune and Stratton, 1948.
- Birley, J. and Brown, G.: "Crisis and life changes preceding the onset or relapse of acute schizophrenia: Clinical aspects." *British Journal of Psychiatry*, 116: 327-333, 1970.
- Bleuler, Manfred: *Die Schizophrenen Geistesstörungen in Lichte Langjähriger Kranken - und Familiengeschichten*. Stuttgart. George Thieme Verlag (New York: Intercontinental Medical Book Corp., U.S. Distribution) 1972, pp. 673, as reviewed by Teschke, G.C., in *Schizophrenia Bulletin*. 4(#1): 48-55, 1978.
- Brooks, G.W. & Deane, W.N.: "The chronic mental patient in the community," *Diseases of the Nervous System*. 26: 85-90, February 1965.
- Brown, G.W., Bone, M., Dalison, B., et al: *Schizophrenia and Social Care*. London, Oxford University Press, 1966.
- Chittick, R.A., Brooks, G.W., Irons, F.A., and Deane, W.N.: *The Vermont Story*. Burlington, Vermont, Queen City Printers, 1961. (out of print).
- Clayton, P., and Hirschfeld, R.: *Personal Resource Inventory (P.R.I.)*. Washington University, St. Louis, 1978.
- Clayton, P. and Neu, C.: *Admission Medical History*, Washington University, St. Louis, 1978.
- Daum (now Harding) C.M., Brooks, G.W. and Albee, G.W.: "Twenty year follow-up of 253 schizophrenic patients originally selected for chronic disability: pilot study." *The Psychiatric Journal of the University of Ottawa*. II(3) pp. 129-132, September, 1977.
- Deane, W.N. and Brooks, G.W.: *Five Year Follow-up of Chronic Hospitalized Patients*. Report from Vermont State Hospital, Waterbury, Vermont, September, 1967.
- Endicott, J., Spitzer, R.L., Fleiss, J.L. & Cohen, J. "The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance" *Archives of General Psychiatry*, 33: 766-771, 1976.
- Endicott, J. & Spitzer, R.L.: "A diagnostic interview: The Schedule for Affective Disorders and Schizophrenia." Presented at the annual meeting of American Psychiatric Association, Toronto, Canada, May 1977.
- Gurland, B.J., Yorkston, N.J., Stone, A.R., et al: "The Structured and Scaled Interview to Assess Maladjustment (SSIAM): 1. Description, rationale, and development." *Archives of General Psychiatry*. 27: 259-264, 1972.
- Hawk, A.B., Carpenter, W.T., Jr., Strauss, J.S.: "Diagnostic criteria and 5 year out come in schizophrenia: a report from the International Pilot Study of Schizophrenia. *Archives of General Psychiatry*. 32:334-347, 1975.
- Hirschfeld, R., Klerman, G.L., Schless, A., Endicott, J. & Clayton, P.: (PERI-M) Modified Life Events Section of the Psychiatric Epidemiology Research Interview, NIMH, 1977.
- Hollingshead, A.B. and Redlich, F.C.: *Social Class and Mental Illness: A Community Study*. New York: J. Wiley and Sons, pp. 294-303, 1958.
- Jablensky, A., Gulbinat, W., Sartorius, N., Shapiro, R., Ernberg, G. & Deppenthaler, J.: "Predictors of course and outcome of schizophrenia and affective psychoses." in *World Health Organization's volume: Schizophrenia - An International Follow-up Study*, Chichester, John Wiley & Sons: Chapter 10, pp. 253-303, 1979.
- Jenkins, C.D., Hurst, M.W. and Rose, R.M.: "Life changes - do people remember?" *Archives of General Psychiatry*. 36: 379-384, 1979.
- Langfeldt, G.: "The prognosis of schizophrenia," *Acta Psychiatrica et Neurologica Scandinavica, Supplementum*, 110, Copenhagen, Munksgaard, 1956.
- Leighton, A.H., & D.C.: *Gregorio, The Hand-Trembler: A Psychobiological Personality Study of a Navaho Indian*, Peabody Museum Papers, Cambridge, Mass.: Harvard University Press, 1949.
- NIMH - Branch of Psychopharmacology, *Abnormal Involuntary Movement Scale (AIMS)*.
- Rupp, C. & Fletcher, E.K.: "A five to ten year follow-up study of 641 schizophrenic cases," *American Journal of Psychiatry*. 96: 877-888, 1939.
- Rutman, I.D. and Baron, R.C.: *Community Careers: an assessment of the life adjustment of former mental hospital patients*. Horizon House Institute for Research and Development, Phila., Penn., 1975.
- Schwartz, C., et al: *Community Care Schedule (CCS)* New York State Psychiatric Institute, 1977.
- Shapiro, R., & Shader, R.: "Selective review of results of previous follow-up studies of schizophrenia and other psychoses." in *World Health Organization's volume: Schizophrenia - an International Follow-up Study*, Chichester, John Wiley & Sons: Chapter 2, pp. 11-45, 1979.
- Shapiro, R. and Keller, M.: *Longitudinal Interval Follow-up Evaluation VII (LIFE)* Massachusetts Mental Health Center, 1978.

Spitzer, R.L., Endicott, J., & Robins, E.: Research Diagnostic Criteria (RDC) for a Selected Group of Functional Disorders. Third Edition. Biometrics Research, 1978a.

Spitzer, R.L., Endicott, J., & Robins, E.: "Research Diagnostic Criteria: rationale and reliability." Archives of General Psychiatry. 35: 773-782, 1978b.

Strauss, J.S. and Carpenter, W.T., Jr.: "The prediction of outcome in schizophrenia, I. characteristics of outcome." Archives of General Psychiatry. 27: 739-746, 1972.

Strauss, J.S. and Carpenter, W.T., Jr.: "The prediction of outcome in schizophrenia, II. relationships between prediction and outcome variables." Archives of General Psychiatry. 31: 37-42, 1974.

Stromgren, E.: "Recent studies of prognosis and outcome in mental disorders", in Hoch, P. and Zubin, J. (Eds.) Comparative Epidemiology of the Mental Disorders, New York and London, Grune and Stratton, 1961.

Weissman, M., et al: Social Adjustment Scale for Schizophrenia (SASII), Yale Psychiatric Institute, New Haven, 1974.

Zubin, J. & Spring, B.: "Vulnerability: A new view of schizophrenia." Journal of Abnormal Psychology. 86: 103-126, 1977.

A Long Term Follow-Up Study of Former
Adolescent Psychiatric Patients*

Wentworth Quast

A. Jack Hafner

Michael J. Shea

Joel H. Hetler

*This paper is based on a presentation made by Joel H. Hetler and Michael J. Shea at the American Psychological Association Convention, Toronto, Ontario, Canada, August 30, 1978.

This research was supported in part by the National Institute of Mental Health (MH-12038). Support was also received from the following sources of the University of Minnesota: the Graduate School, the Department of Psychiatry (Division of Child and Adolescent Psychiatry, and Psychiatry Research Unit), and the Minnesota Medical Foundation. Support was also received from the General Mills Foundation and the Dwan Family Fund.

A Long Term Follow-Up Study of Former
Adolescent Psychiatric Patients*

Wentworth Quast

A. Jack Hafner

Michael J. Shea

Joel H. Hetler

Although previous life history studies have supplied pieces of the puzzle regarding the course and outcome of adolescent psychopathology, not enough is known about the relative contribution of different adolescent variables to adult outcomes. The Adolescent Follow-Up Study described in this paper has investigated the relationships of a large number of variables to the outcomes of adolescent psychopathology. Some of the early findings of this project have been described elsewhere (Johnson, L. B., 1972; Olson, D., 1972; Shea, M. J., 1972; Zold, A., 1975).

Method

Sample

The study sample of 1108 adolescents included virtually all of the adolescent psychiatric inpatients and outpatients seen at the University of Minnesota Hospitals during a thirteen-year period from 1938 through 1950, together with a contrast group of pediatric patients individually matched for age, sex, and date of admission. The University of Minnesota Hospitals receive referrals statewide as well as from contiguous states, and the subject population for this study was demographically representative of these areas. These subjects had a mean age of 15.2 and an age range of 13 through 17. The mean time to follow-up was 24.3 years and the mean age at follow-up was 38.7 years. This meant that sufficient time had elapsed between initial contact and follow-up to reveal life patterns and to insure that the subjects had passed through the major portion of the risk period for the problems of primary interest.

CONTINUED

8 OF 10

The three most useful sources for locating subjects were telephone or letter contact with relatives, telephone directory assistance, and information from driver's license bureaus. Ninety percent of the subjects were located. Among the located subjects 73 percent were still in the upper Midwest area. Of those located, only 5 percent refused to cooperate in the study, and 10 percent had died. The percent of psychiatric versus pediatric subjects who had died was equal. However, most of the pediatric group deaths were accounted for by the disorder for which they were initially treated in adolescence. The psychiatric patients died at a much older mean age (27.4 versus 21.4 years). Deaths from suicide and homicide were few and did not differ by group.

Statistical tests were done on a number of adolescent variables to determine biases introduced through selective attrition. Comparisons were made between located and unlocated, deceased and surviving, cooperative and uncooperative, and psychiatric and pediatric subjects. Very few differences were found. The number of uncooperatives in the psychiatric group was somewhat greater than in the pediatric group. Uncooperatives tended to come more often from lower social classes. Overall, the biases introduced are viewed as minimal.

Procedure

Data collection began with the extraction of a large amount of information from the hospital charts. Next, subjects were interviewed using a 200-question, specially designed, structured interview form. Interviewers were not aware of whether subjects were from the psychiatric or pediatric group. A small percent of subjects received mailed questionnaires which had been shown to be equivalent to the structured interview in content, validity, and reliability (Boulger, 1969). In 100 cases where the subject had died or was too disturbed or too retarded to answer 100 questions, interviews were completed using relatives as informants. At the time of the interview, the Minnesota Multiphasic Personality Inventory (MMPI) was left with the subject to be completed and returned by mail. Next, collateral information was obtained from a variety of sources including schools, state hospitals, the military, law enforcement and corrections facilities. Data from the various subjective and objective sources were cross-referenced for reliability and validity.

One of the purposes of this study was to investigate the usefulness of adolescent diagnosis and to compare it with other methods of classification and prediction. Diagnoses used were those given at the time of adolescent contact. MMPIs taken at adolescence and adulthood contacts were scored using a variety of scales and signs. Presenting symptoms at the time of adolescent contact were recorded and classified using various systems including Achenbach's Internalizing versus Externalizing schema (Achenbach, 1966). A multiplicity of history material was also categorized.

Finally, a mental health rating (MHR), similar to a rating used in the Midtown Manhattan Study (Srole, Langner, Michael, Opler and Rennie, 1962), was assigned to each patient. These ratings were made by two experienced clinicians. Ratings were made for the time of adolescent contact and were done blind in that the raters had no access to information about subsequent adjustment. Similarly, a blind rating was made for the interim beginning one year after the University Hospital contact and ending five years before the date of the follow-up interview or questionnaire. A third MHR was made at adult outcome using information spanning the five years previous to and including the follow-up date. Mental health ratings were found to have a high degree of inter-rater reliability (.92) and consisted of five categories, ranging from category 1, symptom free adjustment to category 5, extremely incapacitating symptomatology requiring institutional care or primary support and protection by others. The overriding criterion for ratings was not the symptomatology itself but the degree to which it interfered with the person's daily functioning.

The analyses presented in this paper exclude all subjects diagnosed as retarded, epileptic or having organic brain syndrome. Results for these groups will be reported at a later time.

Discussion

The implications of the study presented here are that, at least for female adolescent psychiatric patients, it is possible to identify those patients with high risk of continued incapacitation. When using as predictors the presence of both severe overall disturbance at adolescence and psychotic or psychopathic personality diagnosis, the majority of errors will be in the direction of incorrectly predicting favorable outcomes.

Prediction error rate analyses for the males and for a variety of other outcome measures will be reported at a later date. It is hoped that replication of predictor-criterion relations across a variety of outcome measures within the project will help to give an estimate of shrinkage.

In summary, this study found that the majority of adolescent psychiatric patients did not continue to be notably disturbed in adult life. However, a number of former psychiatric patients did continue to have difficulty. Adolescent factors have been identified in this study which might have potential as predictors of adult outcome.

References

- Achenbach, T. M. The classification of children's psychiatric symptoms: A factor analytic study. Psychological Monographs, 1966, 80 (7, Whole No. 615).
- Bennett, I. Delinquent and Neurotic Children. Basic Books, New York, 1960.
- Boulger, J. G. A comparison of two methods of obtaining factual and subjective life history data in follow-up studies: Structured interview vs. questionnaire. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, 1969.
- Bower, E. M. & Shellhamer, T. A. School characteristics of male adolescents who later became schizophrenic. American Journal of Orthopsychiatry, 1960, 30, 712-729.
- Chess, S., Thomas, A., and Birch, H. G. Behavior problems revisited: Findings of an anterospective study. Journal of American Academy of Child Psychiatry, 1967, 6, 321-331.
- Eisenberg, L. The course of childhood schizophrenia. Archives of Neurology and Psychiatry, 1957, 78, 69-83.
- Johnson, L. B. Psychiatric diagnosis, the MMPI psychiatric symptom cluster, and severity of disturbance in adolescence as predictors of adult adjustment. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, 1972.
- Masterson, J. F., Jr. Prognosis in adolescent disorders. American Journal of Psychiatry, 1958, 114, 1097-1103.
- Masterson, J. F., Jr. The Psychiatric Dilemma of Adolescence. Little Brown, Boston. 1967a.
- Masterson, J. F., Jr. The symptomatic adolescent five years later: He didn't grow out of it. American Journal of Psychiatry, 1967b, 123, 1338-1345.
- Olson, D. The severely disturbed adolescent: Prognostic implications and long-term adjustment. Unpublished doctoral dissertation. University of Minnesota, Minneapolis, 1972.
- Ricks, D. P. & Berry, J. C. Family and symptom patterns that precede schizophrenia. In Life History Research in Psychopathology. M. Roff & D. Ricks, (Eds.), University of Minnesota Press, Minneapolis, 1970.
- Robins, L. N. Deviant Children Grow-Up. Williams & Wilkens, Baltimore, 1966.
- Robins, L. N. An actual evaluation of the causes and consequences of deviant behavior in young black men. In Life History Research in Psychopathology, Vol. 2, M. Roff, L. N. Robins, and M. Pollack, (Eds.), University of Minnesota Press, Minneapolis, 1972.
- Roff, J. Adolescent schizophrenia: Variables related to differences in long-term adult outcome. Journal of Consulting and Clinical Psychology, 1974, 42, 180-183.
- Roff, M. Childhood social interactions and young adult bad conduct. Journal of Abnormal and Social Psychology, 1961, 63, 333-337.
- Roff, M. Childhood interactions and young adult psychosis. Journal of Clinical Psychology, 1963, 19, 152-157.
- Roff, M. Some Developmental Aspects of Schizoid Personality. United States Army Medical Research and Development Command, Contract No. DA-49-007-MD-2015, Report No. 65-4, March 1965.
- Roff, M. A two-factor approach to juvenile delinquency and the later histories of juvenile delinquents. In Life History Research in Psychopathology, Vol. 2, M. Roff, L. N. Robins, and M. Pollack, (Eds.), University of Minnesota Press, Minneapolis, 1972.
- Roff, M. & Sells, S. B., Juvenile delinquency in relation to peer acceptance-rejection and socioeconomic status. Psychology in the Schools, 1968, 5, 3-18.
- Shea, M. J. A follow-up study into adulthood of adolescent psychiatric patients in relation to internalizing and externalizing symptoms. MMPI configurations, social competence, and life history variables. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, 1972.
- Shea, M. J., Hafner, A. J., Quast, W., and Hetler, J. H. Outcome of adolescent psychiatric disorders: A long-term follow-up study. In The Child in His Family: The Vulnerable Child, Vol. IV, E. J. Anthony and C. Koupernik, (Eds.). Wiley and Sons, New York, 1978.
- Srole, L., Langner, T. S., Michael, S. T., Opler, M. K., and Rennie, T. A. Mental Health in the Metropolis: The Midtown Manhattan Study, Vol. 1. McGraw Hill, New York, 1962.
- Stennett, R. G. Emotional handicap in the elementary years: Phase or disease. American Journal of Orthopsychiatry, 1966, 36, 444-449.
- Zold, A. The effects of father absence during childhood on later adjustment: A long-term follow-up. Unpublished doctoral dissertation, University of Minnesota, Minneapolis, 1975.

Long-Term Follow-Up of the Major Psychoses

Ming T. Tsuang, M.D., Ph.D.

Dr. Tsuang is Professor of Psychiatry, Department of Psychiatry and Professor of Psychiatric Epidemiology, Department of Preventive Medicine and Environmental Health, University of Iowa College of Medicine, Iowa Psychiatric Hospital, 500 Newton Road, Iowa City, Iowa 52242.

Long-Term Follow-Up of the Major Psychoses

Ming T. Tsuang, M.D., Ph.D.

A ten-year psychiatric epidemiological project is currently under way to study the long-term course and outcome of the major psychoses--schizophrenia, mania, depression and atypical schizophrenia. The period of the follow-up is 30-40 years, beginning with index hospitalization at the Iowa Psychiatric Hospital during the decade from 1934 to 1944. For purposes of comparison matched control groups of surgical patients are also being followed up using the same procedures as for the psychiatric probands. These procedures include studies of all relevant psychiatric and medical hospital records and death certificates, as well as personal interviews with all available probands and their family members using a specially designed structured psychiatric interview form, and with interviewers blind as to which group the respondent belonged. The major goals of the study are:

First, to follow our study subjects using data that are as complete and accurate as possible--by obtaining all available relevant records and by interviewing all available consenting probands and their first degree relatives, regardless of distance; and

Second, to use these data to study the long-term consequences of psychotic illness, including mortality experience; to evaluate the diagnostic validity of the major categories of these psychoses over a period of 30-40 years; and to search for homogeneous subgroups of psychotic illness based on the total fund of information gathered by our project.

We hope that the outcome of our study will be a substantial increase in the amount and quality of our knowledge about the nature and consequences of the major psychoses, as well as a refining of the diagnostic and selection criteria for use in future psychiatric treatment and research. The purpose of this presentation is to describe the scope and orientation of the Iowa project, and to discuss the project's current status and accomplishments to date.

Methods of Procedure

Sample Selection

The subjects of this follow-up study are a total of 835 psychiatric probands and 336 matched controls. In a systematic screening of excellent quality medical records at the Iowa Psychiatric Hospital, we used strict selection criteria (Feighner et al., 1972; Morrison et al., 1972) for schizophrenia, mania and depression to select 525 cases from consecutive admissions to that hospital between 1934 and 1944. There were 200 schizophrenics, 100 manics, and 225 depressives. To provide a reliable baseline for comparison and to achieve blindness, we also selected a stratified random sample of 160 non-psychiatric surgical patients admitted to the University of Iowa General Hospital between 1938 and 1948 for appendectomy or herniorrhaphy. These controls were proportionately matched to the probands on the basis of sex, age at admission, and socio-economic status. Psychiatric cases and controls were followed through all phases of data collection without knowledge of their original diagnostic statuses.

In the process of screening patients for schizophrenia, we excluded 310 who had received a chart diagnosis of schizophrenia, but who did not meet the strict selection criteria for our study. The chief reasons for exclusion were short of duration symptoms, episodic course, or the presence of affective symptoms. We decided to study these atypical schizophrenics using the same methods in order to round out our investigation of the major psychoses. To aid in this purpose, we also selected another stratified random sample of 176 surgical controls, who were matched to the atypical schizophrenics on sex, age at admission, and socio-economic status.

Data Collection

The unusually high quality of the medical records kept at the Iowa Psychiatric Hospital during the period under study is one factor making the patient population used for our project especially suitable for study. The Iowa Psychiatric Hospital is a 60-bed acute care facility serving the State of Iowa, a prosperous farm state that has a very homogeneous and stable population. Since treatment facilities were more limited 40 years ago, patients were admitted primarily for diagnostic purposes. Diagnoses were given by psychiatrists who later became quite distinguished in their field. The hospital records compiled during index admission included, in discrete packets, a complete psychiatric history, a mental status report, diagnostic formulation, reports of laboratory studies and consultations, progress and treatment notes, and a report on discharge and final disposition of the patient. The records were supplemented by complete social histories gathered by hospital social workers.

The probands were traced with information supplied by the index hospital records, and by government documents, directories, and records from other mental hospitals. In telephone interviews with the probands and their first degree relatives, we obtained information regarding current name and address, basic demographics, physical and mental health, residence status, and availability for interview. For deceased probands, and those who refused to be interviewed, we attempted to locate and interview another informant from among the proband's first degree relatives.

Field interviews initiated in 1972 and still ongoing for atypical cases, were conducted with all consenting probands and their first degree relatives using the Iowa Structured Psychiatric Interview (ISPI), a specially designed research instrument that elicits detailed information about social history, drug and alcohol history, psychiatric and physical history and symptoms, and family history. The nucleus of the ISPI is a set of twenty screening questions for schizophrenia, mania, depression and the neuroses. These screening questions assess lifetime experience of symptoms that are a characteristic of the major psychoses, and conveniently summarize the subject's history of such experiences by first episode, latest episode, and the total number of such occurrences. All field interviews were conducted by highly trained personnel, who were kept blind as to the diagnostic status of each informant, i.e., whether former psychiatric patient or surgical control. The ISPI itself has been repeatedly tested for reliability and validity, with highly satisfactory results, particularly for the screening questions for schizophrenia, mania and depression. With the information contained in a complete and accurately executed ISPI, we are sufficiently equipped to make a blind psychiatric diagnosis of the study cases.

Progress to Date

The initial focus of the project was to complete field interviews and other data collection for the 525 schizophrenic, manic and depressive probands and the 160 matched surgical controls. Once this phase was completed, we simultaneously began data analysis for this sample, and concentrated our data collection efforts on the 310 atypical schizophrenics and their 176 matched controls.

Table I shows the results of completed data collection for the schizophrenics, manics, depressives, and their matched controls. The number of subjects traced ranged from 92% of the manics to 99% of the depressives. Substantial numbers of these subjects were deceased at follow-up, especially among the depressives and manics. Of those living, we were able to interview at least 70% in each diagnostic category, including 85% of the living schizophrenics.

One of the major concerns has been to measure the effect of these different psychiatric illnesses on the long-term health and social adjustment of the probands. Longterm outcome of these patients was analyzed in terms of their marital, residential, occupational, and psychiatric status at follow-up. For each category of outcome, outcome status was rated as either Good, Fair or Poor. The criteria used for rating the outcomes are given in Table 2.

On the whole, psychiatric patients showed a significantly poorer outcome than the surgical controls. On the basis of longterm outcome, schizophrenia and affective disorders were significantly different: schizophrenia definitely showed poorer outcome than affective disorders. However, no significant differences in all four outcome variables were found between mania and depression (Tsuang et al., 1979).

In addition to these outcome variables, we compared the mortality experience of the different diagnostic groups. Using sex-age standardized mortality ratios, the mortality experiences of the study population was compared with that of the state of Iowa, the geographical area served by the Iowa Psychiatric Hospital. The data were analyzed for a four-decade period beginning 1935-44, and ending 1965-74. All three psychiatric groups (i.e., schizophrenics, manics, and depressives) had a significant increase in mortality risk. This was most pronounced in the first decade following admission, although schizophrenic patients, especially females, continued to show a significant excess of deaths throughout the entire four decades of the follow-up period. During no decade of the follow-up period did the mortality of the surgical controls differ significantly from that of the Iowa population. Details of these analyses can be found in a paper by Tsuang and Noolson, (1977).

Current Status of the Iowa Project

At present, the resources of the Iowa Project are being concentrated on data collection for the 310 atypical schizophrenics and their 176 matched controls. The same procedures for data collection are being used as for the other diagnostic groups, namely, the use of government documents, hospital records, and telephone interviews to trace the former patients; the use of death certificates for mortality analysis; and the use of the trained interviewers and the Iowa Structured Psychiatric Interview to conduct field interviews of all available and consenting subjects. As with the earlier data collection for the 525 schizophrenics, manics, and depressives, the data collection for the atypical schizophrenics has been highly successful. Table 3 shows the number of atypical schizophrenics and surgical controls who have been traced, and the number who are known to be alive or dead. The total number traced amounted to 97% of the study sample, and of these, 35% were deceased at follow-up. This high success rate for tracing the former patients compares quite favorably to the earlier phase of our project.

The next phase in our project will be to analyze the data for the atypical schizophrenics, and to compare the results with those for the other diagnostic groups.

In future analyses of the completed data for our entire sample of atypical schizophrenic probands, we will extend this outcome study by comparing the groups of major psychoses and controls of all variables collected in follow-up, including mortality data, social and family data, psychiatric and physical history, and psychiatric symptomatology at the time of index admission and at follow-up. The ultimate aim of our project is to outline homogeneous subgroups within the major psychoses, using all the types of data collected during our study. By identifying homogeneous subgroups within the major psychoses, we will then be able to specify the clinical characteristics and course of illness for the various subgroups.

Therefore, for each particular subgroup we will be able to determine which treatment would be most appropriate to use. Furthermore, we will be able to predict the short and long-term outcome for each specific group, and establish objective criteria for determining which psychiatric patients should be classified in the various subgroups. The specified criteria developed from our study, finally, will be crucial for selecting homogeneous samples in future studies of the major psychoses.

Acknowledgements

This study was supported in part by funds from the National Institute of Mental Health, United States Public Service Grants: MH 24189 and MH 31673. Dr. John C. Simpson assisted in the preparation of this manuscript. Other persons who have been professionally associated with this project are: Thomas Bray, Colette Brodersen, C. Kice Brown, Kathleen Bucher, John Clancy, William Coryell, Raymond Crowe, Richard Ebbens, Luke Falherty, Jerome A. Fleming, Cindy Jones, James Kellogg, Barbara Norton, Ralph Ramer, Lawrence Scriven, Lawrence Urban, Robert Wiese, George Winokur, and Robert F. Woolson.

References

- Coryell, W. & Tsuang, M.T. Should "Non-Feighner schizophrenia" be classified with affective disorder? Journal of Affective Disorders, 1979, 1, 3-8.
- Feighner, J.P., Robins, E. & Guze, S.B. et al. Diagnostic criteria for use in psychiatric research. Archives of General Psychiatry, 1972, 26, 57-63.
- Morrison, J., Clancy, J. & Crowe, R. et al. The Iowa 500: Diagnostic validity in mania, depression, and schizophrenia. Archives of General Psychiatry, 1972, 27, 457-461.
- Tsuang, M.T., Dempsey, G.M. & Rauscher, F. A study of "atypical schizophrenia": Comparison with schizophrenia and affective disorder by sex, age of admission, precipitant, outcome, and family history. Archives of General Psychiatry, 1976, 33, 1157-1160.
- Tsuang, M.T. & Dempsey, G.M. Longterm outcome of major psychoses: A 30-40 year field follow-up study. II. "Schizoaffective disorder compared with schizophrenia, affective disorders, and a surgical control group. Archives of General Psychiatry, 1979, 36, 1302-1304.
- Tsuang, M.T. & Woolson, R.F. Mortality in patients with schizophrenia, mania, depression and surgical conditions: A comparison with general population mortality. British Journal of Psychiatry, 1977, 130, 162-166.
- Tsuang, M.T. & Woolson, R.F. Excess mortality in schizophrenia and affective disorders: Do suicides and accidental deaths solely account for this excess? Archives of General Psychiatry, 1978, 35, 1181-1185.
- Tsuang, M.T., Woolson, R.F. & Fleming, J.A. Longterm outcome of major psychoses. I. Schizophrenia and affective disorders compared with psychiatrically symptom-free surgical conditions. Archives of General Psychiatry, 1979, 36, 1295-1301.

Table 1

Results of Data Collection for Schizophrenic, Manic, and Depressive Probands and their Matched Controls

	Diagnostic Group			
	S	M	D	C
Sample Size (N)	200	100	225	160
Mean age at Adm.	29 ± 8	34 ± 13	44 ± 12	32 ± 14
Traced	195	92	223	153
(% N)	(98)	(92)	(99)	(96)
Rated for Outcome*	186	86	212	144
(% Traced)	(95)	(91)	(95)	(95)
Deceased	78	54	162	54
(% Traced)	(40)	(59)	(73)	(35)
Living	117	38	61	99
(% Traced)	(60)	(41)	(27)	(65)
Interviewed	93	25	35	68
(% Living)	(85)	(76)	(70)	(76)

S = Schizophrenia; M = Mania; C = Control; O = Depressive

*Adequate information available to rate all four outcome variables as of the end of 1976.

Table 2

Definition of Ratings for Outcome

Status	Rating		
	Good = 3	Fair = 2	Poor = 1
Marital	Married or widowed	Divorced or separated	Single, never married
Residential	Own home or relatives' residence	Nursing or country home	Mental hospital
Occupational	Employed, retired, housewife or student	Incapacitated due to physical illness	Incapacitated due to mental illness
Psychiatric Symptoms	None	Some	Incapacitating

Table 3

Results of Data Collection for Atypical Schizophrenics and their Matched Controls

	AS	C	Total
Sample Size (N)	310	176	486
Traced	305	166	471
(% N)	(98)	(94)	(97)
Deceased	116	48	164
(% Traced)	(38)	(29)	(35)
Living	189	118	307
(% Traced)	(62)	(71)	(65)

AS = Atypical Schizophrenia; C = Control

Who's in That Dark Alley?

Simon Dinitz
Professor, Ohio State University
Associate Fellow, Academy for Contemporary Problems

John P. Conrad
American Justice Institute

Who's in That Dark Alley?

Simon Dinitz

Professor, Ohio State University

Associate Fellow, Academy for Contemporary Problems

John P. Conrad

American Justice Institute

The Columbus Study

Antecedents

About 50 years of delinquency research have been aimed at a body of belief comprising the following assumptions about violent delinquents:

- (1) A small proportion of young people are disproportionately violent.
- (2) These chronic offenders tend to be poor and black.
- (3) Fear of the serious crimes committed by the young largely arises from the deeds of these relatively few juveniles.
- (4) These few juveniles are intractable-heartless creatures who commit crimes without incurring guilt; they are not amenable to psychological or other treatment.
- (5) The earlier a child begins a criminal career, the longer that career is likely to last.
- (6) There is a progression from less to more serious crimes as the delinquent continues his or her career.

- (7) Delinquents, particularly those who are violent, tend to persist in their specialties, e.g., robbery, assault, rape.
- (8) The dispositions of violent delinquents tend to reflect a balanced judgment which takes into consideration the nature of the criminal history of the violent offender.
- (9) Whatever its other consequences, incarceration tends to slow down the criminal activity of the youthful offender.

The Study Group

A five-year cohort consisting of all the youths born in Columbus between the years 1956-60 and who had committed at least one violent crime was assembled for study. These years were chosen because the individuals born during that period would have completed their juvenile careers in 1977. The principal results of our study were based on the arrest histories of those who were born in 1956-58 (Hamparian, Schuster, Dinitz, & Conrad, 1978).

This population was drawn from the records of the Juvenile Bureau of the Columbus Police. The mechanics of this process proved to be laborious, but we are confident that omissions, if any, were few. The Columbus Police initiate a Case Record for every person at the time of the first arrest and maintain the same file for all subsequent police contacts. Our task was to identify all records containing one or more arrests for violent offenses, and to obtain a copy of each such case record. Each form contained spaces for name, address, sex, age, race, date of birth, place of birth, charges, date and time of arrest, names and addresses of parents or legal guardian, disposition and previous record.

An additional source of data was the commitment files of the Ohio Youth Commission, to which a substantial number of our youths were committed. All these data were coded, placed in our computerized file, and analyzed. A full account of our findings will be found in Hamparian, et al. (1978); for the purposes of this paper, we abstract those which are significant for our argument.

The Offenders

The 1956-60 cohort numbered 1,138 juveniles, each of whom had been arrested for at least one violent offense in Columbus before January 1, 1976. But this group had been arrested for a total of 4,499 offenses of all descriptions, most of them less serious than the arrest that qualified them for our cohort. Within the total for the five-year cohort, 811 were born in the 1956-58 period and had completed their juvenile careers in 1975. These 811 individuals had been arrested on 3,393 occasions. They constituted 1.6% of the age-eligible youths in Columbus; leaving out the girls, the cohort represented 2.3% of all the boys of the same age.

There are 205 census tracts in Columbus. Ten tracts with 7.8% of the age-eligible population contributed nearly a third of our subjects. These tracts had a population whose median income was less than 80% of the median family income in the county. They were poor and heavily black.

Over 84% of those arrested were boys, and over 54% were black, as compared with 15.4% of the age-eligible population. Only half came from families with two parents, and a third came from female-headed households. About an eighth of the sub-cohort of 811 had at least one sibling who was also part of the research study. There were 62 pairs of siblings and seven sets of three siblings--a remarkable showing, considering the age span of about 10-17. The probability of such great sibling representation is low. It suggests the conclusion that some family settings are so destructive that they produce and socialize offspring who commit delinquent acts in much the same way as other families produce college-bound children. Another possibility is that biologically defective parents produce impaired offspring. There are also other organic possibilities, such as inadequate prenatal care (poor nutrition, for example), resulting in minimally brain-damaged offspring.

The Offenders and Their Offspring

We stress that the 1956-58 cohort of 811 constitutes all the juveniles arrested for violent crimes in the age-eligible population. Comparisons with other cities are difficult to make. In the Philadelphia birth cohort studied by Wolfgang, et al. (1972), the chronic delinquents, i.e., the recidivists responsible for the majority of the serious offenses, constituted about 6% of the age-eligible birth cohort. The comparison is suggestive but hardly exact. We think it is reasonable to suppose that there is a significant difference in the rates of violent delinquency in middle-sized cities like Columbus from what is encountered in the megalopolitan cities that are the usual cities for studies such as this. Social disorganization does not reach the anomic levels of the South Bronx or North Philadelphia. About 30,000,000 people live in cities more like Columbus than they are like the much larger conurbations. There is also evidence that the problems of violent delinquency to be encountered in Columbus are generally less serious than in New York and Philadelphia, and are more typical of the hundreds of smaller cities scattered across the country (Dinitz, 1973). What we have to tell about violent delinquency is more typical of America than the frightful turmoil described by criminologists based in such cities as New York, Chicago, and Philadelphia.

With this perspective in mind, then, our cohort of 811 looks like this:

- They were arrested for 985 violent offenses (1.2 per subject), and charged with 1,087 violent offenses (1.3 per subject), during the course of their adolescent years.

- Over 83%--677 subjects--were arrested only once for a violent offense; 103 or 12.7% were arrested twice; 31, or not quite 4% had three or more arrests. At the top of the pile was one youth who was charged with six robberies, among other offenses.
- There were 270 arrests for aggravated crimes: homicide, forcible rape, armed robbery, and aggravated assault. There were 366 for "strong-arm" robbery, where no weapon was used. Finally, there were 455 arrests for molesting and assault, sometimes involving nothing more than a schoolyard fight that got sufficiently out of hand to require that the police be summoned.
- Except in the 31 repetitive cases, the violent offenses did not seem to be part of a specialized pattern of criminal involvement. Our delinquents were "generalists," taking up criminal opportunities of any kind that they came across. A great many of them were first and only offenders. The range of arrests per offender was 1-23. The mean was 4.3, the median 3.4.
- Of all those who were arrested for a violent offense, only about 30% were arrested for only that one crime. The remaining 70% had either been arrested several times for violent crimes or had a violent and nonviolent arrest pattern. Considering the life circumstances of this group of children from the economic bottom of the community, this finding has a significance that deserves more intensive scrutiny than we were able to afford with the data we collected. Sixteen percent of the cohort were not heard from after their second arrest; 10% dropped out after the third arrest, and about 11% dropped out after the fourth. The remainder, about a third, persisted and most, it must be supposed, have continued to adult criminal careers.

Conclusion

Who's in that dark alley? Americans have been too quick to accept the conclusion that a violent youth lies in wait, perhaps accompanied by a vicious gang of peers, prepared to commit unmentionable atrocities on innocent citizens. Atrocities do occur, just as they always have in our cities. It is quite probable in our largest cities, where residents live in unemployment, and poverty, the rate of juvenile violence may be much higher than is encountered in a city like Columbus--of which there are many. But whatever the true incidence and prevalence of violent crime among our youth may be, it is clear that the juvenile justice system, as it now functions, is largely irrelevant to its control and reduction.

Five major points emerge for the policymaker's consideration. First,

the juvenile crime problem is not significantly a problem of violence. Most delinquents never commit a crime of violence, and the notion that they might progress to such a serious level if not severely checked is not warranted by the data. Second, in Columbus, at least--and everywhere else, too, we suspect--the very dangerous juvenile is an exceptional phenomenon. There is no denying that he or she sometimes appears in all ghastly fascination to create a media-event. Almost always his or her mental and emotional condition can be traced to years of mutilated childhood, usually preventable with sensitive and timely intervention. When matters have gone so far that a major crime against others has been committed, special efforts must be made to restrain and treat the youth. Those special efforts should never include legislation to increase the severity of the juvenile law. Third, the considerable number of chronic offenders--five or more arrests--presents the correctional theorist with a major unsolved problem. Clearly, more incarceration is no answer; the training school experience merely confirms many delinquents in their acceptance of a criminal career as a way of life. Unless social invention can devise constructive work and activity for unemployed youth in our inner cities, it is highly unlikely that any improvement can be expected in the numbers and prospects for recidivist young offenders. There is really nothing that a correctional agency can do to prepare a young man or a young woman for life of inoffensive idleness. Fourth, our statistical methods could not justify nor rationalize the disposition of our offenders. If statistical hindsight cannot do better than we did, it is unreasonable to expect that young people committing crimes can predict what's going to happen to them if they are caught. Where the odds are so uncertain, why not take a chance? Fifth, the harsher the penalty the less street time will elapse before the next arrest.

We do not minimize the problem. Most reports of the skirmishes against juvenile delinquency come from our four largest cities--New York, Chicago, Philadelphia, and Los Angeles. The magnitude of the problems described, and the lack of progress in resolving them have led to an attitude of pessimism among both experts and laymen. Nothing can be done except to make the laws more severe while the rates of delinquency decline because of the aging of the population of the country. This confession of defeat is a new thing in American culture, unseemly and inconsistent with our tradition as an optimistic problem-solving society. Nobody needs to lurk in that alley--and it need not be dark. In the dynamic society that Americans could create, young people would be too busy to while away hours on street corners, and the alleys would be well lighted.

References

- Alexander, F., & Healy, W. Roots of crime. New York: Knopf, 1935.
- Bartollas, C., Miller, S., & Dinitz, S. Juvenile victimization: The Institutional paradox. New York: Sage-Halsted-Wiley, 1976.
- Collins, J. Chronic offender careers. Paper presented at the American Society of Criminology, Tucson, Arizona, November 4-7, 1976.
- Conrad, J.P. When the state is the teacher. National Symposium on the Serious Juvenile Offender, Academy for Contemporary Problems, Columbus, Ohio, 1977.
- Conrad, J.P., & Dinitz, S. In fear of each other. Lexington, Mass.: Lexington Books, 1977.
- Dinitz, S. Progress, crime and the folk ethic: Portrait of a small town. Criminology, 1973, 11, 3-21.
- Ferracuti, F., Dinitz, S., & Acosta de Brenes, E. Delinquents and non-delinquents in the Puerto Rico slum culture. Columbus: Ohio State University Press, 1975.
- Friedlander, K. The psychoanalytic approach to juvenile delinquency. New York: International Universities Press, 1947.
- Glueck, S., & Glueck, E. One thousand juvenile delinquents. Cambridge: Harvard University Press, 1934.
- Glueck, S., & Glueck, E. Juvenile delinquents grown up. New York: Commonwealth Fund, 1940.
- Glueck, S., & Glueck, E. Unraveling juvenile delinquency. New York: Commonwealth Fund, 1950.
- Glueck, S., & Glueck, E. Delinquents and nondelinquents in perspective. Cambridge: Harvard University Press, 1968.
- Hamparian, D., Schuster, R.L., Dinitz, S., & Conrad, J.P. The dangerous few. Lexington, Mass.: Lexington Books, 1978.
- Healy, W. The individual delinquent. Boston: Little, Brown, 1915.
- In Re Gault 376 U.S. 1, 1967.
- Platt, A. The child savers. Chicago: University of Chicago Press, 1969.
- Reckless, W., & Dinitz, S. The prevention of juvenile delinquency. Columbus: Ohio State University Press, 1972.

- West, S.J., & Farrington, D.P. Delinquent way of life. London: Heinemann, 1977.
- West, D.J., & Farrington, D.P. Who becomes delinquent. London: Heinemann, 1973.
- Wolfgang, M., Figlig, R., & Sellin, T. Delinquency in a birth cohort. Chicago: University of Chicago Press, 1972.

A Longitudinal Study of Delinquency and Dropout*

Delbert S. Elliott
Behavioral Research Institute

*The full report of the study described here is found in D. S. Elliott and H. L. Voss, Delinquency and Dropout (Lexington, MA: Lexington Books, D. C. Heath & Co, 1974). This study was supported by grants from the Center for Studies in Crime and Delinquency, NIMH (MH 07173 and MH 15285). Address all communications to Delbert S. Elliott, Director, Behavioral Research Institute, 2305 Canyon Blvd., Boulder CO 80302. The assistance of Judy Beth Berg-Hansen in the preparation of this manuscript is gratefully acknowledged.

A Longitudinal Study of Delinquency and Dropout

Delbert S. Elliott
Behavioral Research Institute

This chapter reports on a longitudinal study, "Delinquency and Dropout," which was undertaken with two basic objectives: (1) to test an integrated theoretical model which views delinquency and dropping out of school as alternative responses or adaptations to school failure; and (2) to determine if there is a causal relationship between these two behaviors.

The research is of particular interest since its theoretical perspective suggests that dropping out of school should reduce the likelihood of subsequent delinquency; yet conventional wisdom and prior cross-sectional research on this issue have led most investigators to hypothesize the opposite causal ordering.

The Sample

In this research, a type of cluster sample was used in which the basic sampling unit was a school rather than a person. Our selection of schools was not a probability sample of all schools, but a purposive selection of eight schools located in two metropolitan areas. The first area was in southern California and was comprised of three contiguous suburban communities and surrounding sections of the county. A single, unified school district served the area at the junior and senior high school levels. This unified district operated seven junior high schools (grades 7-9) and five senior high schools (grade 10-12). The second study area was located in northern California and consisted of a single suburban community served by a single four-year high school (grades 9-12). This school was also a part of a large, unified school district, but only the one school in this administrative unit was included in the study.

The subjects for this study included all students in the ninth grade in these seven junior high schools and the one high school. This sampling design was used for several reasons. First, it permitted selection of a variety of school settings and guaranteed inclusion of students with a wide range of social, economic, and racial or ethnic characteristics. A classification of milieu, similar to that developed by Reiss and Rhodes (1961), was constructed on the basis of the distribution of occupational, educational, and racial/ethnic characteristics of the population in each school's attendance area.

The selection of specific types of milieu permitted comparisons of lower- and middle-class subjects in school contexts where each was the dominant class and where each was a numerical minority. Similarly, schools were selected to include Anglos, Negroes, and Mexican-Americans in school contexts where each was dominant and in schools where each was a minority. This permitted assessment of the school milieu while controlling for class and ethnicity.

Second, the inclusion of all students in a particular grade level within each school allowed a relatively complete sociometric mapping of peer groups within each school context. Comparable data were available for each person in the network; this would not have been the case if a random sample of students from each school had been selected. These comparable data were crucially important in this research. Information on each subject's associations in school allowed direct assessment of peer influence.

Third, certain practical advantages were accrued from cluster sampling. Particularly noteworthy was the reduction in both the cost and effort required for interviewing, administering questionnaires, searching records, and obtaining official clearance for research from school administrators.

Age was an important factor involved in selection of the grade cohort to be followed. Various investigations show that the first major period of dropout occurs between the ninth and tenth grade (Dillon, 1949; U. S. Department of Labor, 1960; Van Dyke & Hoyt, 1958; Varner, 1967). Initiation of the study while the respondents were in the ninth grade thus allowed for baseline measures of our predictor variables prior to most dropout.

A slightly different approach was taken with respect to delinquency, which was conceptualized as a more continuous variable with a much earlier onset. Since it was economically impractical to obtain measures of the predictor variables prior to all of the subjects' initial involvement in delinquent behavior, our objective was to demonstrate the effect of these variables on subsequent rates of delinquent behavior, i.e., on changes in the frequency and seriousness of delinquent acts over time.

In some proportion of cases we hoped to be able to examine the onset of delinquency, but clearly this was not considered a possibility for the majority of respondents. We thus focused upon the age range in which maximum rather than initial involvement in delinquent behavior takes place. We concluded that the study should begin prior to and extend beyond the period of maximum involve-

ment in delinquency. To the extent that official delinquency statistics reflect this relationship, we expected the rates of delinquent behavior to be highest for persons between the ages of 15 and 17 (Perlman, 1964). It appeared that if data were gathered initially during the ninth grade and at regular intervals until the usual date of high school graduation, this would provide observations prior to the time the vast majority of dropouts left school, as well as during the years of maximum involvement of adolescents in delinquent behavior.

On the basis of these considerations, all students who entered these schools as ninth graders in September, 1963, comprised the study population. The potential number of respondents was 2,721. The attendance areas for the seven junior high schools were congruent with five of the six senior high schools; this greatly facilitated the maintenance of contact with the study cohort. Except for dropouts and out-of-district transfers (who were maintained in the study population), the study cohort progressed through the ninth, tenth, eleventh, and twelfth grades in these schools and remained essentially intact throughout the entire study period. The sex, ethnic, and social class distributions of the respondents by school are presented in Table 1.

The Research Design

The study was designed as a four-wave panel, with the initial data collection in September, 1963, when the cohort entered the ninth grade. The entire panel was recontacted once annually: mid tenth grade, mid eleventh grade, and late twelfth grade (1967).

Data Gathering Procedures

In each of these annual data collection waves, subjects in the cohort completed standardized questionnaires in groups of 20-30 students. Each group of subjects was given oral instructions, and questions were clarified for individuals as they arose. Each questionnaire was checked for completeness before it was accepted. For subjects out of school, a trained interviewer contacted them individually with an alternate form of the questionnaire designed to be completed with the interviewer's guidance. In this questionnaire, inquiries about current school activities were deleted; items pertaining to employment, marriage, and other out-of-school experiences were substituted.

In addition to the annual questionnaire, an interview with one of each subject's parents (in most instances the subject's mother or mother surrogate) was completed early in 1964. These interviews were conducted in the subject's home. Each dropout was also interviewed as soon as it was determined that he or she had left school, usually within a month of the date of withdrawal. This interview dealt specifically with the circumstances surrounding the withdrawal or dismissal from school and the subject's immediate plans for work or further education.

Table 1

Sex, Ethnicity, and Social Class Distributions in Study Schools
(Percentages)

Junior High School	N ^a	Sex		Ethnicity				Class				
		Male	Female	Anglo	Negro	Mexican- American	Other	1	2	3	4	5
A	(302)	50	50	81	-	11	8	0	5	31	46	18
B	(455)	49	51	93	-	6	1	7	23	37	24	9
C	(212)	49	51	61	1	32	6	2	1	19	44	34
D	(376)	51	49	88	1	7	4	1	8	38	41	12
E	(414)	52	48	73	1	22	4	-	4	30	48	18
F	(342)	49	51	86	-	9	5	1	9	23	53	14
G	(173)	56	44	53	-	43	4	2	-	24	42	32
Senior High School												
A	(284)	46	54	73	1	21	5	-	3	29	50	18
B	(381)	52	48	87	-	10	3	0	8	38	42	12
C	(481)	51	49	94	-	4	2	8	21	41	21	9
D	(366)	52	48	75	-	21	4	1	6	25	50	18
E	(343)	50	50	30	57	3	10	3	4	22	39	32
F	(419)	48	52	72	0	21	7	1	4	28	44	23

^aWith the exception of High School E, the four-year high, the high school N's reflect enrollment in September, 1964, when the cohort entered the tenth grade. Dropouts and out-of-district transfers are not included. For High School E and the junior high schools, these figures reflect enrollment at the beginning of the study.

During 1964, we also obtained teacher evaluations of all subjects with respect to their (1) academic or classroom performance, (2) school adjustment, (3) hostility toward teachers, (4) reading ability, (5) likelihood of dropout, and (6) probability of contact with the police. Finally, school records (IQ, achievement test scores, grades, absenteeism, and disciplinary actions) and police records were also obtained for each subject for the study period.

Cohort Attrition

A fundamental problem in longitudinal studies involves the loss of subjects over the course of the investigation. The seriousness of this problem is related to the rate of attrition and selectivity in the loss of cases and constitutes a crucial problem for studies in which attrition is correlated with the variables being measured.

Ninety-six percent of the 2,724 eligible students participated in the study. Initial losses included: five students who refused to participate; 58 students we were unable to make contact with due to their repeated absence from school; and 41 students whose parents refused to allow their participation. The study thus involved 2,617 youths and their parents, each of whom completed questionnaires/interviews in 1963.

Over the course of the study, 19 subjects -- 10 males and 9 females -- refused to continue their participation. These were distributed evenly: six in the second year, seven in the third year, and six in the fourth year. Eleven of these subjects eventually graduated, whereas eight either dropped out prior to or subsequent to refusing to provide information. Clearly, a sizable proportion of the refusals (42 percent) were dropouts.

To minimize case attrition due to residential mobility, rather elaborate tracking techniques were developed. This was essential because 23 percent of the students left the schools in the two study areas. The procedures used to locate these subjects included the usual checks through post office forwarding addresses, public utility addresses, and requests for school transcripts. Using a card similar to those employed by the schools to obtain names of persons to contact in case of emergency, each year we asked subjects to provide the names and addresses of persons who would always know their whereabouts. In most instances, one or more of these techniques proved to be satisfactory in locating mobile subjects. Because most mobile subjects entered a school in their new community, efforts to follow them were greatly facilitated. In many instances, however, we had to rely on the ingenuity of our interviewers in obtaining leads on subjects temporarily lost.

At the conclusion of the study, there were only eight subjects whose whereabouts were completely unknown. There was no evidence of selectivity by sex, ethnicity, or initial involvement in delinquency on the part of these subjects; however, all eight were from the lower half of the class distribution, all were dropouts, and all but one came from a school milieu that provided limited opportunity. When the project was terminated, there were 235 subjects.

including the 19 refusals, for whom we had incomplete data for the last data-gathering phase, even though we had addresses for them and knew their graduate status. Limitations of time did not permit further efforts to secure complete information from these subjects.

Persons for whom information was missing appeared to be similar to the study cohort with respect to ethnicity and class. However, they were predominantly male (62 percent) and reported significantly higher initial involvement in delinquent activities than the cohort in general. Nevertheless, an attrition rate of 9 percent is quite low, given the length of time the subjects were followed. Further, the effect of incomplete data for 235 subjects was not as serious as would have been the case in a study of shorter duration, because data gathered during the first three years were available for most of these subjects. Consequently, trends for major variables could be established.

Measures

Failure and Anticipated Failure

We hypothesized that the causal sequence leading to delinquency and dropout begins with failure or anticipated failure to achieve culturally valued goals. We identified three distinct sets of goals that might serve as instigators for delinquency and dropout in the context of the community, school, and home. In each annual questionnaire a series of inquiries related to success were made, and we developed success or failure scales for each social context.

The community success/failure scale reflected the individual's assessment of his chances to achieve his educational and occupational aspirations. Success in the home was measured by a two-item scale pertaining to parental acceptance.

Three measures of success in school were developed: (1) academic success, (2) activity success, and (3) peer success. The first scale, a composite of achievement test scores, grade point averages, and teachers' evaluations of academic performance, reflected academic success or achievement in the formal system of the school. Teachers' evaluations were obtained only in the first year. Although comparable data were not available for later years, the teachers' evaluations were included in the first-year scale because these judgments were relevant to academic success, and this scale served as the initial predictor.

According to the values of the youth culture, involvement in school organizations and activities is an important criterion of success in the informal system of the school. Gordon (1957) says that achievement in activities is the most important determinant of a student's status in school. The activity scale was based on two measures developed by Coleman (1961) which assess the discrepancy between a student's desired and actual level of participation in school activities. The third measure of success in school reflected status deprivation or informal distinctions in relation to peers. The scale was based on the questions developed by Reiss and Rhodes (1963) to measure status deprivation.

The Explanation of Failure

We suggested that the explanation of failure is an important conditional variable in the relationship between failure and delinquency. Specifically, the relationship was thought to be contingent upon external, rather than internal, attribution of blame. The approach employed in this research was similar to Rosenzweig's (1938, 1963). Cartoons depicting frustrating situations were presented to respondents.

Responses to these situations were scored according to the direction of aggression by three independent judges. The Coefficient of Agreement between judges' ratings was .79, an indication of a fairly high degree of consensus (Robinson, 1957, 1959). A composite score was derived for all situations to classify respondents either as persons who attribute blame externally (extrapunitive), attribute blame to themselves (intropunitive), or avoid imputation of blame (impunitive).

We proposed that normlessness and social isolation should be related to specific contexts or institutional settings. Institutions are viewed as normative systems organized around a particular type of human activity. Normlessness or social isolation with reference to one context need not imply similar feelings with reference to other contexts. For example, the dropout may not be alienated from the home and community as much as from the school. There may be different reactions to the normative structure in diverse settings. We developed normlessness and social isolation scales for the community, school, and home from a pool of approximately 90 items. The normlessness scales focused upon normative consensus, whereas the social isolation scales centered upon withdrawal or separation from groups and activities in each social context.

Patterns of Differential Association

Presumably differential exposure to delinquency and dropout may occur in several contexts, and the extent to which specific patterns of exposure are conducive to delinquent behavior and dropout may vary. We argued that unless individuals have access to social environments where they can learn appropriate definitions of delinquency or dropout as well as receive positive social reinforcement for these behaviors, they are not likely to engage extensively in delinquent behavior or drop out of school. Rather than limiting our attention to the influence of highly structured gangs, we were concerned with individual's exposure to everyday instances of delinquent behavior and dropout in their environment and to the attitudes, values, and social definitions supportive of such behavior. We accepted Sutherland's (1947) premise that the most significant influences are found in primary groups, and our measures of exposure focused upon these groups as found in the school and home.

(1) Measures of Exposure to Delinquency

Differential exposure to delinquency in the community was measured by a single item designed to tap the individual's subjective impression about the amount of delinquent activity in the community. This item did not measure primary group associations with delinquent youth. It was assumed that subjects who perceive a great deal of delinquency in the community have more exposure to delinquent influences. This question was asked only in the first and fourth years.

Exposure to delinquency in the school was measured in terms of the subject's perception of delinquency among immediate friends and with an objective measure of delinquent behavior on the part of the persons listed as friends in a sociometric question. The first measure, which we referred to as differential association in the school, was based on three items employed by Short (1957, 1958, 1960) and Voss (1964) to tap association with delinquent friends. These items were asked only in the first year; consequently, we did not have gain measures for this scale.

The second measure of exposure was based on the self-report of involvement and officially recorded delinquency of the subject's friends. To differentiate these measures, this one was referred to as sociometric exposure. The proportion of the subject's friends who either reported serious offenses on the self-reported delinquency (SRD) measure or had an official delinquency record constituted the sociometric exposure score for each year.

A measure of exposure to criminal influences in the home was obtained for a representative sample of the study population. The names of siblings and parents were compiled for this sample of 296 respondents. These names were checked against the juvenile and adult records of local police departments as well as appropriate municipal and superior courts. The record check extended to January 1, 1959, when 9 was the modal age of the subjects.

The measure of exposure to criminality in the family consisted of three indicators: official offense rates, the maximum seriousness of any offense, and the type of legal action involved. The first score was simply the total number of criminal or delinquent acts committed by family members, excluding the subject. Scores for the seriousness of the offense and legal action were based upon the single most serious recorded offense for the family. Each of the three indicators was standardized and then summed to provide a single score for family exposure. This measure was obtained only for Period I.

(2) Association with Dropouts

Three measures of exposure to dropout were developed -- two for the school and one for the home. The former were similar to the measures of exposure to delinquency; they reflected the extent of contact with dropout peers. The second measure involved the sociometric question described earlier with reference to delinquency among the subject's friends. For each respondent,

the proportion of friends who dropped out in the following study period constituted the sociometric measure of exposure to dropout. A scale measuring exposure to dropout influences in the home was constructed for Periods I and IV. It reflected the level of parental support for dropout.

(3) Commitment to Peers and Parents

In an effort to tap the dimension of differential association Sutherland (1947) referred to as intensity, we developed two measures of attachment. Commitment to peers was measured by a single item that followed the sociometric question and referred directly to the respondent's list of friends. The item asked, If you found that this group of friends was leading you into trouble, would you still run around with them? The commitment to parents scale consisted of three items: (1) Who do you confide in when you get into trouble? (2) When making a decision about whether a certain action is right or wrong, whose opinion are you most likely to accept? (3) Which one of these things would be hardest for you to take -- your parent's disapproval, your teacher's disapproval, or your friend's disapproval?

Self-Reported Delinquent Behavior

Although the self-report approach raises certain methodological issues, particularly in the areas of reliability and validity, it appeared to be more appropriate for this research than exclusive reliance on official records or the use of other techniques, such as direct observation. Consequently, a self-report instrument was developed to provide a direct measure of delinquent behavior. This instrument, a modification of the Nye-Short delinquency check list, consisted of 12 of Nye and Short's original 21 items (1957). The general format of the questions was similar to that used by Nye and Short. The items pertained to auto theft, petty theft (worth less than \$2), thefts of medium value (\$2 to \$50), gang fights, grand theft (over \$50), robbery, purchasing liquor, destroying property, truancy, and running away from home. Two additional items from the Nye-Short check list -- driving a car without a license and defying parental authority -- were included in the questionnaire but were not included in the measure of self-reported delinquency (SRD).

The SRD items were included in the questionnaires administered to the study population during the first and fourth data-collection waves. Like the question format, the instructions were similar to those provided by Nye and Short. The instructions in the first wave indicated that the time period covered by the self-report questions included the three years of junior high school (grades 7-9), whereas the time period covered in the instructions in the fourth period included the three senior high school years (grades 10-12). With the exception of an additional item concerning drug use, the instrument used in the fourth period was identical to the one employed initially.

Two essential dimensions of delinquent behavior were measured by these SRD items. The basic analytic procedure involved calculation of the frequency

of self-reported offenses by a serious and nonserious classification and by type of offense. To score the frequency of delinquent acts we transformed the response categories of the SRD items into frequencies, as follows: "no" or "none" equals 0; "once or twice" equals 1; "several times" or "three times" equals 3; and "very often" equals 4. This scoring procedure provided a conservative estimate of the actual frequency of reported delinquent acts.

We explored various approaches to the classification of offenses with respect to seriousness (Sellin & Wolfgang, 1964; Robin, 1964; McEachern & Bauzer, 1967; Erickson, 1966; California Penal Code, 1963) and found that the different measures produced somewhat divergent orderings. We resolved the seriousness issue by adopting the distinction between felonies and misdemeanors as specified in the California Penal Code (1963).

Official Delinquent Behavior

We examined school records to obtain I. Q. scores, achievement test scores, grades, and information on absenteeism and disciplinary action for each respondent. The records of the local police and sheriffs, juvenile divisions of the county probation departments, and juvenile courts were also checked to determine which individuals in the cohort had official contact with law enforcement agencies. Information was recorded on the nature of the offense, the date of the offense, and the disposition of the case by the police. The probation department records included information on the filing of petitions and court action for cases referred to the probation department by the local police or sheriff.

In view of the mobility of the population, it was also necessary to have records examined in a variety of other locations. To obtain comparable information about respondents who left the two primary study areas, letters were sent to the police and probation departments in every location in which a subject had resided, according to our records, during the course of the study. Each school a subject entered was also contacted to obtain school records. Response to these requests was excellent; 92 percent of the record checks were completed satisfactorily.

The official measure incorporated the same dimensions of behavior reflected in the SRD measure -- the frequency and seriousness of delinquent acts and the type of offense. Not all police contacts involved delinquent acts, as defined in this research, and only offenses involving a violation of the California Penal Code were included in the official measure of delinquent behavior. In addition, all traffic violations were excluded. The criterion of seriousness was again the distinction between felonies and misdemeanors.

To estimate rates of delinquent behavior, it is desirable to have both self-reported and official data because these sources are complementary. Together they provide a more complete picture of the frequency and nature of delinquent behavior. In addition, the availability of both types of data permits exploration of the relationship between these divergent sources.

Reliability and Homogeneity

For each of the above measures which involve scales, measures of reliability and homogeneity were calculated for each year and are presented in the report (Elliott & Voss, 1974). Alpha (Cronbach, 1951) was the measure of reliability and Homogeneity Ratio (Scott, 1960) was the measure of homogeneity employed.

Summary of Major Findings

In this investigation a dropout was defined as a person who left school for a period of at least 30 consecutive days for a reason other than death. There were 2 deaths in the cohort during the study, and for convenience they were classified as involuntary dropouts. In terms of this operational definition, 558 (21.3 percent) persons were classified as dropouts during the course of the study; these respondents left school 674 times. This definition adequately conveyed the extent to which students interrupt their high school education, but did not indicate their final educational status. At the end of the study, 79 percent of the cohort were graduates, 19 percent were nongraduates, and 2 percent were still enrolled in school. It is interesting to note that 10 percent of those who dropped out of school at some point eventually graduated.

Three types of dropouts -- involuntary, educationally handicapped, and capable -- were identified. Only 2 percent of the dropouts left school involuntarily; 32 percent were classified as educationally handicapped, and 66 percent were judged capable of completing high school. Within the category of capable dropouts, approximately one-fifth were forced to leave school by administrative order; the remainder left voluntarily. Two implications of these findings deserve mention. First, few students, at least in the setting in which our research was conducted, left school because of financial pressure or serious illness. Second, dropout among intellectually capable youths was not always voluntary. While troublesome behavior undoubtedly triggered the administrative response, a sizable proportion of the capable dropouts were forced to leave school, i.e., were "pushouts."

The study population reported a considerable volume of delinquent behavior on the ten item SRD measure. For the three years of junior high school, the 2,617 subjects reported an estimated 10,073 delinquent acts, whereas their responses in senior high school reflected involvement in 13,141 offenses. Over the six-year period, the study population reported 23,214 delinquent acts or an average of 9.63 infractions for each subject. The males reported almost twice as many offenses as the females -- 12.24 in comparison with 6.90 for females.

For the study population, 1,486 police contacts were recorded. The self-report and official data permitted estimation of the amount of hidden delinquency, the probability of police contact according to the frequency of delinquent behavior, and conditional probabilities of police contact in terms of the sex, ethnicity, and social class of the offender. These analyses revealed that there were approximately 5 police contacts for every 100 self-reported offenses, indicating that there was a substantial amount of hidden delinquency in this population. Further, police contact rates varied

by sex, ethnicity, and social class; males, members of minority groups, and lower-class juveniles had a relatively greater risk of police contact for each delinquent act they reported on the SRD.

Over the total study period, dropouts, both male and female, had four times as many police contacts as graduates. Dropouts also reported more delinquent behavior on the SRD measure. Nongraduates, who were in school at the end of the study but did not graduate, were even more highly involved in officially recorded delinquency. Involuntary dropouts reported about the same number of delinquent acts as graduates. In comparison, the educationally handicapped and capable dropouts reported more delinquent behavior than graduates.

Analyzing the effect of dropout on delinquency, we found that dropouts consistently had higher police contact rates than graduates while they were in school, and these rates increased with time while they were in school. However, the police contact rate systematically declined in the period after which dropout occurred, and it continued to decline over time. Out of school rates for dropouts were substantially lower than the rates for graduates. Use of the SRD measure produced similar findings. The dropouts reached a high level of involvement in delinquency prior to leaving school, but once they were out of school, their involvement in serious offenses declined substantially, and the total number of offenses reported declined slightly.

In comparison, graduates reported an increasing number of delinquent acts, nonserious as well as serious offenses, with time. Whether delinquency was measured in terms of police contacts or self-reports, similar patterns were revealed. The relationship between delinquency and dropout could not be explained by class or sex differences among dropouts and graduates, by differential visibility, or by the deterrent effect of adjudication.

In general, delinquency increased the probability of dropout, which in turn decreased the probability of further delinquency. This causal sequence was supported with measures of delinquency based on police contacts and self-reports. In each study year the probability of future dropout was more than two times higher for those who experienced police contact than for subjects with no police contact. Extensive involvement in delinquent behavior, whether or not it led to official action, increased the probability of dropout.

The dropouts' out-of-school experiences with respect to marriage and employment were also related to their involvement in delinquency. With marriage and employment, the dropout made the transition from adolescence into conventional adult roles. It was this factor, we believe, that accounted for the general decline in the dropouts' rates of delinquent behavior once out of school. This interpretation is consistent with the finding that marriage was a more significant deterrent for delinquency than employment, as marriage is a less ambiguous indicator of adult status than employment, particularly when sporadic and part-time employment to some extent characterize student roles.

The data generally supported our hypothesis that the school is the critical social context for the generation of delinquent behavior. For males, the most powerful predictors of delinquency were limited academic achievement, school normlessness, association with delinquent classmates, and a heavy commitment to peers. For females, the best predictors were parental rejection, school normlessness, association with delinquent classmates, and a heavy commitment to peers. However, academic failure as well as normlessness and social isolation in the home were also predictive of female delinquency. Clearly, the home was a more important setting for females than for males -- failure, alienation, and exposure in the school and home were conducive to female delinquency. In contrast, none of the measures for the home context were strongly predictive of male delinquency.

The strongest predictors of dropout were academic failure, school normlessness and social isolation, exposure to dropout in the home, and commitment to peers; there were no significant sex differences. The fact that exposure to dropout in the home and commitment to school peers were both predictive of dropout was not inconsistent, as association with juveniles who dropped out of school was also predictive of dropout. The data did not support the hypothesis that dropout is precipitated by problems in the home. Rather, the major instigating forces in dropout were to be found in academic failure and alienation from the school. Exposure to dropout, whether it occurred in the school or home, was generally conducive to dropout.

Finally, our data provided an explanation for the fact that dropout was related to class while self-reported delinquency was not. A number of our findings, taken together, made it possible to account for the class differential in dropout. Social class was unrelated to school normlessness but was associated with social isolation in school. It was also the case that exposure to delinquent friends was only weakly associated with class, but exposure to dropout in the home was strongly related to class. Lower-class youths were more likely to experience isolation in school, and they had more exposure to dropout influences in the home. Both of these conditions were conducive to dropout. On the other hand, the conditions most conducive to delinquency were either unrelated or weakly related to class. We would, therefore, not expect a class differential in delinquent behavior.

Our findings challenged the conventional stereotypes of delinquents and dropouts. First, we found little support for the view that dropout is motivated by difficulties at home. At the point of leaving school, dropouts were no more likely than graduates to experience rejection or alienation from their parents. In most instances, the immediate stimulus for voluntary dropout was some difficulty or crisis in school. Typically, this involved conflict with a teacher resulting in a temporary suspension from which the student never returned. Since approximately one-fifth of the dropouts were pushed out of school, the phenomenon of dropout could not be viewed strictly in terms of personal decisions on the part of youths to terminate their education.

Although a substantial number of dropouts were in fact pushouts, formal expulsion was rare. Pushout resulted from the enforcement of rules prohibiting pregnant girls, married students, and troublesome 18-year old students from attending the regular day-school program. Few students dropped out of school because they were attracted to work or had to work to help support their families. Thus, there was little evidence that dropouts rejected the benefits of a high school diploma.

Second, although dropouts had limited involvement in extracurricular activities, this was not predictive of dropout. Third, dropout tended to be a group adaptation to school problems. In several instances whole groups dropped out together. Association with dropout peers and limited support for education in the home were both causally related to dropout.

Finally, with respect to delinquency, there was reason to question whether dropouts necessarily suffered all of the negative consequences usually ascribed to dropout. There was considerable evidence that the dropout's involvement in delinquency began while the subject was in school and was primarily a response to experiences in the school. Departure from school reduced the dropouts' rate of delinquent behavior and the likelihood of official police contact.

Our data challenged similar assumptions about the causes of delinquency. First, delinquent behavior did not appear to be related to social class or ethnicity. Status deprivation -- the perception that others have nicer clothes and homes -- was related to social class and ethnicity but was not predictive of delinquency. On the other hand, the rate of police contact was related to social class, ethnicity, and perceptions of status deprivation.

Second, the level of participation in the school's extracurricular activities was not predictive of delinquency. In fact, youths who were highly involved in school activities reported rates of delinquency slightly higher than those with limited involvement. Delinquent behavior was not restricted to those labeled "hoods" or "outsiders" by the peer culture, but appeared to be a more general feature of adolescent culture. Third, delinquency among males was less of a response to rejection and alienation from the home than female delinquency. While the school was the most critical social context for both sexes, it was particularly influential with respect to male delinquency.

Finally, our analysis suggested that the relationships between failure, normlessness, association with delinquent peers, and delinquent behavior were mutually reinforcing. Failure, normlessness, and association with delinquent friends were both causes and consequences of involvement in delinquent behavior. Delinquency increased the likelihood that youths would do poorly in school and perceive themselves as rejected by their parents. Involvement in delinquent behavior had a particularly strong influence on feelings of normlessness in school as well as friendship choices.

In view of these findings, current attempts to raise the age of compulsory attendance as a means of keeping youths in school might well be reconsidered. This policy is based on the assumption that the school functions as a positive form of social control and is conducive to better life adjustment. Our findings are at odds with this assumption and imply that the school often aggravates a youth's problems rather than alleviating them.

It is not a coincidence that the rate of delinquency is inversely related to the rate of dropout. As the holding power of our schools has increased, so has the rate of delinquency. Compulsory school attendance facilitates delinquency by forcing youths to remain in what is sometimes a frustrating situation in which they are stigmatized as failures. In the final analysis, escape either through dropout or graduation appears to be the only satisfactory resolution of this problem. For the dropout and the graduate, rates of delinquency decline upon leaving the compulsory school setting. Delinquency, on the other hand, is not an adequate solution, but serves only to set into motion reciprocal processes of rejection and alienation and thereby increases the probabilities of failure.

It does not necessarily follow from these observations that all students who are alienated and frustrated by their experiences in school or who perceive the school as meaningless should be encouraged to drop out of high school. However, in some cases this may be an appropriate course of action and one which should not be restricted by law. A more important strategy would be to change the structure of the school -- to explore new types of learning environments in which competition is minimized and in which failure ceases to be a functional prerequisite of the educational system.

References

- California Penal Code. The penal code of the state of California. Los Angeles: Legal Bookstore, 1963.
- Cloward, R. A. & Ohlin, L. E. Delinquency and opportunity. Glencoe IL: Free Press, 1960.
- Coleman, J. S. The adolescent society. New York: Free Press, 1961.
- Cronbach, L. J. Coefficient alpha and the internal structure of tests. Psychometrika, 1951, 16, 297-334.
- Dillon, H. J. Early school leavers: A major educational problem. Publication No. 401 (October). New York: National Child Labor Committee, 1949.
- Elliott, D. S. & Voss, H. L. Delinquency and dropout. A summary report to the National Institute of Mental Health. Grant Numbers MH-170173 and RO1 15285 (November). Boulder: University of Colorado Institute of Behavioral Science, 1971.
- _____. Delinquency and dropout. Lexington MA: Lexington Books, D. C. Heath & Co., 1974.
- Elliott, D. S. & Angeton, S. S. Reconciling race and class differences in self-reported and official estimates of delinquency. American Sociological Review, 1980, 45, 95-110.
- Erickson, G. E. Bias in the processing of juvenile delinquents. Unpublished Master's thesis. San Diego: San Diego State College, 1966.
- Gordon, C. W. The social system of the high school. Glencoe IL: Free Press, 1957.
- Hardt, R. H. & Bodine, G. E. Development of self-report instruments in delinquency research. Syracuse: Syracuse University Youth Development Center, 1965.
- Hardt, R. H. & Peterson-Hardt, S. On determining the quality of the delinquency self-report method. Journal of Research in Crime and Delinquency, 1977, 14, 247-261.
- Hindelang, M. J., Hirschi, T. & Weis, J. G. Self-reported delinquency: Methods and substance. Proposal submitted to the National Institute of Mental Health, Department of Health, Education and Welfare, MH 27778, 1975.
- Krane, S. School dropout: A response to aspiration opportunity discrepancy? Unpublished monograph. Boulder: University of Colorado, 1976.

- McEachern, A. W. & Bauzer, R. Factors related to disposition in juvenile police contacts. In M. W. Klein (Ed.). Juvenile gangs in context. Englewood Cliffs NJ: Prentice-Hall, 1967.
- Nettler, G. Explaining crime. New York: McGraw-Hill, 1974.
- Nye, F. I. & Short, J. F., Jr. Scaling delinquent behavior. American Sociological Review, 1957, 22, 326-331.
- Perlman, I. R. Antisocial behavior of the minor in the United States. Federal Probation, 1964, 28, 23-30.
- Reiss, A. J., Jr. & Rhodes, A. L. The distribution of juvenile delinquency in the social class structure. American Sociological Review, 1961, 26, 720-732.
- _____. Status deprivation and delinquent behavior. Sociological Quarterly, 1963, 4, 135-149.
- Robin, G. D. Gang member delinquency: Its extent, sequence and typology. Journal of Criminal Law, Criminology and Police Science, 1964, 55, 59-69.
- Robinson, W. S. The statistical measurement of agreement. American Sociological Review, 1959, 24, 338-345.
- Rosenzweig, S. The experimental measurement of types of reaction to frustration. In H. A. Murray (Ed.). Explorations in personality. New York: Oxford University Press, 1938.
- _____. Validity of the Rosenzweig picture-frustration study with felons and delinquents. Journal of Consulting Psychology, 1963, 27, 535-536.
- Scott, W. A. Measures of test homogeneity. Educational and Psychological Measurement, 1960, 20, 751-757.
- Sellin, T. & Wolfgang, M. E. The measurement of delinquency. New York: Wiley, 1964.
- Severy, L. J. Exposure to antisocial behavior. Unpublished Ph. D. dissertation. Boulder: University of Colorado, 1970.
- Short, J. F., Jr. Differential association and delinquency. Social Problems, 1957, 4, 233-239.
- _____. Differential association with delinquent friends and delinquent behavior. Pacific Sociological Review, 1958, 1, 20-25.
- _____. Differential associations as a hypothesis: Problems of empirical testing. Social Problems, 1960, 8, 14-25.

Sutherland, L. H. Principles of criminology, fourth edition. Philadelphia: Lippincott, 1947.

U. S. Department of Labor. School and early employment experience of youth: A report on seven communities, 1952-1957. Bulletin No. 1277. Washington DC: U. S. Government Printing Office, 1960.

Van Dyke, L. A. & Hoyt, K. B. The drop-out problem in Iowa high schools. Des Moines: Iowa State Department of Public Instruction, 1958.

Varner, S. E. School dropouts. Washington DC: National Education Association, 1967.

Voss, H. L. Differential association and reported delinquent behavior: A replication. Social Problems, 1964, 12, 78-85.

From Boy To Man - From Delinquency To Crime

Marvin E. Wolfgang
Center for Studies in Criminology
and Criminal Law
University of Pennsylvania

From Boy To Man - From Delinquency To Crime

Marvin E. Wolfgang

Center for Studies in Criminology

and Criminal Law

University of Pennsylvania

Purpose and Background

The purpose of this paper is to examine the relationship between juvenile and adult offense probabilities, offense types and offense seriousness. Although the probability statements may sound predictive, I am not suggesting a juvenile-to-adult predictive model to be used by criminal justice. The lucid and comprehensive summary of prediction studies in criminology by Monahan (1977) stands firm in its conclusions beside any data I present here. I therefore wish at the outset to caution against unwarranted prediction inferences being made from the findings I report. On the other hand, there are some strong assertions, supported by statistical analysis, to be made about adult offensivity and adult assaults based on juvenile offensivity and juvenile assaults. The degree of boldness of the claims is a function of the rigor of the data and the robustness of the methodology, not the subjective leaps beyond the confines of the data.

The material presented here is derived from the birth cohort study conducted at the Center for Studies in Criminology and Criminal Law at the University of Pennsylvania. The first display of this work was published as Delinquency in a Birth Cohort in 1972 (Wolfgang, Figlio & Sellin). That study involved analysis of a cohort of nearly 10,000 males born in 1945 who lived in Philadelphia at least from their tenth to their eighteenth birthdays. Through the use of school, police and Selective Service files, we were able to locate and gather data on 9945 boys. Since 1968 we have followed a ten per cent random sample of the original cohort. The sample drawn consisted of 975 subjects who were representative of white and nonwhite delinquents and nondelinquents. After three years of diligent searching for the sample subjects, many could not be found. The process resulted in a working sample of 567 respondents who were interviewed on a variety of items regarding

educational, marital, occupational history, earlier gang membership, and social psychological variables. The interview was approximately one to two hours; no one located refused to respond. Of relevance to this particular paper, questions were asked about "hidden" offenses, those which were committed but for which the subjects were not arrested. Each person was asked if and how often he had committed any of 24 specific crimes before age 18 and after his 18th birthday. These items cover a full range of offenses from the very minor (disturbing the peace) to the very serious (homicide and rape). All subjects were interviewed around the time of their 25th birthday and all names were checked through police files at the time of their 26th birthday. At the time of subject's 22nd birthday a check had been made of the Philadelphia police files (Cardarelli, 1974). And since the interviews, we have investigated the 975, the ten per cent sample, for previous arrests and dispositions up to age 30. Hence we have several data banks about continued criminality to which I shall refer. They all include knowledge about juvenile (under age 18) official arrest record and juvenile self-reported offenses. For adults (18 and over) there are the following files:

1. 18-22: official arrest records
2. 18-26: official and self-reported offenses
3. 18-30: official arrest records.

Special computer runs are still being made since we received reports on our subjects up to age 30. Some of these runs had been made at earlier ages; this is the main reason that some of my findings are drawn from different files at this time.

Methodologically, there is one additional comment to be made, and that is about the application of weighted seriousness scores for each of the offenses committed by our cohort subjects. Derived from the work Sellin and I had done previously and reported in The Measurement of Delinquency (Sellin & Wolfgang, 1964), a psychological scaling study, the seriousness scores denote relative mathematical weights of the gravity of different crimes.

I shall not discuss here issues about reliability and validity of the sample, nor of the official or self-reported material. In our forthcoming book we cover these topics in detail. In short, however, we believe that the traditional scientific requirements of validity and reliability are satisfactorily met; we have been as comprehensively self-critical as possible and have had the benefit of distinguished colleagues.

Some Findings

Table 1 shows the relationship between juvenile and adult offender status by race in the analysis of five years into adulthood, or from ages 18-22 (Table 1 is derived from Cardarelli, 1974). Nearly 60 per cent of the birth cohort had no record of arrest, but 41 per cent did. Of this latter arrest-record group, 35 per cent had a record before 18; 22 per cent only as juveniles, 14 per cent before and after age 18. But it is important to note that only five per cent (4.82) had an arrest record only as adults, or after age 18.

It is also important to point out the differences between whites and nonwhites in this array. Among cohort subjects who had an official arrest record after age 18, or as adults, there are no statistically significant differences between the rates. That is, about five per cent of whites and six per cent of nonwhites obtain an arrest record only after age 18. But the socially and statistically significant fact is that blacks, or nonwhites, are four times more likely to have an arrest record before and after age 18 than are whites.

Moreover, when we examine the mean number of offenses for all subjects with both juvenile and adult arrest records (6.37) we note that the figure is about three times greater than for those who have only a juvenile record and more than three times as great for those with an adult record (1.94).

Conclusion

Serious offenses are committed frequently by a relatively small number of offenders: up to age 30 in a birth cohort, approximately 14 per cent. Serious offenses, officially known and self-reported, committed by juveniles, have a higher probability of being committed by these same persons as adults. Race is significantly associated with this finding, which is to say that proportionately many more nonwhites than whites will be involved in this serious juvenile/serious adult offender status grouping. But the transition stability also occurs among the proportionately smaller number of whites. The chronic offender continues to be the most important category with which the criminal justice system should deal in its concern about serious, particularly personal injury, offenses.

Perhaps as meaningful as anything to emerge from this longitudinal study thus far is that with respect to chronicity of offenders, the juvenile/adult statutory dichotomy has little justification. At whatever age the chronic offender begins his fourth or fifth offense, he will commit further offenses with very high probabilities, and, on the average, the next offense will be an index offense nearly half the time. It may be, therefore, that if the severity of the sanction is proportionate to the gravity of the crime and to the cumulative history of serious crime, the sanction should be similar for chronic serious offenders whatever their age.

Table 1,

Number and Per Cent of Cohort Subjects
By Offender Status and Race

OFFENDER STATUS	Race		Totals N/%
	White N/%	Non-White N/%	
I. Subjects with No Arrest Record	473 (66.71)	103 (38.72)	576 (59.08)
II. Subjects with Arrest Record	236 (33.29)	163 (61.28)	399 (40.92)
A. Before Age 18 Only	147 (10.73)	67 (25.19)	214 (21.95)
B. Before & After Age 18	58 (8.18)	80 (30.07)	138 (14.15)
C. After Age 18 Only	31 (4.37)	16 (6.01)	47 (4.82)
Totals	709 (100.00)	266 (100.00)	975 (100.00)

References

Cardarelli, A. P. Socio-economic status and delinquency and adult criminality in a birth cohort. Unpublished doctoral dissertation. University of Pennsylvania, 1974.

Collins, J. J., Jr. Offender careers and restraint: the probabilities and policy implications. Unpublished doctoral dissertation. University of Pennsylvania, 1977.

Monahan, J. The prediction of violent behavior in juveniles. Paper presented at the National Symposium on the Serious Juvenile Offender, Department of Corrections, State of Minnesota, Minneapolis, Minnesota, September, 1977.

Sellin, T. & Wolfgang, M. E. The measurement of delinquency. New York: John Wiley & Sons, Inc., 1964.

Tracy, P. E. A self-report study of delinquent and criminal behavior. Unpublished doctoral dissertation. University of Pennsylvania, 1977.

Wolfgang, M. E. Crime in a birth cohort. Proceedings of the American Philosophical Society, 1973, 117 (5), 404-411.

Wolfgang, M. E., Figlio, R., & Sellin, T. Delinquency in a birth cohort. Chicago, IL: University of Chicago, 1972.

Health Consequences of Drug Use:
A Longitudinal Study of Urban Black Youth

Ann F. Brunswick, Ph.D.
Public Health (Sociomedical Sciences) &
Center for Sociocultural Research on Drug Use
Columbia University

Health Consequences of Drug Use:
A Longitudinal Study of Urban Black Youth

Ann F. Brunswick, Ph.D.
Public Health (Sociomedical Sciences) &
Center for Sociocultural Research on Drug Use
Columbia University

Introduction

The chapter which follows describes a prospective study of a normal population in a natural setting. The normal population is a special population: inner city black adolescents. A community representative sample numbering 536 young black men and women was studied two times, six to eight years apart -- first when they were adolescents, ages 12-17, inclusive, and then as young adults, ages 18-23, inclusive. Information has been obtained in the course of this research which helps to illuminate the transition from adolescence to adulthood in a subgroup which is frequently, if not always, underrepresented in population based studies -- be they census or national epidemiologic surveys based on households, or the more frequent studies of youth drawn from school rosters.

The age transition has been examined here from the perspective of health stability and change (dependent variable), to test the role of drug use (independent predictor variable) in affecting the normal trajectory of health change over adolescence and into the early adult years. The reader hardly needs to be reminded that the study population is at particular risk on both matters: a propensity to poor health and to using drugs nonmedically abetted by its marginal position in society,

itself stemming from the twin "minority" jeopardies of race and youth. Such social marginality not only entails economic disadvantage, per se, but along with it severely limited opportunities for mobility -- at a time in the life span when society expects expanding competency and growth into responsible roles in education and/or work. Nor surprisingly, induction into drug use and/or dependence has been reported to peak in the youthful years spanned by this research (Brunswick & Boyle, 1979; Chambers & Ball, 1971; Lukoff & Brock, 1974; Robins, Darvish & Murphy, 1970).

The first wave of this research was conducted in 1968-70, with support from the Children's Bureau, subsequently reorganized into the Office of Maternal and Child Health, Dept. H.E.W. (Grant No. Mc-R-360040). Its purpose was to provide population-based data on health status and needs for health care, along with related attitudes and behavior. The data collected in that first wave provided a baseline in-depth study (from personal interviews with and medical examinations of adolescents, from data abstracted from their official school records, and from interviews conducted with one of their parents) around a wide range of health problems -- physical, psychophysical and emotional; around self attitudes; around aspirations, expectations and actual educational achievement. These comprise a biopsychosocial perspective on health (Brunswick, 1976, 1977 & 1979; Brunswick & Collette, 1977; Brunswick & Josephson, 1972; Brunswick & Nikias, 1975; Brunswick & Tarica, 1974; Brunswick, Boyle & Tarica, 1979).

The second wave, follow up interviews, was conducted in 1975-76 with support from the National Institute on Drug Abuse (Grant No. R01 DA 00852 and P01 DA 01097). Building on the detailed health base, patterns of nonmedical drug use in this normal population of black youth and the effects of those patterns on their subsequent health became the focus of restudy.

Because the chief concern of the present research is in how drug use affects the expected sequence of health, besides extensive measurement of the dependent variable (health status) and of the independent variable (drug use), the data set includes four domains of possible mediating variables: family background characteristics, role attainment, social influences (family, peer, institutional and media) and psychosocial attitudes. Finally, it is important to note the gender specific focus of this research, analysis of separate developmental sequences of drug use behavior and health outcomes among young black men and women.

Study Sample

In this section, procedures for drawing the initial study sample, the initial interview response rate and reasons for noninterview, follow up response rate and reasons for noninterview, the composition of the follow up sample, and the representativeness of the completed follow up sample will be discussed in turn.

Source of Sample

In this study, the sample of adolescents was drawn from a representative cross-section of households in Central Harlem. These households had been studied in a Community Health Survey jointly conducted by Harlem Hospital and Columbia University School of Public Health. The Community Survey was conducted between 1967-70 by the Harlem Hospital Center Department of Patient Care and Program Evaluation, and was headed by Professor Jack Elinson. The Community Health Survey was directed by Ms. Patricia Collette. Households were selected by means of a multi-stage probability sampling design. Each survey year, one out of twenty-five households was drawn into the sample, thus constituting an annual four percent sample of all housing units in Central Harlem. This was then divided into twelve representative monthly subsamples.

The adolescent health study used samples of housing units from the first two years of the larger community health survey. Households in the first monthly subsample were used for pretesting the procedures in the adolescent health study; thus, the first year of the adolescent study included monthly subsamples 2-13, the second year monthly subsamples 14-25. The first year, all youngsters in each household were listed if they would be, at the time of interview, no less than a month away from their twelfth birthday and no more than one month beyond their sixteenth. The study was expanded in the second year to include 16 and 17 year olds. In the second year's sample, therefore, the upper age limit was the eighteenth birthday. The original sample listing for this study was done by hand directly from the files of the larger Harlem survey. As a safeguard against errors in listing, when sample data from the larger survey became available on machine printouts, the hand listing was checked against the processed data; any cases which fell within the appropriate birthdates and had been missed earlier were then added to the lists. A small number of youngsters (25) were subsequently dropped from the sample because their ages were found to be beyond the limits set for this study.

In each of the two years in the initial study, personal interviews were completed with 83 percent of the listed sample, providing a total of 752 obtained interviews. Eleven percent of this total were black adolescents from hispanic backgrounds. For purposes of sample homogeneity, and since this group was too small for separate analysis, the hispanic background black youths were excluded from analysis. The effective sample base then became 668 black youths from nonhispanic backgrounds. The interview completion rate was 84 percent for girls and for the younger, 12-15 year old boys. Among 16 and 17 year old boys, interview completion amounted to 78 percent, providing an overall completion of 83.5 percent among nonhispanic black youth.

Differences between the interviewed sample (N = 688) and noninterviewed (N = 132) were examined with regard to a number of background variables. "As to mother's educational attainment, there was a tendency for better completion among adolescents whose mothers had not completed high school than among those whose mothers had more education. This was particularly true among adolescent girls: 87 percent who had mothers with less than a complete high school education were interviewed, compared to 75 percent of girls with mothers who had more education. Among boys the rates were 86 percent and 80 percent, respectively. Girls whose families were on welfare tended to show a higher interview completion rate than girls from families not receiving welfare assistance (90% and 79%, respectively). The difference among boys, though in the same direction, was small (87% and 82%, respectively). Girls with southern-born mothers were more likely to complete the interview than those whose mothers were born in New York City (87% and 76%, respectively). Again, the difference in interview completion rates among boys was in the same direction but only amounted to five percent, a difference easily attributable to random sampling fluctuation." (Note that the Harlem black sample in general derived from the American south, about three-quarters having mothers who were southern born.)

"The rate of interview completion during the two study years, 83 percent, was high for a community in a central city. The 17 percent who were not interviewed were accounted for as follows: refusal, 7 percent; moved and not located, 4 percent; out of area, 5 percent; never at home after five or more attempts, 1 percent. Considerable time, effort and resources were spent in locating and completing as many cases as possible."

Follow Up Response Rate

Rigorous procedures were similarly followed in locating the sample for follow up interviews, which were conducted in 1975-76 when the study group was aged 18-23. Interviewing was divided so that the older segment of the original study group (the original 16-17 year olds) were reinterviewed in 1975, and the others in 1976. In all, 536 of the 668 in the

original sample were reinterviewed, for a total completion rate of 80 percent. This represented 89 percent of those young people who were still alive and who had not been identified as living outside of the metropolitan New York City area. For financial reasons, the follow up sample was restricted to metropolitan New York City. Included also were ten interviews with respondents currently in jails outside of New York City but whose "usual" home address was in the city. In point of fact, current whereabouts were ascertained for 94 percent of the original sample. A further point of interest about the follow up sample is the probability that some young males between ages 18-23 who often are undercounted in decennial census have probably been retained in the follow up, an inference based on the equivalent male and female follow up rates.

A similar proportion of males and females were followed up. Of the potentially interviewable (those alive and not located outside of New York City), 90 percent of the males completed interviews compared to 88 percent of females. More females (4%) than males (1%) refused to be interviewed. Of interest, too, was the group of so-called "nomads" (3% of males, 1% of females) whose families could not direct us to their address because they had no known place of residence.

Longitudinal Sample

Reviewing what has been stated earlier, the sample for this analysis consists of 536 nonhispanic black young men and women whose familial origins were largely from the southern USA. They were ages 18-23 years, inclusive, at reinterview; when initially selected and studied 6-8 years earlier, at ages 12-17, they all were living in the Central Harlem Health District of New York City. At follow up, as in the initial study, the sample was 52 percent male and 48 percent female. Just fewer than two-thirds of both males and females (63% and 62%, respectively) still had residences in Central Harlem at reinterview. Moving patterns within the New York area were remarkably similar for both sexes.

By the follow up, about a third of the sample (34%) had not yet completed high school, just under a third (30%) had been graduated from high school but had gone no further, slightly more than a third (36%) had been graduated and continued studies beyond high school. At the time of reinterview, a quarter of the sample was attending school full time, and another seven percent were attending part time. Somewhat more than a third of the sample (39%) had jobs (three-quarters of these were full-time). More men than women were working: 44 percent of men (including 13% who were also attending school) and 33 percent of women (including 7% who were also in school). Four percent of the 18-23 year

old males were in jail when interviewed and two percent were in the armed services. (Remember that 5% of the original male sample were lost to interview because they were in the armed forces and outside the interviewing area.) Not quite a fifth of the females were married (18%) compared to four percent of the males. Half the females (52%) and a quarter of the males (28%) had at least one child.

To what extent can we accept the follow up sample as representing the entire initial study sample? The response rates, especially the low two percent refusal, suggest that the reinterviewed sample, that is, the sample available for longitudinal analysis, is substantially representative of all those in the initial sample. This was tested by comparing the reinterviewed sample with the full initial sample on approximately twenty characteristics at baseline. There was only one variable on which the reinterviewed group differed from the full initial sample by as much as five percentage points, with the difference at $p < .05$: whether or not the young person had come in for the medical examination. The direction of the difference ran counter to what might have been expected: the reinterviewed sample included slightly more, proportionately, of those who had not cooperated in the medical examination.

Data Collection and Measures Taken

As mentioned earlier, there were four modalities through which initial data were collected: personal interviews with adolescents, interviews with a parent of the adolescent, comprehensive medical screening examinations arranged by the research project staff and conducted at the neighborhood municipal hospital, and abstracts of official school records. (More detail on these procedures can be found in Brunswick & Josephson, 1972.) The follow up was conducted through personal interviews only. Procedures for the personal interview were similar during each of the two study waves (except that respondents were paid \$10 at the second interview, in lieu of a free medical examination). Procedures used in the follow up will be described below.

Field work or data collection in the follow up, as in the first wave, was conducted in two stages. The first 18 months of the restudy was a "feasibility" phase, to include the 126 oldest adolescents from the initial study. The decision to partition the research into a first stage "feasibility" study was made by the funding agency which questioned whether the study group could be relocated. (In the initial study, first year support similarly had been awarded at a reduced level until skepticism about successfully interviewing and

examining a community cross section of Harlem youth had been satisfied. That is why the first study year included only 12-15 year olds; funds were increased in the second year which enabled the addition of 16-17 year olds to the base sample.)

The personal interviews were conducted at home (except for a few cases where respondents preferred locations outside the home) by specially trained interviewers, recruited from the black community, and matched to respondents on race and gender. The first 87 interviews in the restudy were conducted between April-September, 1975, as the feasibility phase recommended before undertaking the full scale follow up. The remaining 449 (84%) of the interviews were conducted from April to September, 1976. In the first restudy year, interviewing was supervised directly from the Principal Investigator's office. (This had been the case also during the years of interviewing in the initial study.) In the second and larger restudy year, interviewing was subcontracted to the National Opinion Research Center, University of Chicago (New York office), following the same procedures as were used in the feasibility phase.

Sampling weights of the original study were retained in the restudy. This consisted of doubling the weights of older adolescents because of their lower probability of inclusion in the sample, i.e., in only one of the original two study years. Fractional weights were applied, in appropriate ratios and separately for males and females, so that weighting did not inflate the number of cases. Due to rounding, the weighted number of female interviews is one less than was actually conducted (258 rather than 259) and the total weighted count of interviews is 535 instead of the 536 actually conducted.

The 54-page, 162-question personal interview was developed after: a) conducting exploratory, unstructured interviews with a small number of black youths in Harlem and Washington Heights for their accounts of how young people talk about and use drugs; b) culling relevant items from other drug survey instruments; c) pretesting procedures in the structured interviews scheduled with a small number of respondents who were part of the pretest sample in the first stage of the study in 1968.

The structured personal interview schedule's contents included the following measures:

- a. General questions on current life conditions. (Questions (Q.'s) 1-2)

b. Health, including perceptions of health, rating of own health status, use of health services, reports of health problems and treatment for them. (Q.'s 3-27)

c. FOR WOMEN ONLY. Pregnancy history and use of obstetrics and abortion services. (Q.'s 28-35)

d. DRUGS: (Q.'s 36-73)

Q.'s 36-39: Perceptions and opinions about drug use and trends in drug use in Harlem.

Q.'s 40-46: Own marijuana use.

Q.'s 47-57: Own nonmedical use of other drugs (opiates, barbiturates, etc.) including combinations with alcohol and marijuana.

Q.'s 58-67: Problems resulting from drug use; drug and alcohol treatment experiences and attitudes toward them.

Q.'s 68-73: Attitudes of drug users and of non-users regarding motivations for using and for stopping drug use.

e. Eating, sleeping and cigarette smoking behavior. (Q.'s 74-80)

f. Alcohol drinking. (Q.'s 81-90)

g. Work and educational aspirations, experiences and satisfactions; military and institutional experiences. (Q.'s 92-122)

Q.'s 92-94: Career aspirations and expectations.

Q.'s 95-98: Major activity if not working or in school.

Q.'s 99-105: Current (or most recent) work.

Q.'s 106-113: Education achievement, aspirations and expectations.

Q.'s 114-120: Experience in the Armed Forces.

Q. 121: Other institutional experiences.

Q. 122: Five-year summary of work and school history.

h. Emotional and psychosocial indicators: mood, outlook on future and feelings about self. (Q.'s 123-133)

i. Interpersonal relationships, satisfactions and difficulties. (Q.'s 134-140)

j. Other personal and background information: age, birthplace, parity, fertility desires/expectations, residential mobility, household composition and marital status, sources of income. (Q.'s 141-154)

k. Leisure time, political and religious involvement. (Q.'s 155-160)

l. Interest in further health and/or drug information. (Q.'s 161-162)

m. Interviewer's report including observations relevant to possible drug use (filled in immediately after leaving respondent). (Items 1-15)

These measures comprise the variables tested as independent predictor (drug use); dependent health outcome (scaled into emotional, psychophysical and residual physical symptom measures); and mediating domains of family background, role attainment and current living situation, social influences including family-peer contacts/informal and formal group affiliations/leisure activities, psychosocial attitudes. These variables and measures on which they were based are listed in Table 1. With minor exceptions, all health and mediating variables were measured at initial interview as well as at restudy. Baseline health data also include findings from physical examination.

Retrospective Report on Independent Variable

Only the independent variable (drug use) was not assessed systematically at the initial interview. Given the importance of establishing the time of onset of drug use, onset was inquired about in different ways to provide multiple measures that could be used as checks against each other, thereby increasing the reliability of the measure.

In the personal interview, respondents were asked whether or not they had used each of nine drugs or classes of drugs which were read to them one at a time: marijuana, acid or other psychedelic, cocaine, heroin, methadone, "uppers," "downers," "glue or some other inhalant," alcohol. To establish when use of a given drug was initiated, respondents were asked how old they were when they first took it, how long ago it was, and what their major activity was at the time. As noted above, all three questions were asked in order to improve measurement of onset and to facilitate accuracy in recall. Consistency of the three independent measures was checked before time of onset was coded for a respondent. Usual frequency and recency or last time used were among other items of information obtained about the use of each drug.

Reliability of drug reports on personal interview has been adequately established (Ball, 1967; O'Donnell, 1976; Robins, 1973; Robins & Murphy, 1967; Robins, Davis & Nurco, 1974; Single, et al., 1975; Stephens, 1972; Whitehead & Smart, 1972). Indeed, O'Donnell has shown that opiate users and/or addicts, i.e., those who are most drug involved, are more likely than others to be candid about socially disapproved behaviors. Generally, to the extent that there is bias in self report, it tends toward underreporting not overreporting or nonreporting and it affects report of current use rather than of earlier use.

Reliability of report in the current study was tested by checking replies on reinterview of thirteen earlier identified drug users. They were among 17 drug users who had been so identified at the initial interview (of the remaining, two had died and the other two had moved out of the area). Ten of these thirteen, or 77 percent, reported retrospectively on reinterview dates of onset which were consistent with actual initial findings. One other respondent's report differed by one year and another by two years. All but one corroborated earlier reported use. That one exception turned out to be a young man who had earlier reported a single use of cocaine which had made him ill and he had no further contact with drugs.

In sum, the retrospective report of time of onset of drug use was accurate when a one year margin of error was allowed. In the analysis, therefore, onset or reported first use of a drug which fell within one year of the time of initial interview was conservatively considered use by time one. Relatively few cases were contingent on this kind of a decision.

Furthermore, the advisability of considering retrospective recall of dates as occurring within a one-year margin of error was consistent with a reliability check performed on another variable -- reported data of first pregnancy among young women who had had a pregnancy by the time

of the initial interview. When reports of age at first pregnancy taken in adolescence were compared with those reported retrospectively six to eight years later, 9 of 15 (60%) were correct to the year; for another third (5) of the cases, reports on reinterview fell within one year of the initial report, with no consistency in earlier or later attribution. Thus, taken with the one-year margin of error, retrospective recall matched earlier report in 14 of 15, or 93 percent of cases of early adolescent pregnancies.

Results to Date

This paper has emphasized procedural and methodological issues, since those would seem to be the unifying theme of this book. Areas of results will be adumbrated here, with reference to an already published source when one is available:

1. Health decline, measured by an increase in the number of self reported physical and psychophysical complaints occurred two years earlier among young black women than men. For this reason, the health decline between adolescence and early adulthood was greater for young men than women. While girls reported significantly more symptoms than boys as adolescents, by young adulthood the gender difference in number of self reported health problems was small and not significant (Brunswick, 1980).
2. Greater prevalence of illicit drug use was demonstrated in this sample than in similarly aged white samples. As to patterns of use, licit drugs (alcohol and cigarettes) are used more heavily by those who used illicit drugs; "softer" illicit drugs (e.g., marijuana) are used more heavily by those who used harder illicit drugs as well. That is, substances are not "substituted" for one another but incrementally used as harder drugs are used (Brunswick, 1979, 1980).
3. A significant "cohort" effect was observed in relation to heroin use, indicating the importance of analyzing historic and secular time influences on drug use patterns (Boyle & Brunswick, 1980).
4. More generally, cohort or historic time influences prevalence of use of particular drugs, but not its incidence, i.e., age at first use (Brunswick & Boyle, 1979).
5. Roughly half of the young heroin users in this inner city sample stopped using heroin without treatment. But when treatment was

used, it was used differently by men and women, as suggested by the finding that women used heroin longer before entering treatment and remained in treatment longer than did men (Brunswick, 1979).

6. The longitudinal analyses for drug effects on health are still underway. Two findings that emerge through many different analytic approaches are: glue (inhalant) use by adolescent boys has a direct impact on psychophysical health; heroin use by adolescent girls, especially early use, has a significant and lagged effect on their psychophysical symptoms.

The experience in longitudinal research which has been chronicled in this chapter has opened new vistas, both of analytic problems and possibilities. Areas of new awareness are numerous; one in particular is worth closing this chapter with. That concerns adequately identifying the temporal relationships between variables in longitudinal research. Ideally, longitudinal studies should be designed with short enough intervals between measurement so as to capture the onset of the behaviors of interest while the overall study span is long enough to capture its effects. But just what "short enough" and "long enough" are for different variables has yet to be determined.

References

Bass, J.C. The reliability and validity of interview data obtained from 59 narcotic drug addicts. American Journal of Sociology, 1967, 72(6), 650-654.

Boyle, J.M. & Brunswick, A.F. What happened in Harlem? Analysis of a decline in heroin use among a generation unit of urban black youth. Journal of Drug Issues, 1980, 10(1). Also in E. Josephson and O. Ochs, (Eds.), Drug policy and social research, Washington, D.C.: Hemisphere Publishing Co., 1980.

Brunswick, A.F. Indicators of health status in adolescence. International Journal of Health Services, 1976, 6(3), 475-492. Also in J. Elinson and A. Siegmann, (Eds.), Indicators of health status, Farmingdale, N.Y.: Baywood Publishing Co., 1979.

Brunswick, A.F. Health and drug behavior: Preliminary findings from a study of urban black adolescents. Addictive Diseases, 1977, 3(2), 197-214.

Brunswick, A.F. Black youths and drug-use behavior. In G. Beschner & A. Friedman, (Eds.), Youth and drug abuse: Problems, issues, and treatment, Lexington, Mass.: Lexington Books, Inc., 1979.

Brunswick, A.F. Health stability and change: A study of urban black youth. American Journal of Public Health, 1980 (a), 70(5), 504-513.

Brunswick, A.F. Social meanings and developmental needs: Perspectives on black youths' drug use. Youth and Society, 1980 (b), 11(4), in press.

Brunswick, A.F. & Boyle, J.M. Patterns of drug involvement: Developmental and secular influences on age at initiation. Youth and Society, 1979, 11(2), 139-162.

Brunswick, A.F., Boyle, J.M. & Tarica, C. Who sees the doctor? A study of urban black adolescents. Social Science and Medicine, 1979, 13A(1), 45-56.

Brunswick, A.F. & Collette, P. Psychophysical correlates of elevated blood pressure. Journal of Human Stress, 1977, 3(4), 19-31.

Brunswick, A.F. & Elinson, J. A bridge over troubled water: New roles for old researchers. Paper presented at Twenty-Eighth Annual Conference, American Association for Public Opinion Research, Asheville, North Carolina, May 1973.

Brunswick, A.F. & Josephson, E. A Study of adolescent health in Harlem. Supplement to American Journal of Public Health, 1972, 62.

Brunswick, A.F. & Nikias, M. Dentists' ratings and adolescents' perceptions of oral health. Journal of Dental Research, 1975, 54(4), 836-843.

Brunswick, A.F. & Tarica, C. Drinking and health: A study of urban black adolescents. Addictive Diseases, 1974, 1(10), 21-42.

Chambers, C.D. & Ball, J.C. An assessment of drug use in the general population. Special Report No. 7, New York State Narcotics Addiction Control Commission, May 1971.

Josephson, E. Resistance to community surveys. Social Problems, 1970, 18(1), 117-129, Summer. Reprinted in R.K. Yin (Ed.), The city in the seventies, Itasca, Illinois: F.E. Peacock Publishers, 1972.

Kaplan, H.B. & Meyerowitz, J.H. Social and psychological correlates of drug abuse: a comparison of addict and non-addict populations from the perspective of self-theory. Social Science & Medicine, 1970, 4, 203-335.

Lukoff, I.F. & Brook, J.S. A sociocultural exploration of reported heroin use. In C. Winick, (Ed.), Sociological aspects of drug dependence, Cleveland, Ohio: CRC Press, 1974.

O'Donnell, J.A. Survey data as contributors to estimation. In The epidemiology of heroin and other narcotics. NIDA Research Monograph #16, Washington, D.C.: Government Printing Office, 1976.

Robins, L.N. A follow-up of Vietnam drug users. (Interim Final Report, Special Action Office Monograph, Series A, No. 1), Washington, D.C.: Government Printing Office, 1973.

Robins, L.N., Darvish, H.S. & Murphy, G.E. The long-term outcome for adolescent drug users: a follow-up study of 76 users and 146 non-users. In J. Zubin & A.M. Freedmen, (Eds.), The psychopathology of adolescence, London: Grune and Stratton, 1970.

Robins, L.N., Davis, D.H. & Nurco, D.N. How permanent was Vietnam drug addiction? Supplement to American Journal of Public Health, 1974, 64, 38-44.

Robins, L.N. & Murphy, G.E. Drug use in a normal population of young negro men. American Journal of Public Health, 57, 1580-1596.

Single, E., Kandel, D.B. & Johnson, B.D. The reliability and validity of drug use responses in a large scale longitudinal survey. Journal of Drug Issues, 1975, 5(4), 426-443.

Smith, B.H. Competence and socialization. In J.A. Clausen, (Ed.), Socialization and Society, Boston: Little, Brown & Co., 1968.

Stephens, R. The truthfulness of addict respondents in research projects. International Journal of Addiction, 1972, 7, 549-558

White, R.W. Motivation reconsidered: The concept of competence. Psychological Review, 1959, 66, 297-333.

Whitehead, P.C. & Smart, R.G. Validity and reliability of self reported drug use. Canadian Journal of Criminology and Corrections, 1972, 14, 108.

Zigler, E. & Trickett, P.K. IQ, social competence, and evaluation of early childhood intervention programs. American Psychologist, 1978, 33(9), 789-798.

Table 1

List of Variables and Their Measures in The Data Set
(Grouped according to role in a causal model
of the effect of drug use on health)

A. Predictor Variables: Drug Use Behavior

- (1) Patterns of drug use for individual drugs/drug types: alcohol, marijuana, cocaine, heroin, methadone, other opiates, psychedelics, amphetamines, barbiturates, glue and other inhalants.

<u>VARIABLE</u>	<u>MEASURE IN DATA SET</u>
Ever use	Ever use specified drugs.
Frequency of use	Current/past usual use: daily, few times a week, few times a month, once a month, few times a year, or once/twice only.
Recency of use	Time since last use: number of days, weeks, months or years.
Onset of use	Age at first use. Chronological year of first use.
Duration of use	Time between first use and last use.
Mode of administration	Manner in which drug usually taken: smoking, swallowing, skin popping, snorting, injecting.
Self classification	Constant user, occasional user, former user, experimenter only.
Future use intent	Will use again: yes, no, maybe.
Peak use	Ever used more than current/past usual use. Age at onset of peak use. Time between onset and peak use. Frequency at peak use. Mode of administration at peak use.
Trouble as a result of use	Drug(s) causing trouble. Specified types of mental and physical impairments, accidents and trouble with law.

Table 1 (continued)

A. Predictor Variables (continued):

<u>VARIABLE</u>	<u>MEASURE IN DATA SET</u>
Treatment	Ever been in treatment: number of times in treatment. Drug treated for. Reason for treatment and treatment objective (each episode). Type of treatment facilities and services provided (with each episode). Duration of treatment (each episode). Outcome of treatment (each episode).
(2) Patterns of polydrug use or total drug use pattern: aggregate use pattern across alcohol, marijuana, cocaine, heroin, methadone, other opiates, psychedelics, amphetamines, barbiturates and glue and other inhalants.	
<u>VARIABLE</u>	<u>MEASURE IN DATA SET</u>
Polydrug use	Number of drugs ever used. Types of drugs ever used. Number of drugs used non-experimentally. Types of drugs used non-experimentally.
Onset of drug use	Age first began using any drugs. Age first began using drugs non-experimentally.
Sequence of drug use	Order of onset of drugs used.
Recency of drug use	Time since last use of any drug: number of days, weeks, months or years.
Duration of drug use	Time between use of first drug and most recent use of any drug.
Total drug use score	Frequency of use (converted into number of times used per year) multiplied by the duration of use, for each drug, summed across all drugs.
Drug combinations	Ever take drugs together (at the same time). Drug combinations used. Reason for using specific combinations. Frequency of use of specific combinations. Age at first use of specific combinations. Recency of last use of specific combinations.

Table 1 (continued)

A. Predictor Variables (continued):

<u>VARIABLE</u>	<u>MEASURE IN DATA SET</u>
Developmental drug use	Onset of drug use for specific ages. Current drug use at specific ages. Termination of drug use at specific ages. Treatment of drug use at specific ages.
Chronological drug use	Onset of drug use by chronological year (e.g. 1970). Current drug use by chronological year. Termination of drug use by chronological year. Treatment of drug use by chronological year.

B. Outcome Variables: Health Conditions (Physical, nervous and psychosomatic), General Health Status.

<u>VARIABLE</u>	<u>MEASURE IN DATA SET</u>
Health troubles and conditions experienced in past year	Thirty-five specific health problems and conditions: (repeated) headaches, trouble seeing, nervous problems, menstrual, dizziness, speech, nosebleeds, colds, sore throats, earaches, hearing, sinus, cough, chest pains, asthma, shortness of breath, heart thumping, other heart, shaking/trembling, backaches, pains in limbs, skin problems, allergies, hernia, constipation, diarrhea, stomach pains, stomach ulcer, vomiting, indigestion, blood in bowel, worms, blood in urine, painful urination, serious accident/injury, and a residual category for other volunteered conditions.
Health conditions ever experienced	Eleven critical health problems: diabetes, tuberculosis, anemia, syphilis, gonorrhoea, polio, high blood pressure, hypertension, jaundice/hepatitis, epilepsy, rheumatic fever, and a residual category for other volunteered conditions of this type (e.g. meningitis).

Table 1 (continued)

B. Outcome Variables (continued):

<u>VARIABLE</u>	<u>MEASURE IN DATA SET</u>
Development and maturation	Height. Weight. Nutritional self appraisal (overweight, underweight). Physician appraisal of maturation, including body development and weight (T ₁) only.
Dental conditions	Problem with teeth and/or gums.
Number of self reported health conditions	Number of health conditions in past year (maximum of 36) and number of chronic health conditions (maximum of 12) and current dental problems (maximum of 1).
Disability	Number of days in hospital. Number of days at home in bed. Number of days unable to carry on usual activities. Summary measure of the above.
Limitation of activities	Health stops from going to work/school. Health requires special kind of work/school. Health limits time or activities at work/school. Any of the above limitations.
Handicap	Type of handicap. Financial aid for handicap. Interviewer's observation.
Self evaluation of health	General health rating. Current health compared to 5 years ago.
Psychosomatic disorders	Bad dreams and nightmares. Trouble going to sleep. Trouble staying asleep. How good is appetite. Psychosomatic conditions from current health conditions: nervous/emotional problems, heart thumping, shaking/trembling. Number of above psychosomatic problems.

Table 1 (continued)

B. Outcome Variables (continued):

Psychological Wellbeing

<u>VARIABLE</u>	<u>MEASURE IN DATA SET</u>
Depressed outlook	Sort of time having. How happy in general. How often feel alone among friends. How often think about dying. What think might die from. Wonder if anything worthwhile.
Anxiety/worry	Worry about the future. Upset by personal problems.
Self esteem	Want to change self. Feel useless at times. Times feel no good at all. Able to perform as well as others. Satisfaction with weight. Satisfaction with height.

Symptom measures were reduced into three scales as described below.

Psychological Health

PSY.

8 Items, Guttman Lower-bound Reliability = .66 (Males=.62, Females=.69)

1. In general, what sort of time are you having these days -- not so good, pretty good or a very good time?
2. And would you say most of the time you feel very happy, pretty happy or not so happy?
3. Some people worry a lot about what their lives will be like five or ten years from now and some people don't. Would you say you worry a lot about that, that you worry a little about it, or don't you worry about it at all?
4. About how often do you have a personal problem that upsets you -- does that happen very often, does it happen sometimes or are you hardly ever upset by a personal problem?
5. I certainly feel useless at times: agree, strongly agree/disagree, strongly disagree.

Table 1 (continued)

Scales (continued):

6. At times I think I am no good at all: agree, strongly agree/disagree, strongly disagree.
7. I sometimes can't help wondering if anything is worthwhile anymore: agree, strongly agree/disagree, strongly disagree.
8. Do you feel somewhat alone or apart, even among friends?

Psychophysical Health

PSY.-PHYS.

9 Items, Guttman Lower-bound Reliability = .7r (Males=.73, Females=.75)

During the past 12 months, have you had: repeated headaches; nervous or emotional trouble; dizziness, fainting spells or blackouts; chest pains; short of breath even without exercise; heart thumping or racing; shaking or trembling; stomach pains; indigestion, acid in the stomach.

Physical Health

PHY.

38 Items for males, 39 for females; Guttman Lower-bound Reliability = .64 (Males=.67, Females=.69)

During the past 12 months, have you had: trouble seeing (even when wearing glasses); problem with your menstrual period; problem with your speech; repeated nosebleeds; frequent colds; repeated sore throats; frequent earaches or discharge from ears; trouble hearing; repeated sinus trouble; long-lasting cough or chronic bronchitis; asthma, wheezing; anything else wrong with the heart; problem with your skin; hay fever or other allergy; hernia, rupture or swelling in the groin; frequent constipation; frequent diarrhea; stomach ulcer; vomiting, feeling sick to your stomach; blood in the bowel movement; worms or parasites; blood in your urine; difficult, painful, frequent urinating; serious accident or injury where you had to stay away from your usual activities for at least one day or longer; musculoskeletal problem (backaches and limb problems).

And have you ever had: diabetes; tuberculosis; anemia (thin blood); syphilis; gonorrhea; polio; high blood pressure; low blood pressure; hypertension; jaundice, hepatitis; epilepsy, convulsions, fits, seizures; rheumatic fever; miscellaneous.

How often would you say you have problems with your teeth and gums -- like cavities, toothaches or bleeding gums? How do you feel about your weight (underweight, overweight)?

Table 1 (continued)

- C. Mediating Variables: (1) Intrapersonal Measures.
 (2) Interpersonal Relations and Social Activities.
 (3) Social Roles and Life Settings.
 (4) Demographic and Social Structural.

<u>VARIABLE</u>	<u>MEASURE IN DATA SET</u>
(1) Intrapersonal Measures	
Sense of Personal Efficacy	Always stopped from getting ahead. Little chance to be successful. Good luck more important than hard work.
Life Satisfaction	How things have been going last 5 years. Satisfaction with number of friends. Satisfaction with job. Satisfaction with military service.
Alienation/Hopelessness	Can't make a better world. People would sooner help others. Public officials aren't interested.
Time Orientation	Able to look ahead in life -- 5 years. Able to look ahead in life -- 10 years. Perception of best time in life. Ability to defer gratification.
Aspirations/Expectations	Education. Career. Age at marriage. Length of life.
Disparity/Congruity in Aspiration/Expectations	Above listed regarding education, occupation, marriage, length of life.
Religiosity	Importance of religion to self.
(2) Interpersonal Relations and Social Activities	
Social Networks	Number of close friends. Time spent with friends, family, alone. Number of social activities. Type of social activities. Formal, informal group memberships. Party (political) identification. Voting behavior. Church attendance. Number of friends from outside neighborhood.

Table 1 (continued):

- C. Mediating Variables (continued):

<u>VARIABLE</u>	<u>MEASURE IN DATA SET</u>
(3) Social Roles and Life Settings	
Economic Status	Household income. Amount of own earnings. Per person household income. Source of income (if welfare, how long). Poverty/non-poverty area of residence.
Educational Status	Currently in school. Years of education completed. Diploma currently working for. Diploma or degree completed. School grades (marks) (from initial study).
Employment/Occupation Status	Currently employed. Would like to work/looking for work, look for another job. Hours per week work. Occupational classification. Continuity of employment. Length of present employment. Length of present unemployment. Employment pattern over past 5 years.
Marital Status and Household Composition	Now/ever/never married. Age at marriage. Living alone or number of persons in household. Head of household. Living with parents. Number of people in household.
Mobility	Time in present building. Time in present neighborhood. Number of neighborhoods/past 5 years. Time in New York City. Time outside New York City/past 5 years. Number of different schools attended. Number of different jobs held/past 5 years.

Table 1 (continued)

C. Mediating Variables (continued):

VARIABLE	MEASURE IN DATA SET
(3) Social Roles and Life Settings (continued)	
Fertility: Pregnancy Experience (females only)	Ever pregnant. Age at first pregnancy. Number of pregnancies. Pregnancy spacing.
Pregnancy Outcomes (females only)	Number of live births. Number of abortions. Number of miscarriages. Number of stillbirths. Number of neonatal deaths.
Children (males and females)	Any children. Number of children. Age at first birth.
Fertility Aspirations and Expectations (males and females)	Number of children want to have. Number of children expect to have. Desired age at first birth (if no children). Disparity/Congruity in Aspirations/Expectations.
(4) Demographic and Social Structural (Antecedent Background)	
Age	Birthdate.
Birthplace	State or country.
Mother's birthplace	State or country.
Mother's education	Years of education completed.
Female headed household	No older male in household.
Family on welfare	Family receiving welfare assistance.
Household density	Number in household.

Interim Report on a Prospective Study of
Children at Risk for Schizophrenia

L. Erlenmeyer-Kimling, Ph.D.^{1,2/}

Barbara Cornblatt, Ph.D.

Joseph Fleiss, Ph.D.^{3/}

David Friedman, Ph.D.

Yvonne Marcuse, Ph.D.

John D. Rainer, M.D.

Jacques Rutschmann, D.Sc.

Thomas Chin, M.A.

Ulla Hilldoff, M.A.

Ruth Meyerson

Anne Moscato

Stuart Troy

^{1/} All authors except J. Fleiss: New York State Psychiatric Institute, Department of Medical Genetics, 722 West 168th Street, New York, N.Y. 10032

^{2/} Department of Psychiatry, Columbia University: L. Erlenmeyer-Kimling, Professor; B. Cornblatt, D. Friedman, Y. Marcuse, Research Associate; J.D. Rainer, Professor; J. Rutschmann, Associate Professor.

^{3/} J. Fleiss: Professor and Head, Division of Biostatistics, Columbia University, School of Public Health.

Interim Report on a Prospective Study of
Children at Risk for Schizophrenia

L. Erlenmeyer-Kimling, Ph.D.

Barbara Cornblatt, Ph.D.

Joseph Fleiss, Ph.D.

David Friedman, Ph.D.

Yvonne Marcuse, Ph.D.

John D. Rainer, M.D.

Jacques Rutschmann, D.Sc.

Thomas Chin, M.D.

Ulla Hilldoff, M.A.

Ruth Meyerson

Anne Moscato

Stuart Troy

In the past two decades, investigators in schizophrenia research have become aware of the difficulties of disentangling etiological variables through retrospective studies of schizophrenic patients. Investigators have recognized, moreover, that preventive intervention will be possible only when a better understanding of early precursors of the disorder is achieved. For this reason, a number of workers have turned to the prospective study of populations that are at higher than average statistical risk for developing schizophrenia at some time

during the lifespan. The high risk model for schizophrenia research was first suggested by Pearson and Kley (1957) who pointed out that children of two schizophrenic parents were a strategic population for study. Such children have a risk of about 39% for developing schizophrenia (Erlenmeyer-Kimling, 1968) as opposed to a general population risk of 1-2%. In the late 1950s and the 1960s, high risk studies were initiated by Fish (1962), Sobel (1961), Kallmann (Kallmann et al., 1964) and Mednick and Schulsinger (1968). Some of the studies focussed on the offspring of schizophrenic mothers (the Fish, Sobel and Mednick & Schulsinger studies), with a risk of about 12% (Erlenmeyer-Kimling, 1977). The risk to offspring of schizophrenic fathers is the same as the risk to offspring of schizophrenic mothers. The Kallmann study examined the offspring of two schizophrenic parents who are known to have a very high risk.

The early studies have been followed by a spate of high risk investigations, many of which examine their subjects longitudinally. Most, but not all, of the studies take children of schizophrenic parents as the population at risk. Although the majority of schizophrenics do not have an affected parent -- and it may therefore be argued that the children of schizophrenic parents represent biased samples -- such children are, at least for the present, the most economical and accessible group of high risk subjects for prospective investigations. Screening of unselected populations of children would require initial samples at least ten times as large as those drawn on children with one schizophrenic parent to obtain a comparable number of preschizophrenic individuals (Pearson & Kley, 1957).

The various high risk studies cover a wide range of ages, from infancy in some projects, to early adulthood in others, and approach the problem of schizophrenia from differing theoretical orientations. Many of the studies now in progress have been reviewed by Garnezy (1974), Erlenmeyer-Kimling (1975) and Rieder (in press).

High risk research has several possible goals. The first is the development of the natural history of the disorder under study. The second is concerned with the identification of the specific members of the high risk group who are most likely to develop schizophrenia; in other words, the research goal is to find predictors of later negative outcome. When one parent is schizophrenic, some 10 to 15 percent of the group should be discriminated, if the measurement battery were working accurately.

The third goal is to determine whether early predictors of schizophrenia are indicative of biological deficits underlying the disorder. While the second goal would make possible the identification of individuals in whom intervention might be attempted, goal three might suggest what form the intervention should take.

A fourth goal may prove very difficult to achieve in practice. Ideally, it should be possible to separate high risk subjects into three groups: those who are not genetically predisposed -- that is, subjects who have not received the critical gene, or the critical number of genes, from their schizophrenic parent -- versus subjects who are genetically predisposed and eventually develop schizophrenia versus genetically predisposed subjects who remain free of the disorder. If such distinctions were possible, we might learn a great deal about the environmental variables that interact with the genotype to produce schizophrenia, as well as environmental variables that help to buffer against the prospect of becoming ill.

These are all long term goals that depend on followup over a period of some years and on the validation of eventual outcome. High risk research is still too new to have achieved its goals but should begin to do so within the next decade.

The New York High Risk Study

The New York high risk study was initiated in 1971 with the aim of longitudinally examining developmental characteristics of high risk and comparison children in the New York metropolitan area. The study called for the examination of a large number of variables over a period of several years. Most of the variables were selected because they were thought to reflect functions that had been reported to be disturbed in schizophrenic patients, and at the same time because they had some developmental continuity over a wide span of ages. In particular, we were interested in measures that might reveal deficits in information processing, either behaviorally or physiologically, such as the deficits frequently observed in schizophrenic patients. Our laboratory procedures thus concentrate rather heavily on measures of attention, distractibility, response latency and neurophysiological functioning.

Description of the Samples

A first sample of subjects (Sample A) has been followed longitudinally since its recruitment in 1971-72. Assessment of a second sample (Sample B) is nearing completion.

Sample A consists of 205 children who were between the ages of 7 and 12 years at the time of intake into the sample (see Table 1). They are: 80 subjects at high risk for schizophrenia (44 children of schizophrenic mothers, 23 children of schizophrenic fathers and 13 children at exceptionally high risk by virtue of having two schizophrenic parents) and 125 subjects at low risk for schizophrenia. Of the low risk children, 100 had parents who had never been hospitalized or treated for a psychiatric disorder at the time of intake into the study and 25 had a parent

with another psychiatric disorder which is not considered to be related to schizophrenia -- usually affective disorder. While the latter subjects may be at risk for conditions other than schizophrenia, they constitute an important comparison group. Thus, a goal of the study is to identify variables that are not simply characteristics of children of mentally ill parents in general, but that differentiate the children of schizophrenic parents from children of parents with other psychiatric disorders. (Group labels are: HR = high risk group, subjects with one and with two schizophrenic parents; CP = psychiatric comparison group; CN = normal comparison group.)

Children who were in the target age range at the time of intake are designated the "study children" to distinguish them from their 122 younger (below age 7) and older (above age 12) siblings who have been studied less intensively. Developmental data and followup histories, but no laboratory data, have been collected on these siblings.

Sample B now includes 150 study children (see Table 1). As in Sample A, Sample B children are aged 7-12 years at the time of intake. Developmental and followup data are also being gathered on 86 younger and older siblings of the study children. Sample B affords a test of the replicability of findings derived from the study of Sample A as well as a means of testing hypotheses that may be suggested by data gathered on Sample A.

In both samples, the children of mentally ill parents were ascertained through the admission of the parent to one of several state psychiatric facilities in the New York metropolitan area. Consecutive admissions at these facilities were screened to identify patients who were white, English speaking, still married to the biological parent of a child or children in the target age range and who had at least one 7-12-year-old child. (The criterion of an intact marriage was waived if both parents were schizophrenic.) An effort was made to obtain families with two or more target aged children so that sibling comparisons might be made in analyses of gene-environment interaction. Patients with diagnoses of chronic alcoholism, drug addiction, brain trauma or psychoses of toxic origin were excluded. As an initial diagnostic evaluation, records of patients who passed the screening criteria were reviewed independently by two psychiatrists in our group after all references to hospital diagnoses and medications had been removed from the records. Each of the reviewers assigned a diagnosis based on the record materials, completed a symptom checklist and scored the 100-point Global Assessment Scale (Endicott et al., 1976) assessing the severity of functional impairment. Only those cases in which there was a full diagnostic agreement were retained for the study. In Sample B, in addition, Research Diagnostic Criteria (Spitzer et al., 1977) are used in the initial diagnostic evaluation.

A second diagnostic evaluation is being obtained in Sample A and will be obtained in Sample B by interview with the patients using the Schedule for Affective Disorders and Schizophrenia -- Lifetime Version (SADS-L, Spitzer & Endicott, 1977). Further evaluations will take into account all of the accumulated material on each patient, including the Minnesota Multiphasic Personality Inventory (MMPI) which is administered during a home visit.

In Sample A, the normal comparison (CN) group was obtained with the cooperation of two large school systems that agreed to send our letter asking for participants to families that met the screening criteria described above (i.e., white, English speaking, intact family, one or more 7 to 12-year-old children). The obvious disadvantage was that the CN group was a volunteer group. For Sample B, the latter problem was avoided by enlisting a population sampling firm to conduct a survey for the purpose of obtaining demographic information on a large number of families, which then formed a pool from which matches from the Sample B high risk subjects are being drawn. Matching is based on the family's socioeconomic level and the age and sex of the children. In both Samples A and B, families are excluded from the CN group if either parent is found to have had psychiatric treatment.

Study Design and Procedures

The plan of study calls for reexamination of the study children every two to three years. Sample A has now been examined on three separate occasions, and plans have been developed for a fourth round of examinations for this sample as well as a second round for Sample B. Each round of examinations consists of a home visit to the family and a visit to our laboratory by the study children. Between each round of testing, contact is maintained with the families by telephone at three to six month intervals. In addition, teachers' evaluations and school record data are obtained on the study children once they reach junior high school age in a separate substudy.

The initial home visit includes a two to three hour structured interview conducted with the well parent (or the more functionally intact of two schizophrenic parents or a randomly chosen parent in the normal comparison group). The interview covers the family histories on each parent and complete developmental histories on each child in the family, including older and younger siblings of the "study children." Each parent also receives a clinical interview and the MMPI. At the same time, a second interviewer obtains a shorter interview with the other parent regarding his or her perceptions of the children and also conducts an interview with each study child separately. The Bender-Gestalt and Human Figure Drawing tests are administered to the study children.

Subsequent home visits include interviews with the parents focusing on changes in the family situation and developmental changes in all of the children. The study children are administered a life events questionnaire, an adaptation of the Physical Anhedonia Scale (Chapman et al., 1976) and, if aged 14 or older, the MMPI.

Following the home visit, the study children are tested in the laboratory. (For budgetary reasons, the order of home and laboratory visits was reversed in the third round of examination for Sample A.) Table 2 lists the principal measures that have been administered in the laboratory during the various rounds of testing.

Outcome Assessments

Although the ultimate assessment of clinical outcomes in the high risk group cannot be made until all of the subjects have passed through, or at least are well into the schizophrenia risk period, interim assessments of the current clinical status of the subjects can be made. One means of evaluating current clinical status is by way of a semi-structured 30 minute psychiatric interview (developed for this program by Drs. Clarice Kestenbaum and Hector Bird) which is videotaped and subsequently rated for clinical deviance and symptomatology by three child psychiatrists who are blind with respect to a given subject's parental group. Administration of the interview during each of the laboratory visits will make available a longitudinal picture of the development of overall psychopathology and of individual symptom patterns. In addition to the videotaped interview, a standardized psychiatric interview (such as the SADS-L) with the study children is planned for later rounds of examinations.

A second type of assessment (developed by Dr. B. Cornblatt) is based on the followup data obtained from the parents on the study children. The followup data are rated on a 5-point scale (BGAS = Behavioral Global Assessment Scale) ranging from 1 = gross behavior disturbance requiring psychiatric hospitalization to 5 = above average functioning in all areas. In this evaluation system, raters are required to consider three major areas of adjustment: (a) family relationships and the subject's general emotional development; (b) peer interactions; and (c) school functioning (or functioning in a work situation if the subject has left school). (Interrater reliability among three raters for 50 subjects selected at random from Sample A was 0.94.) Raters are blind with respect to the subject's group. Because the rating system is quite simple, ratings for each subject can be readily updated if new information is obtained that would indicate a change in the subject's clinical status.

For the purpose of further analyses, the 5-point scale can be further collapsed into three categories: (a) major impairment of functioning in more than one area (scale points 1 and 2); (b) moderate difficulty in

functioning in some areas (scale point 3); (c) good functioning in family, school and peer relations (scale points 4 and 5). BGAS scores as of September 1979 for the Sample A high risk subjects (HR group) differ significantly ($p < .001$) from those for the combined comparison groups (CP and CN groups), with the HR group showing significantly more deviance in mid adolescence than the comparison subjects. These findings are based on a riddit analysis (Fleiss, 1973). This analysis takes into account the natural ordering of the BGAS clinical categories without making assumptions about scaling.

Finally, the current status of Sample A (which now ranges from 14 to 20 years of age, with an average of 17 years) can be described in terms of hospitalizations and psychiatric treatment not (yet) involving hospitalization.

Note that the percentage of CN study children with psychiatric hospitalizations and psychiatric treatment combined is 14 percent compared with only 7 percent who were categorized as "impaired" on the BGAS. This is because "psychiatric treatment" in the CN group frequently consists of guidance counseling for relatively minor problems which would lead to a BGAS of "moderate difficulty." When psychiatric treatment has been received by HR and CP subjects, however, it has been for relatively serious problems. Thus for these groups, the BGAS ratings are in complete agreement with the classification based on psychiatric hospitalization and treatment.

Discussion

At this stage of the longitudinal study, it appears that we have been able to identify subgroups of high risk children who exhibit attentional and cognitive dysfunctions at early ages, early neuromotor disturbances and evoked response patterns which differ from those of the comparison subjects or the remainder of the high risk group. The subgroup showing early attentional and cognitive deficits overlaps significantly with the subgroup of high risk subjects that is now demonstrating psychopathology in late adolescence, suggesting that deviance on such measures in childhood may be predictive of later clinical disturbance.

As not all subjects were tested with the psychophysiological measures, it is not possible to determine the degree of overlap between deviance in the evoked response patterns and other assessments of deviance at present. Nevertheless, some of the same subjects who exhibited early attentional and cognitive dysfunctions were found on later testing to be the ones who had unusual evoked response patterns showing deviations in components reflecting cognitive and attention related processes.

Because relatively few subjects displayed neuromotor dysfunctions, it is also difficult to judge how well the neuromotor data relate to the other findings of deviance. As noted above, however, the subjects with impaired neurological scores did overlap to some extent with those who had deviant attentional cognitive performance and with subjects currently classified as clinically deviant.

Whether a significant number of the subjects belonging to any of these deviant subgroups will later become schizophrenic will be seen only after further followup. In view of the fact that we expect schizophrenia to be genetically heterogenous, we would not be surprised to find more than one subgroup, with differing response characteristics emerging from the overall high risk group. Thus, there may be more than one set of characteristics that will be found to predict to schizophrenia. Moreover, our subgroups may contain some false positives who will not become schizophrenic. False positives may be very important for further scrutiny because they may be genetically predisposed individuals who are fortunate enough to have environmental buffering that enables them to avoid becoming ill. If so, such subjects might provide an important key to the understanding of environmental variables that help individuals with a "schizophrenic genotype" to avoid becoming ill.

The preliminary indication of a deficit in attention -- both as measured directly and as reflected in the evoked response components -- in some young children who are at risk for schizophrenia is especially interesting in view of (1) the many reports of impaired attention in schizophrenic patients (cf. Garmezy, 1977; Grinker & Holzman, 1973; Gunderson et al., 1974; Matthysse, 1977), (2) the strong suggestion of a correlation between attentional dysfunction and lowered Monoamine Oxidase (MAO) and Dopamine-Beta-Hydroxylase (DBH) activity levels (Buchsbaum et al., 1978), and (3) the reported reduction of MAO activity in schizophrenic patients (cf. Wyatt et al., 1979).

In our sample, the fact that attentional dysfunction is seen only in a subgroup of children of schizophrenic parents and that it does not cluster in siblings (Erlenmeyer-Kimling & Cornblatt, in press) seems to rule out the possibility that this deficit is of strictly environmental origin as a result of living with a schizophrenic parent. Thus, the current picture is compatible with the hypothesis that impairment in attention is an early indicator of vulnerability to schizophrenia -- in at least some cases -- and that it is possibly an early phenotype sign of disturbed neurochemical functioning. A planned study on MAO and DBH activity levels, with further evaluation of attentional processes, in the high risk and comparison subjects may help to establish this relationship further.

Our long term goal is preventive intervention. Intervention is as yet premature because we do not yet know how to identify specific individuals for whom intervention is indicated, what problems really characterize such children, or what form intervention should take. If, however, we can learn to identify the truly vulnerable members of the high risk group and to understand what characteristics need to be modified in them, then we will be in a position to develop rationally based strategies for intervention.

References

- Buschsbaum, M.S., Murphy, D.L., Coursey, R.D., Lake, C.R. & Zeigler, M.G. Platelet monoamine oxidase, plasma dopamine-beta-hydroxylase and attention in a "biochemical high risk" sample. Journal of Psychiatric Research, 1978, 14, 215-224.
- Chapman, L.J., Chapman, J.P. & Ravlin, M.L. Scales for physical and social anhedonia. Journal of Abnormal Psychology, 1976, 85, 374-382.
- Cornblatt, B., Marcuse, Y. & Erlenmeyer-Kimling, L. Age trends on two measures of cognition in children at risk for schizophrenia. In press.
- Endicott, J., Spitzer, R.L., Fleiss, J.L. & Cohen, J. The global assessment scale: A procedure for measuring overall severity of psychiatric disturbance. Archives of General Psychiatry, 1976, 33, 766-771.
- Erlenmeyer-Kimling, L. Studies on the offspring of two schizophrenic parents. Journal of Psychiatric Research, 1968, 6 (Supplement 1), 65-83.
- Erlenmeyer-Kimling, L. A prospective study of children at risk for schizophrenia: Methodological considerations and some preliminary findings. In R. Wirt, G. Winokur & M. Roff (Eds.), Life history research in psychopathology, Vol. 4. Minneapolis: University of Minnesota Press, 1975.
- Erlenmeyer-Kimling, L. Issues pertaining to prevention and interaction in genetic disorders affecting human behavior. In G.W. Albee & J.M. Joffe (Eds.), Primary prevention in psychopathology. Hanover, New Hampshire: University Press of New England, 1977.
- Erlenmeyer-Kimling, L. & Cornblatt, B. Attentional measures in a study of children at high risk for schizophrenia. In L. Wynne & R. Cromwell (Eds.), The nature of schizophrenia. New York: John Wiley & Sons, Inc., 1978. Also in Journal of Psychiatric Research, 1978.
- Erlenmeyer-Kimling, L. & Cornblatt, B. Children at risk for schizophrenia: Cognitive and clinical assessments at early ages. Genetics, in press.
- Fish, B. Abnormal states of consciousness and muscle tone in infants born to schizophrenic mothers. American Journal of Psychiatry, 1962, 119, 439-445.
- Fleiss, J.L. Statistical methods for rates and proportions. New York: John Wiley & Sons, 1973.
- Friedman, D., Frosch, A. & Erlenmeyer-Kimling, L. Auditory evoked potentials at high risk for schizophrenia. In M. Begleiter (Ed.), Evoked brain potentials and behavior. New York: Plenum, 1979a.
- Friedman, D., Vaughan, H.G., Jr. & Erlenmeyer-Kimling, L. Stimulus and response related components of the late positive complex in visual discrimination tasks. Electroencephalography and Clinical Neurophysiology, 1978, 45, 319-330.
- Friedman, D., Vaughan, H.G., Jr. & Erlenmeyer-Kimling, L. Event-related potential investigations in children at high risk for schizophrenia. In D. Lehmann & E. Callaway (Eds.), Human evoked potentials: Applications and problems. New York: Plenum, 1979b.
- Friedman, D., Vaughan, H.G., Jr. & Erlenmeyer-Kimling, L. The late positive complex to unpredictable auditory events in children at high risk for schizophrenia. Presented at the annual meeting of the Society for Psychophysiological Research, Cincinnati, Ohio, October 1979c.
- Garmezy, N. Children at risk: The search for the antecedents of schizophrenia. Part II: Ongoing research programs, issues and intervention. Schizophrenia Bulletin, 1974, 9, 55-125.
- Garmezy, N. The psychology and psychopathology of attention. Schizophrenia Bulletin, 1977, 3, 360-369.
- Grinker, R.R. & Holzman, P.S. Schizophrenia pathology in young adults. Archives of General Psychiatry, 1973, 28, 168-175.
- Gunderson, J.G., Autry, J.H., Mosher, L.R. & Buchsbaum, S. Special report: Schizophrenia. Schizophrenia Bulletin, 1974, 15.
- Kallmann, F.J., Galek, A., Hurler, M. & Erlenmeyer-Kimling, L. The developmental aspects of children with two schizophrenic parents. In D. Solomon & B.C. Glueck, Jr. (Eds.), Recent research in schizophrenia. Psychiatric Research Report 10. Washington D.C.: American Psychiatric Association, 1964.
- Koppitz, E.M. The Bender gestalt test for young children. New York: Grune & Stratton, 1963.
- Koppitz, E.M. Psychological evaluation of children's human figure drawings. New York: Grune & Stratton, 1968.
- Koppitz, E.M. Visual aural digit span test performance of boys with emotional and learning problems. Journal of Clinical Psychology, 1973, 29, 463-466.

Marcus, J. Cerebral functioning in offspring of schizophrenics: A possible genetic factor. International Journal of Mental Health, 1974, 3, 57-73.

Marcuse, Y., Cornblatt, B. & Erlenmeyer-Kimling, L. Evidence of "soft" neurological dysfunctions in children at risk for schizophrenia. In preparation.

Matthysse, S. The biology of attention. Schizophrenia Bulletin, 1977, 3, 370-372.

Mednick, S.A. & Schulsinger, F. Some premorbid characteristics related to breakdown in children with schizophrenic mothers. In D. Rosenthal & S. Kety (Eds.), The transmission of schizophrenia. Oxford: Pergamon Press, 1968.

Pass, H.L., Klorman, R., Slazman, L.F., Klein, R.H. & Kaskev, G.B. The late positive component of the evoked response in acute schizophrenics during a test of sustained attention. Society for Psychophysiology Research Abstracts, 1977.

Pearson, J.S. & Kley, I.B. On the application of genetic expectancies as age-specific base rates in the study of human behavior disorders. Psychological Bulletin, 1957, 54, 406-420.

Rieder, R.O. Children at risk. In L. Bellack (Ed.), The schizophrenia syndrome, 2nd edition. In press.

Rutschmann, J., Cornblatt, B. & Erlenmeyer-Kimling, L. Sustained attention in children at risk for schizophrenia. Archives of General Psychiatry, 1977, 34, 571-575.

Sobel, D.E. Children of schizophrenic parents: Preliminary observations on early development. American Journal of Psychiatry, 1961, 118, 512-519.

Spitzer, R.L., Endicott, J. & Robins, E. Research diagnostic criteria (RDC) for a selected group of functional disorders, Third edition, 1977.

Spitzer, R.L. & Endicott, J. Schedule for affective disorders and schizophrenia -- Life-time version (SADS-L), Third edition, June 1977.

Wyatt, R.J., Potkin, S.G. & Murphy, D.L. Platelet monoamine oxidase activity in schizophrenia: A review of the data. American Journal of Psychiatry, 1979, 136 (4A), 377-386.

Table 1
Samples in the New York High Risk Study

Group	Sample A		Sample B	
	study children	siblings	study children	siblings
High risk (HR)*	80	53	44	34
Psychiatric comparison (CP)	25	20	40	19
Normal comparison (CN)	100	49	66	33
Total	205	122	150	86

*In Sample A, 13 high risk study children are offspring of two schizophrenic parents, and in Sample B, 8. Remaining high risk children are offspring of a schizophrenic mother or father.

Table 2
Laboratory Procedures

Procedure	Sample A Round			Sample B Round
	1	2	3	1
Anthropometric measures & photographs	X	X	X	X
Attentional-cognitive measures				
Wechsler Intelligence scales				
CPT (Carousel presented card version)	X		X	X
ATS (Attention span)	X	X		
CPT1 (Computerized vigilance task)	X			
CPT2 (Computerized discrimination task)		X	X	X
IOT (Information overload test)		X	X	X
STM-Lag (Short term memory lag test)			X	X
VADS (Visual aural design test)			X	
Neurological, motoric and laterality				X
Neurological examination				
Lincoln-Oseretsky Test of Motor Impairment	X			X
Purdue Pegboard	X			X
Dichotic Listening Test			X	X
Psychophysiology				
Conditioning paradigm, GSR				
Resting EEG	X			
Auditory event-related potentials		X		
-to repetitive stimulation		X		
-to task-relevant stimulation		X		
Visual evoked potentials to CPT1 and CPT2			X	X
Auditory thresholds (absolute and aversion)			X	X
-heart rate and skin conductance		X	X	X
Magnitude estimation and psychophysical judgments		X	X	X
Stimulus level intensity gradients			X	
Clinical and Social Measures				X
Structured videotaped interview				
Videotaped psychiatric interview	X			
Friendship and Intimacy interview		X	X	X
			X	X

-749-

The Consortium for Longitudinal Studies:
A Follow-up of Participants in Early Childhood Programs*

Harry W. Murray, Jacqueline M. Royce, Irving Lazar
Department of Human Service Studies, Cornell University

Richard B. Darlington
Department of Psychology, Cornell University

*Many people over many years contributed to this research. We are grateful to Consortium members E. Kuno Beller, Cynthia and Martin Deutsch, Ira Gordon, Susan Gray, Merle Karnes, Phyllis Levenstein, Louise Miller, Francis Palmer, David Weikart, Myron Woolman, and Edward Zigler for their vision and their commitment to a cooperative effort. We thank Dr. Edith Grotberg and Dr. Bernard Brown of the Administration for Children, Youth and Families for their encouragement and their dedication to children.

Our present and former Cornell colleagues Ann Bell, Ruth Hubbell, Daniel Koretz, Marilyn Rosche, Ann Snipper, Reuben Snipper, Judy Vopava, and Marjorie Wikerd made important contributions to this work. We also thank the supervisors and staff at each project site for sharing in this effort. We have been supported by DHEW (OHDS) grants Nos. 90-C-1311 and 18-76-07843 and the William and Flora Hewlett Foundation.

The Consortium for Longitudinal Studies:
A Follow-up of Participants in Early Childhood Programs*

Harry W. Murray
Jacqueline M. Royce
Irving Lazar
Department of Human Service Studies
Cornell University

Richard B. Darlington
Department of Psychology
Cornell University

The Consortium for Longitudinal Studies is a cooperative effort of over a dozen investigators endeavoring to assess the long-term effects of early education on the lives of children from low-income families. The principal investigators, who conducted 14 longitudinal research and demonstration projects in the 1960s, agreed to pool their original data and to collect common follow-up measures for joint analysis in 1976 and 1979. This chapter describes the Consortium, the design of the collaborative study and the characteristics of the data base.

Description of Consortium Early Child Development Studies

The early education programs in the Consortium studies were located in urban and rural sites in the Northeast, Southeast, and Midwest (see Table 1). The programs were active from 1962 to 1972; only Levenstein's Verbal Interaction Project, Weikart's Perry Preschool Program, and Karnes' GOAL program are in operation at the present time. Depending on the theoretical interest of the investigator, curricula included programs based on the Bank Street

child development model, on Montessori methods, on Piagetian theory, on the Bereiter-Engelmann method and others. For full description of Consortium programs, see Consortium (1977); Consortium (1978); Consortium (in press); Beller (1974); Deutsch, Taleporos, and Victor (1974); Gordon, Guinagh, and Jester (1977); Gray (1974); Karnes, Zehrback, and Teska (1977); Levenstein (1977); Miller and Dyer (1975); Palmer and Siegel (1977); Seitz, Apfel, and Efron (1978); Weikart, Bond, and McNeil (1978); and Woolman (1971).

The Consortium studies can be categorized as center-based, home-based, or combination home visits/center programs. Center-based studies typically provided nursery school programs for 3- or 4-year-olds with varying degrees of structure in the program curriculum. Instruction generally took place in small groups but in some cases was on a one-to-one basis.

Home-based studies directed their educational efforts primarily toward the parent, usually the mother, as the major instrument of change and influence in the child's life. Activities, toys, and games were brought to the family home by a parent educator or home visitor, who taught the mother to use the activities and to promote her child's development through parent-child interaction. These programs served infants and toddlers.

The third group of studies combined these approaches, providing a center-based nursery school program along with a periodic home visit which involved both parent and child. In some programs the emphasis during the home visit was closely related to the center programs; in others, it had a more general content.

Description of the Sample

Since all projects were originally conducted independently, each had different sample selection criteria and procedures for assignment to program or control groups. Table 2 summarizes sample selection criteria. The criteria were similar for most projects, since projects generally selected children from low-income families. Selection criteria differentiated among projects as follows. The Palmer project selected half lower-class and half middle-class black males. Deutsch, Gray-ETP, Palmer, and Weikart-Perry selected only black subjects. Karnes specified a minimum IQ of 70, while the Weikart-Perry and Curriculum Demonstration projects specified a maximum IQ score of 84. All projects except the three in the New York City area were open to non-English-speaking children, but only the Woolman project actually served a significant number of such children. Most of these criteria were reflected in the demographic characteristics of the original samples.

The projects also differed in the manner in which they assigned children to program or control groups. The importance of adequate procedures for assignment to program or control cannot be overemphasized. Therefore, the sample selection procedures for each project were reviewed and classified (Consortium, 1977). Seven projects were classified as experimental--Deutsch, Gordon, both Gray projects, Palmer, Weikart-Perry, and Weikart-Carnegie.

Because they were originally designed to answer questions concerning mothers other than a simple program/control comparison, five were classified as quasi-experimental for our purposes--Beller, Levenstein, Miller, Woolman, and Zigler. Beller, for example, compared three groups on age of entry to schooling (nursery, kindergarten, and first grade). Miller's project had an experimental design for comparing different curriculum groups, but a quasi-experimental design for comparing program to control. Two projects were concerned solely with curricula comparisons and had no control groups--Karnes and Weikart-Curriculum Demonstration.

The Consortium data base is unusual: it has both cross-sectional and longitudinal dimensions; furthermore, because it is composed of a number of research and development projects, it constitutes a data base permitting meta-analysis. Its complexities become evident as soon as one attempts to "describe the sample." Due to attrition (both in the form of inability to locate subjects and of decisions not to follow certain groups or projects because of design or monetary constraints) numerous "samples" could be delineated. Two such "samples" are described here: the original sample (on whom initial data was collected) and the 1978 follow-up sample (on whom any of the common Consortium instruments were collected).

The original sample consisted of 3,177 children from 13 projects. For the median project, 67.3 percent of the children were enrolled in a preschool program, 47.4 percent of the children were female, and 90 percent were black. At the time of entry into the program, the medians on background variables were: mother's education, 10.3 years; Hollingshead SES scores in the lowest category; three siblings; and Stanford-Binet pretest IQ score of 90.8. Individual project differences are minor, except where accounted for by sample selection criteria mentioned previously.

The 1976 follow-up sample was composed of 2,008 children from 11 projects. The median project had 64.9 percent children enrolled in the program, 94.8 percent of them were black, and 50.0 percent were female. The medians on the follow-up sample for background variables measured at the time of entry into the projects were: mother's education of 10.4 years; SES (Hollingshead ISP) of 64.0; 3.2 siblings; and pretest Stanford-Binet IQ score of 92.1. Thus, on a macro level, the 1976 follow-up sample was very similar to the original sample despite the attrition and changes in groups and projects included in the follow-up.

The differences between the original and 1976 follow-up samples can be traced to several sources: (a) the exclusion of certain projects from the follow-up (Gray-FOHV, Weikart-Curriculum Demonstration, Weikart-Carnegie); (b) the exclusion of certain groups or waves in a project from the follow-up (e.g., Deutsch groups three through eight); (c) the addition to the follow-up sample of one project (Woolman) that was not contained in the original data base; and (d) the attrition of individuals within the projects which were followed-up.

The samples used for analysis were, of course, subsets of the total

CONTINUED

9 OF 10

samples. Analysis samples could differ from total samples in three ways.

1. Whole projects could be eliminated from the analysis because (a) they did not collect the necessary dependent variable for the analysis (e.g., Beller did not collect WISC scores); (b) the design of the project was inappropriate for the analysis (e.g., the Karnes and Weikart-Curriculum Demonstration projects had no control groups and, hence, were eliminated from the program-control comparisons); or (c) attrition analyses indicated problems which endangered validity (this occurred only for the Deutsch project).

2. Missing data could eliminate some subjects within a project which was included in the analysis.

3. In some cases, projects had two or more comparison groups. In order to make the comparison as closely experimental as possible, we eliminated the control group which we judged to be less valid for program/control analyses (e.g., "after-only" subjects selected in first grade). This procedure reduced the sample size and, hence, the power. Beller, Gray-ETP, and Levenstein each had one control group eliminated from analyses for this reason. The Consortium data base does contain these groups, however.

The Consortium research is a four-wave panel study which consists of: (a) the initial data collected independently by the investigators prior to the formation of the Consortium; (b) postprogram follow-ups collected independently by the investigators before 1976; (c) the 1976 collaborative Consortium follow-up; and (d) the 1979-1980 collaborative Consortium follow-up now in progress.

Wave 1 data were collected when subjects entered the various preschool programs, over the years 1962-1972. Each project had its own research design, generally comparing one program group or a number of program groups to a control group, or comparing among a number of program groups. Except for cross-fertilization through professional communications, each investigator independently planned and implemented his or her own research design and data collection effort. Later programs were influenced by earlier programs in some cases. Wave 1 data were obtained primarily from the intake or subject-selection interviews that each investigator had designed for his/her own purposes at that time. The data included primarily demographic and family background data and preprogram scores on a variety of measures, generally including an intelligence test.

Wave 2 data include the early follow-ups conducted independently by each investigator prior to 1976. Depending on research design and availability of funding, investigators conducted an immediate post-program assessment and a number of follow-up assessments until subjects were 8, 9, or 10 years old. The number of annual follow-ups ranged from two to seven.

Waves 1 and 2 raw data were sent to Cornell in early 1976 to form

the Consortium data bank. Data sent on computer cards or raw data sheets were checked for accuracy and internal consistency. Cross-tabulations of type of data by project site revealed variables which were collected by enough investigators to permit common analyses. After consultation with each project site, IQ test scores and selected demographic data were recoded into a common format and placed on computer file.

Wave 3 data were collected collaboratively in 1976-1977 when the sample ranged in age from 9 to 19 years old. The Wave 3 data included the 1976 Consortium Parent and Youth Interviews, the Wechsler Intelligence Scales (Revised), achievement test scores primarily recorded from school records, and school career information recorded by field workers on Consortium School Record Forms.

Some investigators had maintained contact with their sample over the years through activities such as home visits, Christmas cards, and post office forwarding addresses. Other investigators, however, who had not attempted to reach their sample for up to 10 years, experienced difficulty locating subjects and school records. They examined name and address listings in local schools or Boards of Education. Because of school reorganization or other factors, some investigators were unable to find subjects on school listings and had to rely on other sources, including retired school teachers living in the community, former high school athletes, relatives, neighbors, and subjects already interviewed.

In order to assure consistency of data collection across sites, the group at Cornell conducted training sessions for field site supervisors and provided video tapes to demonstrate interviewing techniques. Most items were precoded, and all instruments were printed on two-part carbonless paper; field sites retained one copy and mailed the original to Cornell where after the fact coding of open-ended responses, editing, and keypunching were done. Each project site received a duplicate set of their data on computer cards.

The Wave 4 follow-up of 1979-1980 is currently underway. All participating investigators are collecting detailed information on subjects' school careers. The information is recorded on Consortium School Forms at each field site. Investigators with subjects over 18 years old are administering the 1979 Consortium Interview, either by telephone or in person. Some project sites have collected achievement test scores, either individually-administered or collected from school records. As in the 1976 follow-up, Cornell staff coordinate data collection at the field sites, primarily by telephone, and then code open-ended responses, edit, and keypunch the data.

Measures

A wide variety of status and process variables were collected by individual investigators over the four waves of data collection. In the first two data collection waves before the formation of the Consortium (1962 to 1975), investigators collected diverse pre-program child and family measures (Wave 1) and then numerous and varied post-program outcome measures over several years (Wave 2). In the Consortium follow-ups of 1976 and 1979 (Waves 3 and 4), a core of common measures was collected. The measures included in the Consortium data bank are described below.

For Wave 1, the independently-collected initial data, all investigators except Beller collected pre-program background data such as education and occupation of the mother and father, number of siblings, and father presence in the home as well as sex and ethnicity of the child, mother's age at childbirth, and other variables. Adding certain variables to the Consortium file was problematic. Variables such as father presence and number of siblings, for example, did not have a standard definition across projects. In some cases, data from a project for a certain variable could not be transformed into a common Consortium format, but usually variables were retained despite some problems in comparability.

In addition to Wave 1 demographic variables, most projects collected a battery of pretests such as Cattell, Stanford-Binet, Peabody Picture Vocabulary Test (PPVT), Bayley, and a variety of other standardized or project-designed measures. Of these measures, only the Stanford-Binet and PPVT pretests were collected by a sufficient number of projects to justify inclusion in the common data base.

The second wave of testing, also conducted independently by the investigators, consisted primarily of standardized IQ achievement tests and project-developed tests. The majority of projects collected Stanford-Binet and PPVT IQ scores at the various Wave 2 follow-up testing periods. All Binet and PPVT scores are included in the Consortium data file. Other tests were collected by several investigators but not included in the Consortium data base. For example, standardized tests collected by more than one project were: ITPA (4 projects); WISC (4 projects); Leiter (2 projects); Bayley (2 projects); Preschool Inventory (2 projects); California (4 projects); Metropolitan (5 projects) and Wide Range (2 projects) achievement tests, and Metropolitan Readiness Test (3 projects). Individual investigators also collected a large number of tests that were not replicated in any other project (e.g., socio-emotional and perceptual measures).

The decisions on the choice of common instruments to be used in the Wave 3 follow-up study of 1976 evolved from lengthy discussion of alternatives by the Consortium members over a 6 month period. Because of financial restrictions and the high cost of locating subjects who had not been interviewed or tested for many years, it was necessary to choose a limited battery of measurements. The Consortium members agreed to administer the Consortium-developed Parent and Youth Interviews and appropriate Weschler

Intelligence Scale and to obtain data from school records and achievement tests administered by the schools. In addition, four projects provided detailed school information (e.g., course marks, tracking, attendance) that they recorded on Consortium Long School Record Forms.

The 1976 Consortium Parent Interview was initially developed from interviews used by individual investigators. It was designed to obtain comprehensive information on household composition; family members; housing and mobility data; parental attitudes toward, aspirations for, and evaluations of their child; information on the child's medical history; the parent's current relationship with the child; and parental assessment of the preschool program.

The 1976 Youth Interview also drew upon interviews developed previously by the principal investigators. It included information on the child's status in school; his or her educational and occupational aspirations; leisure time activities and interests; employment status; and integration into his or her peer group, family, and the larger community.

The 1976 Consortium School Record Form provided data on the subjects' current educational status (graduated from high school, dropped out, currently in school); whether the subject was ever placed in special education classes (and type of class); how often he or she was retained in grade or accelerated; and the code number and type of school attended for 4 specified years.

Diverse achievement tests were administered by the many school systems including the California Achievement Test, the Comprehensive Test of Basic Skills, the Iowa Test of Basic Skills, and the Metropolitan Achievement Test. Depending on the information recorded in school records, field workers recorded raw scores, grade equivalent scores, or standard scores on 1976 Consortium Achievement Test Forms for up to four different academic testing periods.

For each project, the number of cases with Wave 3 follow-up data by instrument is shown in Table 3. The number of cases in the 1976 follow-up with data on at least one instrument was 2,008; the last column on the right indicates the number of cases in the follow-up for each project.

The Wave 4 data collection of 1979-1980 utilizes two Consortium-designed instruments: the Consortium 1979 Youth Interview and the Consortium 1979 School Data Form. Six sites have interviewed or will interview their subjects (Beller, Deutsch, Gray, Karnes, Miller, and Weikart). For some sites, the 1979 Consortium Interview forms part of a longer interview designed by individual investigators. The Consortium Interview focuses on determining subjects' status in the areas of education, employment, and family formation. Additional items include job satisfaction, attitudes toward work, educational and vocational aspirations and expectations, sources of income, desired and expected family size, stresses and supports during childhood, self-evaluation, self-esteem, and life priorities.

The 1979 School Data Form consists of two sections: (a) information for each academic year including attendance, special class placements, grade retentions, and participation in supplemental tutoring or enrichment programs; honors; course marks for 4 years; class rank; suspensions, expulsions, and dropout history and reasons; and (b) detailed information on "problem" children, including diagnoses and reasons for referral to special services and critical home and school events.

As of 1979, measures included in the Consortium data base are: demographic variables and Stanford-Binet and PPVT pretests for Wave 1; Stanford-Binet and PPVT post-program IQ scores for Wave 2; and all Consortium instruments for the Wave 3 follow-up (i.e., 1976 Parent and Youth Interviews, achievement tests for 4 years; 1976 WISC scores, and school record information). We expect to add the Wave 4 data to the Consortium file in early 1980 (i.e., 1979-1980 Consortium interviews, school data, and achievement scores).

Outcomes Studied

We have examined four main outcomes in our study of the effect of Consortium childhood programs: school competence, measured as special education placements and retention in grade; intelligence test scores, measured at several time periods; standardized achievement test scores; and certain non-cognitive outcomes, particularly achievement-orientation.

Improved school competence of children from low-income families was one goal of the Consortium infant and preschool programs, just as it is for Head Start programs nationwide. Progressing through school with age-mates instead of being assigned to special education classes or retained in grade is a direct measure of school competence that is visible and ecologically valid. The child who requires assignment to special education classes or grade retention has not met the requirements of his or her school. Success or failure in the elementary and secondary school years is an important variable in the status attainment process. It has critical implications for an individual's ultimate educational, occupational, and economic attainment. Furthermore, the labeling and ability streaming (tracking) associated with school failure have important developmental implications.

The second outcome measure we studied was intelligence. Intelligence tests have been a focus of controversy ever since they were devised as a method of quantifying the nebulous concept of intelligence. They have been accused of being culturally biased, of reflecting "achievement" and motivational variables as well as cognitive ability, and of being used for social control. Empirically, however, intelligence test scores remain the best available predictor of school performance and they are commonly used in educational evaluations. It is only logical to expect a correlation between IQ scores and school performance since scoring well on an

intelligence test and performing well in school both require similar behaviors, including spontaneous verbalization, persistence at a task, ability to follow instructions, and ability to adapt to structured surroundings. Thus, intelligence test scores are used in this study as an indicator of expected performance in schools as they currently exist in our society.

Achievement tests administered by the schools were also used as outcome measures. Scores on standardized achievement tests provide information on pupil learning that is designed to be independent of local curricula and individual teachers. Like intelligence tests, achievement tests have been a subject of controversy--whether they are primarily measures of ability rather than learning, the utility of norm-referenced versus criterion-referenced tests, etc. Nevertheless, we utilized achievement tests as the most comparable measure of children's mastery of specific academic content across projects.

A number of non-cognitive outcomes (achievement-orientation, maternal and child aspirations, and the child's self-assessment of school performance) were investigated using questions on the 1976 parent and child interviews. We have regarded the analyses of noncognitive outcomes as exploratory because of the difficulties in measuring attitudes and values.

Of the outcomes considered thus far, only intelligence test scores were recorded at several points in time and are feasible to analyze over time. Achievement test scores were also collected at several time periods, but missing data and noncomparability of the various tests would make analysis over time problematic. The Wave 4 data collection of 1979-1980 will provide repeated measures of placement in special education and of retention for each year. We expect these analyses to be methodologically challenging.

Summary

Numerous analyses have been performed, both by the Consortium staff using the Consortium data base and by the individual projects using their own data. Analyses on the Consortium data base have focused on two main questions thus far:

1. Does participation in early childhood program have long-term positive effects on various outcome measures?
2. Do certain types of children benefit more than others from the preschool experience?

The Consortium for Longitudinal Studies has provided evidence that early childhood programs can have an impact on the academic careers of low-income children. This effect can be measured both in terms of

performance on standardized tests and in terms of actually meeting the requirements of normal progression as defined by the schools. The four waves of data collection continue to provide an opportunity to evaluate the effect of preschool on various outcomes as the sample reaches adulthood. Further, the Consortium data can permit the testing of models which explore the interrelationship of variables over time. Research questions can be addressed in areas of cognitive and social development, family background and process effects, family formation, and educational and occupational attitudes and attainments.

References

- Beller, E. K. Impact of early education on disadvantaged children. In S. Ryan (Ed.), A report on longitudinal evaluations of preschool programs, Vol. 1: 15-48. DHEW Publication No. (OHD) 74-24, 1974.
- Campbell, D. T. Temporal changes in treatment-effect correlation: A quasi-experimental model for institutional records and longitudinal studies. In V. Glass (Ed.), Proceedings of the 1970 Invitational Conference on Testing Problems. Princeton, N. J.: Educational Testing Service, 1971.
- Consortium for Longitudinal Studies: Lasting effects of early education. Monographs of the society for research in child development, (in press).
- Consortium for Longitudinal Studies. Lasting effects after preschool. DHEW Publication No. (OHDS) 79-30178, October, 1978.
- Consortium on Developmental Continuity. The persistence of preschool effects. DHEW Publication No. (OHDS) 78-30130, September 1977.
- Darlington, R. B. Multiple comparisons and related problems. Workshop presented at the Annual Meetings of the American Psychological Association, Toronto, August, 1978.
- Deutsch, M., Taleporos, E., & Victor, J. A brief synopsis of an initial enrichment program in early childhood. In S. Ryan (Ed.), A report on longitudinal evaluations of preschool programs, Vol. 1. Washington, D. C.: Office of Child Development, DHEW Publication No. (OHD) 74-24, 1974.
- Glass, V. Primary, secondary, and meta-analysis of research. Educational Researcher, 1976, 5, 3-8.
- Gordon, I. J., Guinagh, B., & Jester, R. E. The Florida parent education infant and toddler programs. In M. C. Day & R. K. Parker (Eds.), The preschool in action (2nd Edition). Boston: Allyn & Bacon, 1977.
- Gray, S. W. Children from three to ten: The early training project. In S. Ryan (Ed.), A report on longitudinal evaluations of preschool programs, Vol. 1. Washington, D. C.: Office of Child Development, DHEW Publication No. (OHD) 74-24, 1974.
- Hilton, L., & Patrick, C. Cross-sectional versus longitudinal data: An empirical comparison of mean differences in academic growth. Journal of Educational Measurement, 1970, 7, 15-24.
- Karnes, M. B., Zehrback, R. R., & Teska, J. A. Conceptualization of the GOAL (game-oriented activities for learning) curriculum. In M. C. Day & R. K. Parker (Eds.), The preschool in action: Exploring early childhood programs (2nd Edition). Boston: Allyn & Bacon, 1977.

Levenstein, P. The mother-child home program. In M. C. Day & R. C. Parker (Eds.), The preschool in action: Exploring early childhood programs. (2nd Edition). Boston: Allyn & Bacon, 1977.

Miller, L., & Dyer, J. L. Four preschool programs: Their dimensions and effects. Monographs of the society for research in child development, 1975, 40 (5-6), Serial No. 162.

Mosteller, F., & Bush, R. R. Selected quantitative techniques. In G. Lindzey (Ed.), Handbook of social psychology, Vol. 1. Reading, Mass.: Addison-Wesley, 1954.

Palmer, F. H. & Siegel, R. J. Minimal intervention at ages two and three and subsequent intellectual changes. In M. C. Day & R. K. Parker (Eds.), The preschool in action: Exploring early childhood programs (2nd Edition). Boston: Allyn & Bacon, 1977.

Seitz, V., Apfel, N., & Efron, C. Long-term effects of early intervention: The New Haven project. In B. Brown (Ed.), Found: Long-term gains from early intervention. Boulder, Colorado: Westview Press, 1978.

Thorndike, R. L. Stanford-Binet intelligence scale: 1972 norms tables. Boston: Houghton Mifflin, 1973.

Weikart, D. P., Bond, J. T., & McNeil, J. T. The Ypsilanti Perry preschool project: Preschool years and longitudinal results. Monographs of High/Scope educational research foundation, No. 3, 1978.

Woolman, M. Learning for cognition: The micro-social learning system. Report to the New Jersey State Department of Education, 1971.

Table 1

Characteristics of Early Education Programs and Ages of Subjects for Each Data Set

Principal Investigator	Early Education Program	Location	Type of Delivery System	Subject Birth Year	Age at Entry to Program	Length of Program (years)	Years of Program
Beller	The Philadelphia Project	Philadelphia, Pennsylvania	Center	1959	4 years	1 year	1963-64
Deutsch	Institute for Developmental Studies	New York, New York	Center	1958-61	4 years	5 years	1963-71
Gordon	The Parent Education Program	Gainesville, Florida	Home	1966-67	3 mos. to 2 years	3 years	1966-70
Gray	The Early Training Project (ETP)	Murfreesboro, Tennessee	Center/Home	1958	3.8 or 4.8 years	14 mos. or 26 mos.	1962-65
Gray	The Family-Oriented Home Visitor Program	Nashville, Tennessee	Home	1971	birth to 1 year	1 or 2 years	1970-73
Karnes	Curriculum Comparison Study	Champaign-Urbana, Illinois	Center	1961-63	4 years	1 year	1965-66 (2 waves)
Levenstein	The Mother-Child Home Program	Glen Cove, Manhasset & Freeport, Long Island, New York	Home	1964-68	2 years & 3 years	1-1 1/2 years	1967-72
Miller	Experimental Variation of Head Start Curricula	Louisville, Kentucky	Center & Center/Home	1964	4 years	1 year	1968-69

Table 1 (Cont.)

Principal Investigator	Early Education Program	Location	Type of Delivery System	Subject Birth Year	Age at Entry to Program	Length of Program (years)	Years of Program
Palmer	Harlem Training Project	New York, New York	Center	1964	2 or 3 years	1 or 2 years	1966-68
Weikart	Perry Preschool Project	Ypsilanti, Michigan	Center/Home	1958-62	3 years (1st wave) 4 years	2 years (1st wave) 1 year	1962-67 (5 waves)
Weikart	Curriculum Demonstration Project	Ypsilanti, Michigan	Center/Home	1964-65	3 years	2 years	1966-70
Weikart	Carnegie Infant Program	Ypsilanti, Michigan	Home	1968	3 mos. to 2 years	16 mos.	1967-70
Woolman	Micro-Social Learning System	Vineland, New Jersey	Center	1966-68	4 to 5 years	1 to 4 years	1969-73
Zigler	New Haven Follow-Through Study	New Haven, Connecticut	Center	1962-64	5 years	4 years	1967-71

Table 2

Sample Selection Criteria Used by Projects

Project	Age (in years)	Sex	Race	IQ Score	Language	SES
Beller	4,5,6					low*
Deutsch	4-8		Black		English	low
Gordon	birth,1,2					low
Gray-ETP	4,5		Black			low
Gray-FOHV	birth,1					low
Karnes	4			≥ 70		low*
Levenstein	2,3				English (early waves)	low
Miller	4					low
Palmer	2,3	Male	Black		English	1/2 low 1/2 middle
Weikart-Perry	3,4		Black	< 85		low
Weikart-CD	3			< 85		low
Weikart- Carnegie	birth,1					low
Woolman	4,5					low*
Zigler	5					low

* Indicates projects which did not have a formal specification of "low income." For other projects, low SES was specified on a measurable scale (e.g., Hollingshead ISP).

Table 3

Number of Cases with 1976 Follow-up Data by Instrument (Wave 3)

Project	Follow-up Age in 1977 (years)	Parent Interviews	Youth Interviews	School Record Form	Ach't. Tests	Long School Record Form	IQ Scores ^a	1976 Follow-up Sample ^b
Beller	18	112	116	119	115	115	--	126
Deutsch	16-19	73	88	54	44	---	88	97
Gordon	10-11	107	106	107	103	---	90	107
Gray-ETP	19	72	69	75	72	74	72	77
Karnes	14-16	165	156	153	143	105	112	168
Levenstein	7-13	98	75	154	115	---	75	186
Miller	13	141	141	139	134	---	141	141
Palmer	13	143	144	223	197	---	132	228
Weikart-Perry	15-19	106	104	123	96	---	110	123
Woolman	9-11	54	97	611	349	200	95	611
Zigler	13-15	---	141	144	142	---	143	144
Totals		1071	1237 ^c	1902	1510	494	1058	2008

^a Scores include WISC-R scores for Gordon, Gray, Karnes, Miller, Palmer, Woolman; WISC scores for Levenstein, Weikart; WAIS scores for Deutsch; and PPVT scores for Zigler.

^b Follow-up sample *n* is the number of cases with data on at least one instrument.

^c Total includes 141 Zigler cases who did not have Consortium Interview.

Two Decades of Research: The High/Scope
Longitudinal Preschool Evaluations*

David P. Weikart and Lawrence J. Schweinhart
High/Scope Educational Research Foundation

* The research reported herein was supported in part by the U.S. Office of Education (1964-1968), the Spencer Foundation (1971-1974), the Carnegie Corporation of New York (1975-1978), and the Administration for Children, Youth and Families (DHEW) through the Education Commission of the States (1977-1979).

Two Decades of Research: The High/Scope
Longitudinal Preschool Evaluations*

David P. Weikart and Lawrence J. Schweinhart
High/Scope Educational Research Foundation

The Context

In the 1960s the idea became prevalent that early life experiences are crucial in the formation of intelligence, or at least one's ability to perform well in school. A first wave of findings on the cognitive benefits of early intervention was very encouraging. Several studies of local programs (Kirk, 1958; Deutsch et al., 1967; Weikart, 1967; Klaus & Gray, 1968) found that preschool education could increase IQ. Since IQ was meant to predict school performance, it appeared that the hypothesis of cognitive benefits of early intervention was being confirmed.

This optimism was shattered by events of the next few years. The local experimental studies began to find that the initial IQ gains disappeared a year or two after the end of the preschool programs. The first national evaluation of Project Head Start (Westinghouse Learning Corporation/Ohio University, 1969), despite methodological shortcomings, made the unmistakable announcement that the project had almost no positive impact on cognitive performance. Zigler and Butterfield (1968) called into question even the ephemeral gains in IQ, carrying out studies which suggested that preschool-induced IQ gains represented motivational factors rather than improvement in formal cognitive processes. A widespread mood of pessimism about compensatory preschool education was developing.

A decade later, however, new evidence is beginning to appear, evidence which seems to reaffirm the lasting effects of early childhood experience, while at the same time taking into account previous results.

Design

In 1962, Weikart and his associates began an experimental study of

preschool intervention called the Ypsilanti Perry Preschool Project. The Perry Study was designed to determine how preschool education can benefit economically disadvantaged children diagnosed at age 3 as educable mentally retarded. To address this question, a sample of 123 such children in Ypsilanti's Perry Elementary School attendance area were placed, by a random assigned procedure, in either of two groups: an experimental group of 58 children who attended the Perry Preschool and a control group of 65 children who did not attend preschool.

In 1967, a second study was begun, called the Ypsilanti Preschool Curriculum Demonstration Project. The Curriculum Study was designed to ask how each of three different preschool curriculum models could benefit disadvantaged children diagnosed at age 3 as educable mentally retarded. To address this question, a sample of 92 such children in Ypsilanti were placed, by a random assignment procedure, in one of three groups: 31 children who participated in High/Scope's Cognitively Oriented Curriculum; 31 children who participated in a behaviorist curriculum; and 30 children who took part in a traditional nursery school curriculum. The subjects from these two studies comprise the longitudinal sample for this research, 215 young people in all.

Sample Selection

Each September the names of all families with three-year-old children were drawn from the appropriate school census (Perry or Ypsilanti). The socioeconomic status of these families was determined from the parents' education, occupational level of the head of the household, and household density (rooms per person). If the SES score of the family was below a specified level, the three-year-old was given the Stanford-Binet Intelligence Scale. Children with low IQs (60 to 90), but no evidence of organic impairment, became a part of the study sample. Reflecting the ethnic makeup of the Perry School attendance area, all the children in the Perry Preschool sample were black. In the Curriculum Study sample, there were 57 black and 35 white children. Thus, the sample consisted of young children of low measured IQ from families of low SES.

A replication element was built into both studies by enlarging the sample annually with waves of children selected by the same criteria and assigned to groups by the same procedures. The Perry Study began in 1962 with the selection of a group of four-year-olds (Wave 0) and a group of three-year-olds (Wave 1). The Perry Study sample was completed over the next three years by the annual selection of additional groups of three-year-olds (Waves 2, 3 and 4). Without breaking pace, the Curriculum Study began with the selection of a group of three-year-olds (Wave 5) in 1966 and annually thereafter for the next three years (Waves 6, 7, and 8). The complete longitudinal sample consists of nine waves of children, each a year older than the next oldest group.

Group Assignment and Comparability

Each year children were randomly assigned to some treatment group. Children were ranked by their pretest IQs, then sorted into two groups in the Perry Study and three groups in the Curriculum Study. The arrangement and size of these groups, by waves, is diagrammed in Table 1. Children with similar scores were arbitrarily exchanged to equate the sex ratio and average SES score of each group. Siblings were assigned to the same group as their older siblings in the studies, in order to maintain the independence of the groups. In the Perry Study, five extra children were transferred from the experimental group to the control group because they were unable to attend preschool due to lack of transportation or a working mother who could not arrange a home visit schedule with the teacher. Once assigned to a particular group, none of the families withdrew from the program. Neither parents nor teachers had a choice about which children attended preschool or which curriculum model they received.

The demographic characteristics of groups in each study are comparable. In the Perry Study, the only statistically significant difference between the groups is that there were more working mothers of children who did not attend preschool. This difference is the result of assigning more of such families to the control group; however, maternal employment in this sample is positively correlated with scholastic outcomes, thereby improving the control group's success on these measures and making the tests for preschool benefit more conservative. In the Curriculum Study, the only statistically significant difference is mothers' years of education completed.

Preschool Program Operation

The preschool programs to which treatment groups were assigned were organized educational programs directed at the intellectual and social development of young children. Each program was staffed by a teaching team who followed an explicit curriculum model and who received extensive managerial support and inservice training.

Children attended preschool for two years: at ages 3 and 4, except for the children in Wave 0 and part of Wave 5 who attended preschool for one year only at age 4 (see Table 1). The school year began in October and ended in May; classes were conducted for 2 1/2 hours a day, five days a week. In the classrooms, the staff-child ratio was one adult for every five to six children. Home visits by a teacher to each mother and child lasted 1 1/2 hours apiece. Home visits were conducted weekly during the Perry Preschool Study; every two weeks for Waves 5, 6, and 7 of the Curriculum Study; and not at all for Wave 8 of the Curriculum Study.

The Cognitively Oriented Curriculum employed in the Perry and Curriculum studies is detailed in The Cognitively Oriented Curriculum (Weikart, Rogers, Adcock, & McClelland, 1971). (The current description of the High/Scope curriculum was authored by Hohmann, Banet, & Weikart,

1979.) The behaviorist curriculum model used in the Curriculum Study is described in Teaching Disadvantaged Children in Preschool (Bereiter & Engelmann, 1966). The traditional American nursery school curriculum is described in a number of textbooks, such as The Nursery School (Read, 1966).

Variables Measured

A total of 71 measures have been collected annually from ages 3 to 11, at age 14, and now at 15 and 19. Family variables were assessed by parent reports. Youth variables were assessed by data from subjects, their teachers, and institutional records. The emphasis on data has consistently been on variables reflecting objective, verifiable status and performance. All these data are stored in operational computer files.

Family variables. Family status variables were measured twice in each study, at ages 3 and 15 in the Perry Preschool Study and at ages 3 and 10 in the Curriculum Study. While most of the instrumentation was developed by project staff, two measures developed elsewhere were employed: the Parent Attitude Research Instrument (Schaefer & Bell, 1958) and the Environmental Process Scale (Wolf, 1964). Birthdates of the father, the mother, and the youth in the study were determined. The sample youth's birth order and the number of siblings in the family were assessed. Education and employment characteristics of mother and father were assessed in detail; these same characteristics of siblings were also tapped at the later assessments. Father presence in the home was determined.

Self-reports of parent practices and attitudes were collected at age 15 in the Perry Preschool study and at age 10 in the Curriculum Study. Parent practices included child-rearing, school involvement, community involvement and voting. Parental attitudes included attitudes toward child-rearing, school, neighborhood, subject's friends, and scholastic aspirations for the subject. Parental attitude toward child-rearing was also assessed when children were aged 3 and 4 in the Perry Preschool Study.

Youth variables. Youth variables through age 15, collected by a youth interview, teacher ratings of classroom behavior, tests, and records are reported in Table 2. The variety of measures and the number of measures repeated annually are extensive. School performance was assessed almost every year by a number of tests of school potential and achievement, by both teacher and self-report of classroom behavior, and by school records of special education placement and grade repetition.

A comprehensive assessment of the Perry Study sample at age 19 and the Curriculum Study sample at age 15 is presently under way. The primary emphasis in this assessment is on real-life status and performance, with a secondary emphasis on test performance and attitudes. The variables are

shown in Table 3. Instrumentation was selected from the following sources: the APL Survey (American College Testing Program, 1976); the WISC-R (Wechsler, 1974); the Neighborhood Youth Corps Questionnaire (Freeberg, 1974); the Monitoring the Future Questionnaire (Bachman & Johnston, 1978); the Rosenberg Self-Esteem Inventory (Rosenberg, 1965); the Arlin-Hills Survey (Psychologists and Educators, Inc.); and the Bialer locus of control questionnaire (Bialer, 1961).

Attrition, i.e., cases for which data were missing, has averaged 9.1 percent across all 71 measures in the studies: 9.4 percent across the 48 measures in the Perry Preschool Study and 8.5 percent across the 23 measures in the Curriculum Study.

Results

School Functioning in the Primary Grades

These studies have produced a variety of evidence that participation in a well-run preschool program can lead to lasting improvement in school functioning. In the primary grades, improved school functioning was evidenced through improved IQ scores, improved achievement, and favorably rated classroom behavior.

School Functioning in Adolescence

Improved school functioning of children attending the Perry Preschool was even more dramatically evidenced in early adolescence. These teenagers who had attended preschool were found to score substantially higher in academic achievement in the eighth grade, to require costly special education programs half as often, and to have a better attitude toward school.

Discussion

The High/Scope longitudinal preschool studies, among field evaluations of educational programs, come closest to being true experiments. Their use of a type of random assignment to groups and low rate of attrition contribute to internal validity and confidence that group differences are correctly attributed to the preschool interventions. The studies meet the scientific requirements to provide a test of the potential of preschool intervention.

Less clear is the generalizability of these findings. The important findings of reduced need for special education has been replicated in projects of the Consortium for Longitudinal Studies for which such data

were available (Consortium, 1977). However, few of these studies are as yet in possession of the data which would confirm or deny the Perry study's findings of preschool effects in adolescence.

The Perry Preschool program preceded the nationwide preschool programs of Project Head Start. As with most of the programs operated by members of the Consortium, it was not intended to be representative of all Head Start programs. If anything, it came to be regarded as a model or demonstration program. Perhaps that same relationship is still appropriate: the Perry study represents the potential for compensatory preschool education and suggests directions for its development.

The tone of this chapter is clearly optimistic. Compensatory preschool education is not the unqualified failure that some have claimed on the basis of short-term evidence. Preschool education can help to overcome the school failure widely experienced by children raised in poverty.

But this optimism must be realistic. The late Ira Gordon pointed out that "we continue to go through a cycle of selecting a cause and designing a single solution to problems in spite of our knowledge that life is complex and there are no simple answers" (Gordon, 1979, p. 2). Not so long ago preschool education was seen by many as the single solution. Then it was not a solution at all. It would indeed be rash to use the evidence of the Perry Preschool study and other such studies as a basis for reinstating preschool education as a single solution. Preschool education is not nearly enough. Poverty still grinds down those who seek to overcome it. But, at least, preschool education may be able to help.

References

- American College Testing Program. User's guide: Adult APL survey. Iowa City, Iowa: Author, 1976.
- Arthur, G. The Arthur adaptation of the Leiter International Performance Scale. Beverly Hills, Calif.: Western Psychological Services, 1952.
- Bachman, J. G. & Johnston, J. The Monitoring the Future Questionnaire. Ann Arbor: University of Michigan, Institute for Social Research, 1978.
- Bachman, J. G., O'Malley, P. M., & Johnston, J. Adolescence to adulthood: Change and stability in the lives of young men (Youth in transition, Vol. 6). Ann Arbor: University of Michigan, Institute for Social Research, 1978.
- Bereiter, C., & Engelmann, S. Teaching disadvantaged children in the preschool. Englewood Cliffs, N.J.: Prentice-Hall, 1966.
- Bialer, I. Conceptualization of success and failure in mentally retarded and normal children. Journal of Personality, 1961, 29, 303-320.
- Consortium on Developmental Continuity. The persistence of preschool effects. Final report, Grant No. 18-76-07843 to the Administration for Children, Youth and Families (OHDS, DHEW), September 1977.
- Deutsch, M. et al. The disadvantaged child. New York: Basic Books, 1967.
- Dunn, L. M. Peabody Picture Vocabulary Test manual. Minneapolis, Minn.: American Guidance Service, 1965.
- Durost, W. N., Bixler, H. H., Wrightstone, J. W., Prescott, G. A., & Balow, I. H. Metropolitan Achievement Test manual. New York: Harcourt, Brace, Jovanovich, 1971.
- Freeberg, N. W. Development of assessment measures for use with youth-work training program enrollees. Phase II: Longitudinal evaluation. Princeton, N. J.: Educational Testing Service, 1974.
- Gordon, I. J. International year of the child. Seek continuity in social policy concerning child achievement. Young Children, 1979, 34, (3), 2-3.
- Hohmann, M., Banet, B., & Weikart, D. P. Young children in action: A manual for preschool educators. Ypsilanti, Mich.: High/Scope Press, 1979.

Kirk, S. A. Early education of the mentally retarded. Urbana: University of Illinois Press, 1958.

Kirk, S. A., McCarthy, J., & Kirk, W. D. Examiner's manual: Illinois Test of Psycholinguistic Abilities (Rev. ed.). Urbana: University of Illinois Press, 1968.

Klaus, R. A., & Gray, S. W. The early training project for disadvantaged children: A report after five years. Monographs of the Society for Research in Child Development, 1968, 33 (4, Serial No. 120).

Psychologists and Educators. Arlin-Hills Survey. Jacksonville, Ill.: Author, undated.

Read, K. H. The nursery school: A human relations laboratory (4th ed.). Philadelphia, Penn.: Saunders, 1966.

Rosenberg, M. Society and the adolescent self-image. Princeton, N. J.: Princeton University Press, 1965.

Schaefer, E. S., & Bell, R. Q. Development of a parental attitude research instrument. Child Development, 1958, 28, 339-361.

Schweinhart, L. J., & Weikart, D. P. Perry Preschool effects nine years later: What do they mean? Paper presented at a conference on Prevention of Retarded Development in Psychosocially Disadvantaged Children at the University of Wisconsin, Madison, July 1978.

Schweinhart, L. J., & Weikart, D. P. Perry Preschool effects in adolescence. Paper presented at the bi-annual meeting of the Society for Research in Child Development, San Francisco, March 1979.

Terman, L. M. & Merrill, M. A. Stanford-Binet Intelligence Scale, Form L-M: Manual for the third revision. Boston, Mass.: Houghton Mifflin, 1960.

Tiegs, E. W., & Clark, W. W. Manual: California Achievement Test, complete battery. Monterey Park: California Test Bureau (McGraw-Hill), 1963.

Tiegs, E. W., & Clark, W. W. Test coordinator's handbook: California Achievement Tests. Monterey, California: CTB/McGraw-Hill, 1970.

Vinter, R. D., Sarri, R. S., Vorwaller, D. J., & Schafer, W. E. Pupil Behavior Inventory: A manual of administration and scoring. Ann Arbor: Campus Publishers, 1966.

Weber, C. U., Foster, P. W., & Weikart, D. P. An economic analysis of the Ypsilanti Perry Preschool Project. Monographs of the High/Scope Educational Research Foundation, 1978, No. 4.

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children. New York: Psychological Corporation, 1949.

Wechsler, D. Manual for the Wechsler Intelligence Scale for Children (Rev. ed.). New York: Psychological Corporation, 1974.

Weikart, D. P. Results and implications of preschool intervention programs. Paper presented for the Assistant Commissioner for Program Planning and Evaluation, U. S. Office of Education, March 1967.

Weikart, D. P., Bond, J. T. & McNeil, J. T. The Ypsilanti Perry Preschool Project: Preschool years and longitudinal results through fourth grade. Monographs of the High/Scope Educational Research Foundation, 1978, No. 3.

Weikart, D. P., Deloria, D., Lawser, S., & Wiegink, R. Longitudinal results of the Ypsilanti Perry Preschool Project. Ypsilanti, Mich.: High/Scope Foundation, 1970.

Weikart, D. P., Epstein, A. S., Schweinhart, L. J., & Bond, J. T. The Ypsilanti Preschool Curriculum Demonstration Project: Preschool years and longitudinal results. Monographs of the High/Scope Educational Research Foundation, 1978, No. 4.

Weikart, D. P., Rogers, L., Adcock, C., & McClelland, D. The Cognitively Oriented Curriculum: A framework for preschool teachers. Urbana: University of Illinois--NAEYC, 1971.

Westinghouse Learning Corporation/Ohio University. The impact of Head Start experience on children's cognitive and affective development. Springfield, Va.: U. S. Department of Commerce Clearinghouse, 1969. (PB 184 328)

Wolf, R. M. The identification and measurement of environmental process variables related to intelligence. Unpublished doctoral dissertation, University of Chicago, 1964.

Zigler, E., & Butterfield, E. C. Motivational aspects of changes in IQ test performance of culturally deprived nursery school children. Child Development, 1968, 39 (1), 1-14.

Table 2
(Continued)

MEASURE/SOURCE/VARIABLES	AGE AT MEASUREMENT*																
	Preschool				Elementary					Middle			High School				
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
<u>TESTS OF ACHIEVEMENT</u> (Tiegs & Clark, 1963, 1970; Durost et al., 1971) Reading, language, mathematics, total school achievement					B	B	P	B	P			P	C				P
<u>HOSPITAL RECORDS</u> (In Weikart, Deloria, Lawser, & Wiegerink, 1970) Perinatal stress	P																
<u>SCHOOL RECORDS</u> Special education, grade repetitions					B	B	B	B	B	B	B	B	B	B	B	B	B

*P = Perry Study C = Curriculum Study B = Both studies

Table 3
 Age 15 and Age 19 Assessments in Progress
 Cross-Classification of Major Variables

Measurement Modes Contexts	Self-Report	Institutional Records	Test Performance	Subject Attitude Toward:	Attitude Toward Subject by:
Delinquency	Delinquent behavior Police contacts	Courts	Coping with government & law	Importance of law	
School	School programs	School (e.g., drop-out, truancy, disciplinary actions) Special education	School aptitude* Basic skills in coping	School Future schooling	Counselor Teacher
Economic Activity	Employment Unearned Income Investments	Employment	Economic coping	Employment Future employment	Employer Supervisor
Family and Community	Live with parents Marital status Children Community involvements	Social agencies	Coping in the community	Family Spouse Future marriage Neighborhood	Parent
Self	Leisure activities (i.e., consumption vs. production) Health		Productive language*	Self-esteem Sense of control*	Interviewer

*Age 15 assessment only--hypothesized curriculum differences.

Table 1
Sample Size by Group and by Wave

	Wave										Group Totals
	Perry Study					Curriculum Study					
	0	1	2	3	4	5	6	7	8		
CHILDREN WITH PRESCHOOL	Number of Cases										
Cognitive	13 ¹	8	12	13	12	11	5	6	9 ²		79
Behaviorist						8 ¹	7	8	8 ²		31
Nursery						8 ¹	7	8	7 ²		30
CHILDREN WITHOUT PRESCHOOL	15	9	14	14	13						65
Wave Totals	23	17	26	27	25	27	19	22	24		215

¹One year of preschool at age four.

²Unlike other waves, this wave did not receive home visits; during their second preschool year, all three groups were exposed to the Cognitive curriculum.

Table 2
Youth Variables

MEASURE/SOURCE/VARIABLES	AGE AT MEASUREMENT*																
	Preschool				Elementary					Middle			High School				
	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
<u>YOUNG ADULT ASSESSMENT</u> (See Table 5)													C				P
<u>YOUTH INTERVIEW</u> (Based on Bachman, O'Malley, & Johnston, 1978) Behavior: delinquency, school troublemaking, homework, employment, friendship, leisure time. Attitudes: toward school, toward parents, aspirations, self-esteem, sense of control													P				P
<u>CLASSROOM BEHAVIOR</u> (Pupil Behavior Inventory by Vinter, Sarri, Vorwaller, & Schafer, 1966; other scales developed by staff) Classroom conduct, personal behavior, academic motivation, verbal skill, emotional adjustment, social development, mother participation, academic potential, sociability, cooperativeness, independence from teacher				P	B	B	P										
<u>TESTS OF APTITUDE</u> (Stanford-Binet test by Terman & Merrill, 1960; nonverbal Leiter adaptation by Arthur, 1952; PPVT by Dunn, 1965; ITPA by Kirk, McCarthy, & Kirk, 1968; WISC by Wechsler, 1949) Academic potential, language abilities, sight vocabulary, various components of aptitude	B	B	B	B	B	B	P	B			P	C					

* P = Perry Study C = Curriculum Study B = Both studies

Monoamine Oxidase Activity and the Longitudinal Biological High
Risk Approach to Schizophrenia and Affective Illnesses*

Robert D. Coursey, Ph.D.
University of Maryland

Monte S. Buchsbaum, M.D.

and

Dennis L. Murphy, M.D.
National Institute of Mental Health

*Editors' Note

The inclusion of biological measures as antecedent measures is not very common in longitudinal data files. This review of the literature describes how one such biological measure has been used in one type of longitudinal research.

Monoamine Oxidase Activity and the Longitudinal Biological High
Risk Approach to Schizophrenia and Affective Illnesses

Robert D. Coursey, Ph.D.
University of Maryland

Monte S. Buchsbaum, M.D.

and

Dennis L. Murphy, M.D.
National Institute of Mental Health

Introduction

The biological assay of blood platelet monoamine oxidase (MAO) activity has gathered strong empirical support as a significant correlate of chronic schizophrenia (for reviews, see Wyatt & Murphy, 1976; Wyatt, Potkin & Murphy, 1979) and possibly of bipolar affective illness (Leckman, Gershon, Nichols & Murphy, 1977; Murphy & Weiss, 1972). The purposes of this chapter are to (1) examine the suitability of platelet MAO activities for longitudinal research, (2) inquire into possible methodological pitfalls and unexamined assumptions of the various research strategies which use this measure, (3) explain some of the difficulties with these research strategies and the reasoning that led to the biological high risk model, and (4) present the biological high risk method and review several recent longitudinal studies using this model.

Requirements of the MAO Factor for Longitudinal High Risk Research

There are at least five major presuppositions that require empirical support before one can accept with some assurance the utility of MAO blood platelet activity assays in longitudinal research, and the underlying neurochemical-psychological mechanisms upon which this research is based.

The first prerequisite is that platelet MAO activities can be reliably measured. The second assumption that requires evaluation is whether platelet MAO activities are generally stable characteristics of individuals, and to what extent can they be influenced by transient environmental factors such as diet, hormones or drug use. A related, more specific question is the extent to which platelet MAO activities are genetically derived. Fourthly, there is the large question of whether platelet MAO activities can be meaningfully related to brain MAO activity and to neuronal function. And finally, the fifth question is whether there exist stable, replicable correlates of MAO activity with psychological traits and behavior.

Measurement Reliability

Human platelets contain readily-assayable amounts of MAO activity when radioactively-labeled amines are used as substrates for the enzyme. A 20-fold range of MAO activities was observed in the largest group of normal subjects (680 individuals) reported in the literature (Murphy et al., 1976). Variations do occur within subjects, but they are small compared to between-subjects differences.

Both assay replicability and test-retest reliabilities are high. One assay using benzylamine as the substrate and using platelet counts obtained with an electronic particle counter yielded high correlations for sample-to-sample comparisons as well as for individuals studies over various time periods (See Table 1, from Post & Murphy, 1979).

Assay sensitivity may be considerably less when certain substrates (e.g., serotonin) are used (Donnelly & Murphy, 1977) and also may be lower when spectrophotometric or fluorometric assay methods rather than radioenzymatic procedures are used. Although high correlations have been found between results obtained with benzylamine vs. tryptamine or tyramine vs. tryptamine as substrates, other studies have reported apparent differences in platelet activities (particularly in individuals with psychopathology) when platelet MAO activities with different substrates were compared (Demisch, Muhlen, Bochnik & Seiler, 1977; Meltzer & Stahl, 1974). Careful attention to substrate concentrations in relation to the Michaelis constant, (K_m 's) of the enzyme and also to produce separation techniques

are required for valid comparisons of this type.

Differences in the collection, preparation, and enumeration of platelets are also important factors potentially limiting assay reproducibility and undoubtedly contribute to the variability in results reported from different laboratories. Platelet centrifugation procedures, in particular, may yield platelet populations of different density and size, with corresponding differences in platelet MAO activity (Murphy, Costa, Shafer & Cerash, 1978). Our group has come to favor basing the assay on the number of platelets in aliquots from the samples examined for enzyme activity. This is because the alternative standardization procedure based upon "platelet" protein measurements actually includes variable amounts of plasma proteins adherent to platelets, as well as varying numbers of leukocytes and erythrocytes (Murphy and Donnelly, 1974; Murphy, Wright, Buchsbaum, Nichols, Costa & Wyatt, 1976).

According to the data available, MAO activity measurements using the assay procedures which we have developed, standardized and evaluated, yield highly reproducible results, and provide a reliable measure of the activity of this platelet enzyme.

Platelet MAO Activity: Stability Over Time and Changes in Relation to Hormones, Age, Diet, Drug Action, and Psychiatric Status

In order to carry out longitudinal research on the relationship of biochemistry to behavioral expression, there must be some stability in the biological factor or at least some known lawfulness about its variability. The intra-individual MAO activity level is generally stable over long periods of time and under a variety of conditions. Yet there are factors which seem to affect it. We shall look at hormones and the menstrual cycle, age, diet, drug effects, and changes in psychiatric status.

Some hormones such as estrogen and testosterone (Luine, Knychevskava, & McEwen, 1975; Luine & McEwen, 1977) may affect MAO activity in animals. Contraceptive steroids have been reported to alter uterine MAO activity in one study (Southgate, Collins, Pryse-Davis & Sandler, 1970), although a relatively non-specific histochemical assay procedure was used in this investigation. In contrast, contraceptive steroids yielded no change in platelet MAO activities in a more recent study (Feldman & Roche, 1976). Changes in platelet MAO have been observed during the human menstrual cycle, although the changes, while consistent in pattern, were small (23% from peak to nadir) and only slightly greater than the 15% variation seen upon repetitive sampling in males.

There also appear to be age-related increases up to 25% in platelet MAO activity beginning perhaps around 45 in males and over 60 in females (Nies, Robinson, Davis & Ravari, 1971; Robinson, Davis, Nies, Ravari & Sylvester, 1971; Robinson, Davis, Niles, Colburn, Davis, Bourne, Bunney, Shaw, & Coppen, 1972; Irving, Coursey, Buchsbaum & Murphy, in preparation) but no significant correlation with ages below 60 (Murphy, Wright, Buchsbaum, et al., 1976). This age-related increase was also noted in Macaca Mulatta (Murphy, Redmond, Baulu & Donnelly, 1978). But all of these studies have been across different-aged individuals rather than within individuals across time. Also note that all of the differences reported here are considerably smaller than the over 20-fold range of values found between individuals.

Little work has been reported on the effects of diet on MAO activity and presumably MAO is little altered by food ingestion (Belmaker, Murphy, Wyatt & Loriaux, 1974; Post & Murphy, 1979), although the effects of short and long-term nutritional differences have yet to be evaluated. An exception is the finding that iron deficiency anemia is associated with reduced blood platelet MAO activity (Yodanis, Grahame-Smith & Woods, 1976).

Despite its general stability, platelet MAO activity can be altered by drugs and, apparently, in association with at least a few bodily changes. Some antidepressants and antihypertensive drugs have as their major mode of action the production of markedly reduced MAO activity in all tissues (including platelets). The question has been raised as to whether other drugs, not widely recognized as specific MAO inhibitors, might also reduce the activity of the enzyme to some extent. Some psychoactive drugs such as tricyclic anti-depressants do seem to inhibit MAO activity, at least in vitro (Edwards & Burns, 1974; Sullivan, Dackis & Stanfield, 1977; Becker & Shaskan, 1977). It has not been demonstrated, however, that sufficiently high concentrations of tricyclics are obtained in patients receiving these drugs to result in MAO inhibition. In contrast, the drugs commonly used to treat schizophrenia, the phenothiazines and butyrophenones, have, if anything, a tendency to increase MAO activity (Brockington, Crow, Johnstone & Owen, 1976; Murphy, Belmaker, & Wyatt, 1974; Wyatt & Murphy, 1976). In addition subcutaneous administration of adrenaline has been reported to increase MAO levels (Gentil, Greenwood & Lader, 1975). To date, none of the common street drugs tested (amphetamine tetrahydrocannabinol) have been reported to significantly lower MAO levels (Donnelly & Murphy, 1977).

Finally there is generally fairly high stability across time in psychiatric patients. Murphy, Belmaker, Carpenter, and Wyatt (1977) measured MAO activities in chronic schizophrenics and unipolar depressed patients over a two-year period and normals over two months, while Carpenter, Murphy, and Wyatt (1975) studied acute schizophrenics over one year. Only 15 to 20% changes over time were found in all of these groups.

This intra-individual stability across time and across marked changes in clinical state ranging from acute and severely psychotic states with changes in physical activity, dietary intake, sleep, drug treatment, and remission (Murphy, Belmaker, Carpenter & Wyatt, 1977; Carpenter, Murphy & Wyatt, 1975) as well as during cycles of severe mania and depression in a bipolar patient (Murphy & Wyatt, 1975) makes dubious some suggestions that short-term changes in MAO activity levels might be related to changes in clinical symptoms (Becker & Shaskan, 1977). It would seem more likely that MAO activity might represent a stable factor which interacts with environmental changes to result in psychological differences. Alternatively, of course, MAO activity may represent only a biochemical marker for vulnerability to a variety of psychiatric disorders.

Are MAO Activities Under Genetic Control

Because both schizophrenia and bipolar affective disorders seem under partial genetic control, the question arises about blood platelet MAO activities. Brain MAO activity in rodents and platelet MAO activity in man have both recently been demonstrated to be under a high degree of genetic control. A number of studies in mice and rats have demonstrated differences in MAO activity between strains, although little work has been accomplished in regard to the mode of genetic transmission.

In man, three independent twin studies in normals have demonstrated significantly higher intraclass correlation coefficients for monozygotic compared to dizygotic twins (Table 2), with the magnitude of difference suggestive that most of the variance in platelet MAO activity can be accounted for on a genetic basis (Murphy, Belmaker & Wyatt, 1974; Nies, Robinson, Lamborn & Lampert, 1973; Winter, Herschel, Propping, Fredi & Vogel, 1978). High correlation coefficients have also been demonstrated for platelet MAO activity in monozygotic twins with affective disorders (Murphy, 1973) and monozygotic twins discordant for schizophrenia (Wyatt, Murphy, Belman, Chen, Donnelly & Pellia, 1973). First degree relatives of patients with affective disorders (Leckman et al., 1977), alcoholism (Sullivan, Caverce, Maitbie, Lister & Zung, et al., 1979), and schizophrenia also have more similar MAO activities than do unrelated individuals. Sib-pair comparisons in normal humans (Murphy & Donnelly, 1974) and in Rhesus monkeys have also demonstrated higher correlation coefficients than pairs matched for age and sex, further suggesting genetic control for platelet MAO activity.

Familial influence on MAO activity was also observed in a study of normal individuals with relatively high or low platelet MAO activity (Puchall, Coursey, Buchsbaum & Murphy, 1980). In this study, however,

no Mendelian pattern of inheritance was revealed. Specifically, mother-son correlations were not higher than father-son correlations, thus not supporting data from somatic cell hybrids grown in tissue culture which suggested that the structural gene for the catalytic subunit of MAO-A is on the X chromosome (Pinter, Barbosa, Hawkings, Castiglioni, Francke & Breakefield, 1980).

Are MAO Blood Platelet Activities Related to Neuronal Transmission?

This is a complex and extremely difficult question to evaluate because many mammalian species including mice, rats, cats and some other common laboratory animals lack MAO activity in platelets (Paasonen & Solatunturi, 1964) rendering many experimental approaches to this problem impossible. The complexity of this assumption can be made manifest by examining the issues involved in how platelet MAO activities relate to brain MAO activities which are available to degrade synaptic transmitters in vivo. It is important to recall that MAO is only one small part of the total chemical chain and structural mechanisms that regulate synaptic transmission. Extremely low MAO activity, for example, may be compensated for or reached to by a variety of feedback mechanisms in another part of the chemical chain.

In regard to even the most elementary question of whether there is a relationship between MAO activities measured in platelets and those measured in brain and other tissues of the same individual, no definitive answer is yet available. MAO is a complex enzyme, or more likely a family of enzymes, with two major forms (type A and type B) now recognized. Platelets contain MAO type B exclusively (Donnelly & Murphy, 1977), whereas almost all other tissues contain mixes of the A and B types in varying proportions. MAO-A is the predominant form in most rodent tissues, including the brain, while there is a higher proportion of MAO-B than MAO-A in human and subhuman primate brain (Murphy, Redmond, Garrick & Baulu, 1979). Moreover, brain regions in rodents and primates differ in their proportions of MAO-A to MAO-B, as well as in total MAO activity. Methods for distinguishing MAO-A from MAO-B activity are not completely specific. In addition, non-neuronal elements in brain tissue, including blood vessels and glial cells grown in tissue culture. Techniques to separate neurons from glial are fairly crude. There are also many different types of neurons, many of which do not use the amines as neurotransmitters, and hence isolation of amine-containing neuronal elements which only contain MAO-B has not yet been accomplished. This level of complexity makes it highly unlikely that simple comparisons of platelet vs. brain MAO activity will prove to be readily interpretable, even if the difficulties in obtaining usable post-mortem brain samples can be overcome. Because intravascular coagulation occurs at the time of death, platelet samples need to be obtained

some time prior to death, again adding to the difficulty of doing such a study in man.

Despite the extreme difficulty in obtaining direct data relating platelet MAO activity to brain MAO, there are some indications that platelet MAO activity may serve as an index of MAO activity in other tissues, including brain. For example, MAO-inhibiting drugs affect MAO-B activity in platelets and brain to a similar extent. Treatment with these drugs in animals and man leads to behavioral changes as well as other biochemical and pharmacologic effects (such as intracellular amine accumulation and blood pressure changes) which are interpretable consequences of the interruption of amine degradation by the inhibitors. However, presuming neuronal transmission is modified, it remains unknown how these changes in neuronal transmission might affect such psychological dimensions as loose associations. On the other hand, there is much animal data to support an association between reduced MAO activity and increased psychomotor activation and related phenomena. Additionally, of course, no one pretends that fundamental cognitive and/or affective symptoms necessarily result in a patient seeking treatment and receiving a diagnosis. This final condition of a diagnosed patient implies a breakdown of social supports, availability and timing of alternative care, stress, and other changes.

Are There Stable, Replicable Correlations of MAO with Psychological Expression?

Because longitudinal research implies the stable commitment of large resources, it would be foolhardy to undertake a biologically-based longitudinal high risk project without some empirically-established probability for payoff.

For the most part, studies of human MAO activity have sought to establish a relationship between MAO and a diagnostic category, in the hope that MAO is etiologically relevant to the fundamental symptoms of these categories. However, a close logical examination of this hypothesis in its simplest form (low MAO activity causes diagnoses x and y) suggests that MAO cannot be the main cause of all the fundamental symptoms of two diagnoses: bipolar affective illness and chronic schizophrenia. While one frequently finds loose associations in both, it is hard to reconcile the notion that low MAO activity in isolation from other factors could be responsible for the elation of mania and the anhedonia in schizophrenia, the social aggressiveness of mania and the social withdrawal of schizophrenia, the sensation seeking quality of mania and the stimulus withdrawal in chronic schizophrenia. The above suggests that MAO research, having established some type of relationship

of MAO with schizophrenia (Wyatt, Potkin & Murphy, 1979) and perhaps other diagnostic categories, needs to become much more analytic in its investigation of this biochemical-psychological relationship. In other words, this simplified hypothesis needs to be made more complex and sophisticated at both the biochemical-neurophysiological end as well as at the behavioral, psychological expression end. The present paper focuses on the psychological correlates (both normal-range traits and psychopathological symptoms) of MAO activity. We will first look at traits in the normal range.

Because blood platelet MAO activities range along a continuum, it seems reasonable to suppose that its psychological expression would be a continuum. And because MAO activities are normally distributed (Buchsbaum & Reider, 1979; Murphy et al., 1976), and most values lie within normal boundaries, then most of the range of psychological expression of MAO should be non-pathological. Indeed, many such normal correlates have been found. The dimensions that have been most replicated seem related to the person's disposition to seek stimulation in the outside environment. Examples of this theme include correlation of low MAO values with high scores on the Sensation Seeking Scale (Murphy & Belmaker, Buchsbaum, Martin, Cirrarelle & Wyatt, 1977; Schooler, Zahn, Murphy, & Buchsbaum, 1978), with increased social interaction in both humans (Coursey, Buchsbaum & Murphy, 1979) and monkeys (Redmond, Murphy & Baulu, 1979), cigarette usage (Coursey, Buchsbaum & Murphy, 1979; Irving, Coursey, Buchsbaum & Murphy, in preparation) as well as drug usage (Coursey, Buchsbaum, & Murphy, 1979) and a tendency toward more stimulating leisure-time activities in man (Schooler, Zahn, Murphy & Buchsbaum, 1978), and psychopathic traits and legal convictions (Buchsbaum, Coursey & Murphy, 1976).

Although MAO correlates with the above dimensions across the whole central range of MAO values; nevertheless, at the extremes of this MAO continuum, there may exist some specific relationships to more pathological states, like mania. Thus, groups of subjects with extremely high MAO activity levels have been found to have negative affect (Schooler et al., 1978) and depression-related diagnoses (Nies, Robinson, Harris & Lamborn, 1974; Puchall, Coursey, Buchsbaum, & Murphy, 1980), or incidences of depressive episodes (Irving, Coursey, Buchsbaum & Murphy, in preparation), and perhaps paranoid characteristics, especially in psychopathological samples (Haier, Murphy & Buchsbaum, 1979; Potkin, Cannon, Murphy & Wyatt, 1978; Demisch, Muhlen, Bochnik et al., 1977; but see Berger, Ginsburg, Barchas, Murphy & Wyatt, 1978; Van Valkenburg & Crowe, 1978).

On the other hand, extremely low MAO males have been found to have higher involvement with mental health treatment (Buchsbaum, Coursey & Murphy, 1976; Coursey, Buchsbaum & Murphy, no date; Irving et al., in preparation), and suicide (Buchsbaum, Heier & Murphy, 1977; Buchsbaum, Coursey & Murphy, 1976). Low MAO values also seem to be associated with

chronic schizophrenia (see reviews Wyatt & Murphy, 1976; Wyatt, Potkin, & Murphy, 1979), bipolar affective disorders (Leckman et al., 1977; Murphy & Weiss, 1972), and alcoholism (Major & Murphy, 1978; Sullivan, Cavenor, Maltbie, Lister & Zung, 1979; Sullivan, Stanfield, Maltbie, Hammett & Cavenar, 1973; Wieberg, Gottfries & Orelund, 1977).

The area where most of the studies have been done, the relationship of MAO activity with diagnostic categories, bears special scrutiny because a diagnostic category consists of multiple psychological traits and behaviors as well as physical and psychological consequences of the essential dysfunctions, combined with adaptive coping efforts and defenses against the ravages of the disorder, all embedded in a psychosocial context (e.g., institutionalization, Mednick & McNeil, 1968) which further shapes the diagnostic picture. Thus, a relationship between some level of MAO activity and this physio-psycho-social conglomerate, upon closer inspection, must mean a relationship with some one or several aspects of the diagnosed person's psychological state' and thus the biochemical variable cannot be automatically assumed to have etiological significance (Schooler et al., 1978; Purchall et al., 1980). It seems plausible that MAO activity could be related to any of six aspects of a diagnosed person's state. For purposes of this paper, we will examine the relationship of MAO with the diagnosis of chronic schizophrenia.

First, MAO might be related to a central deficit such as loose associations (Coursey, Buchsbaum, & Murphy, 1980) or attentional difficulties (Buchsbaum, Murphy, Coursey, Lake & Zeigler, 1978). However, this cannot be an exclusive association with the critical components of schizophrenia alone because low MAO activity is also associated with sensation seeking behavior, which at face value does not seem compatible with chronic schizophrenia (Schooler et al., 1978). Moreover, the low level of MAO found in schizophrenic samples in a statistical rather than an absolute difference, and there are instances of high MAO among subjects carrying diagnoses of chronic schizophrenia (Schildkraut, Orsulak, Schatzberg, Cole, Gudeman, Rhode, 1978) and mania (Purchall et al., 1980). Proposed explanations include the suggestion that MAO may disinhibit or facilitate either biochemically or psychologically some other fundamental etiological factor. For example, it is quite possible that the stimulant effects of low MAO might be so stressful to a schizophrenia-prone person as to cause a breakdown or exacerbation of symptoms (Schooler et al., 1978). Along a similar vein, the weak relationship of MAO and diagnosis may be due to the fact that schizophrenic dysfunction may result from the interaction of MAO with other variables; for example, with the physiological dimensions of augmenting-reducing (Haier, Buchsbaum, Murphy, Gottesman, Coursey, 1980; Coursey, Buchsbaum & Murphy, 1979, 1980), dopamine-beta-hydroxylase (DBH) (Buchsbaum, Murphy, Coursey, Lake & Zeigler, 1978), cell membrane transport characteristics (Dorus, Pandey, Shaughnessy & Davis, 1979; Shaughnessy, Doris, Pandey & Davis, 1980), or with other not-yet examined defects such as GSR factors (Mednick & McNeil, 1968; Saltzman & Klein, 1978) or brain damage (Goldstein,

1978; Mirsky, 1969). Finally, the relationship may not be due to a linear correlation with MAO at all, but simply the extremity of the deviation may set in effect other mechanisms which result in schizophrenia (Coursey, Buchsbaum, & Murphy, 1980; Irving et al., in preparation; Haier, Buchsbaum, Murphy, Gottesman & Coursey, 1980).

A second possible explanation of the MAO-diagnosis connection is that MAO may not be related either directly or indirectly to some fundamental psychiatric symptom, but rather to a subsidiary, related phenomenon. This is suggested by the fact that while reduced MAO activities have been rather consistently found in chronic schizophrenia, in a number of studies of acute schizophrenia, MAO showed little or no reduction (Carpenter, Murphy, & Wyatt, 1975; Meltzer & Stahl, 1974). Thus, MAO may be related to some factor producing chronicity rather than directly to classical schizophrenic symptoms. A related explanation has been put forth by Murphy & Buchsbaum (1978) that low MAO might produce a more generalized vulnerability to manifest psychopathology rather than etiologically cause any particular type (Schooler, et al., 1978). An example of this type of vulnerability factor would be the appearance of mania, depression or thought disorder following the massive endocrinological changes in the post-partum period. Here, a single type of biological insult gives rise to a variety of symptom patterns. Still, the concept of a "generalized vulnerability" needs further development.

A third possibility is that MAO may be related to some third factor which is commonly associated with the diagnosis, but not of any etiological significance, such as treatment effects, associated addictive behavior, etc. For example, it could be postulated that very long term exposure to neuroleptics lowers platelet MAO (as is required for tardive dyskinesia). However, the possibility that this completely explains the clinical reports of low MAO in psychiatric illness is considerably lessened because of the high correlation of MAO levels in first-degree well relatives of patients and in monozygotic twins discordant for schizophrenia (Leckman, Gershon, Nichols, & Murphy, 1977; Wyatt et al., 1973), as well as the association of low MAO with psychopathology in individuals untreated and undiagnosed (Buchsbaum et al., 1977).

Another way in which MAO activity might appear related to psychopathology would be through some mediating variable, such as alcoholism in affective disorders or smoking in chronically institutionalized patients. For instance, tobacco smoking was found to be highly associated with low MAO levels in several studies (Coursey, et al., 1979; Irving et al., in preparation). This relationship requires further exploration, especially since institutionalized patients such as chronic schizophrenics, and alcoholic persons tend to smoke more. Thus, low values of MAO in chronic schizophrenics and other patient groups could conceivably be due to the failure to control for this particular addiction.

Several other studies (e.g. Major & Murphy, 1978; Sullivan et al., 1977; Sullivan et al., 1979; Wieberg et al., 1977) have found low platelet MAO activity levels related to alcoholism. It thus might be the higher incidence of alcohol usage in affective disorders which could be responsible for low MAO activity rather than a direct relationship of MAO activity and affective disorders.

In these alcoholism studies, the investigators were unable to rule out the possibility that alcoholic consumption might reduce MAO activity rather than that low MAO levels constitute a predisposition to alcoholism. However, the results of a study by Puchall et al. (1980) decreases the likelihood that alcohol causes the low MAO levels observed in alcoholics because this study identified people at risk for alcoholism on the basis of robands who had low MAO activity levels but were not alcoholics themselves. A similar question might be raised about the low MAO activity found in 26 marijuana smokers when compared to non-smokers. However, no acute reduction in MAO was found following smoking a marijuana cigarette containing 15 mg of delta-9-tetrahydro cannabinol (THC), nor did THC inhibit MAO activity in vitro concentrations greatly exceeding in vivo plasma concentrations (Stillman, Wyatt, Murphy & Rauscher, 1978). These studies suggest that lowered MAO activity is not a consequence of these habits or addictions, and thus opens the door to a possible relationship of MAO and the habit or addiction prone personality.

A fourth possible explanation of the MAO-diagnosis relationship is that reduced MAO levels are a consequence of the disorder rather than the cause. Psychiatric patients might exercise less, or consume unusual diets affecting monoamine metabolism. However, low MAO levels have not been found in hospitalized controls and control of dietary monoamines has not affected platelet MAO. Another possibility is that changes in MAO activity may be a consequence of behavior activity or cortical or autonomic arousal, rather than a cause. If this were true, one would expect this arousal to be a non-specific stress which might affect other biochemical changes, such as cortisol or norepinephrine, which in turn might affect MAO levels (Gentil, Greenwood & Lader, 1975). Neither norepinephrine or cortisol have reliably differentiated schizophrenics or controls with anywhere near the limited success of MAO.

A fifth possibility is that MAO might be related to some traits such as sensation-seeking, disinhibition of impulse control, and/or rule-breaking that negatively impacts the patient's environment and social support system. Thus, these traits could call attention and concern to persons disturbed by a variety of biological and/or psychosocial factors which, in turn, would facilitate entering the mental health system and receiving a psychiatric diagnosis and treatment. In several studies (Buchsbaum, Coursey, & Murphy, 1976; Coursey, Buchsbaum, & Murphy, submitted; Irving et al., in preparation), low MAO subjects had seen a psychiatrist but were not presently diagnosable. Or these MAO-related traits may simply create

stressful situations (e.g. divorce, jail terms) which precipitate breakdowns in vulnerable persons.

Finally, extreme MAO levels may be simply a marker for vulnerability to psychiatric illness (Murphy & Buchsbaum, 1978; Post & Murphy, 1979) and may be unrelated to whether the person is merely predisposed, severely psychotic, in remission, or may even be unrelated to a single disorder. The evidence that supports this possibility was discussed above under the stability of MAO activities within individuals.

Three Research Methodologies for Use with

Biological Factors: Their Pitfalls and Problems

Psychopathological Group versus Normal Controls

The overwhelming majority of the studies investigating the psychiatric relevance of MAO activity have used the traditional biochemical research paradigm of comparing the assays of a pathological group (e.g., schizophrenia, affective disorders) with a control group. The methodological difficulties with this approach to biochemical differences have been well documented elsewhere (Kety, 1959; Mednick & McNeil, 1968; Bannister, 1971; Post & Murphy, 1979, etc.). But in order to understand the rationale for our new approach using the biological high risk model, we need to enumerate some fundamental strategic difficulties in four areas with this basic research method.

The first set of problems center on the selection criteria for the pathological group. Beyond the problems of establishing clear and reliable categories (e.g. Buchsbaum & Haier, 1978; Falek & Moser, 1975) with symptomatic homogeneity, particular problems arise in using a disease category method in neurochemical investigations of psychological disturbances. The disease model makes most sense when there is some clear abnormality, trauma, or pathogen, and when the symptoms are themselves directly traceable to its pathophysiology. Trauma and pathogens suggest a disjunctive, qualitative difference between health and disease, and diagnostic categories make sense in that context. But many of the neurochemical physiological variables studies, such as MAO, DBH, and evoked potentials, are continuous and normally disturbed. Also it is quite unlikely that an experiential category, defined without regard to a biological substrate, based primarily on language and social phenomena, would correspond even roughly with some underlying biochemical abnormality. Even if a group of subjects were perfectly homogeneous with respect to psychiatric symptoms, there is no guarantee that they would be homogeneous with respect to etiology. Indeed, most researchers and theoreticians assume biological heterogeneity in schizophrenia and in affective disorders (e.g. Buchsbaum & Rieder, 1979; Murphy & Buchsbaum, 1978;

Buchsbaum & Haier, 1978; Bellak, 1979; Kraepelin, 1968; Sullivan, 1947). Given such a situation, Buchsbaum & Reider (1979) have demonstrated through a computer simulation the low probability of finding and replicating biological differences when only a certain percentage of the pathological group has a particular biological etiology.

The second major deficit for this method is establishing an adequate control group which is equivalent on all other relevant dimensions (such as the effects of long-term stress, institutionalization, diet and habits of hospitalized patients, subtle accompaniments such as mild brain damage, etc.).

The third area is untangling essential biological factors from other interactive factors and artifacts. This method simply is unable to sort out a principle biological factor from other relevant, interactive, or irrelevant but concomitant biological and environmental variables. For instance, any attempt to examine family and stress variables in the group comparison strategy are retrospective, and thus subject to considerable bias (Fontana, 1966). Further, one cannot establish the power of a particular biological factor because of these confounding effects. Moreover, etiologically significant variables cannot be distinguished from disease or treatment consequences. Now that we know some of the long-lasting side effects of pharmacological treatment (e.g., tardive dyskinesia) we must also worry about permanent biochemical alterations due to past treatment.

Finally, there is the danger of false attribution. Even if one could definitely prove a relationship between MAO and diagnosis, and rule out all artifacts, one cannot conclude that the specific symptom constellation is caused by this biological factor. The psychological expression of the same biological factor may vary according to environment, learning, coping and defenses, intelligence, culture, and other biological variables. Thus, extreme MAO activity, as pointed out above, may simply be a general mental health vulnerability factor or a genetic marker.

Longitudinal High Risk Research

In its most fundamental form, this research establishes the high risk group on the basis of a psychiatric diagnosis of the mother, and then the children are periodically assessed over a 15-20 year period (Mednick & McNeal, 1968). As outlined by Mednick & McNeil (1968), this approach has many desirable features, including excellent procedures for establishing control groups. It also helps untangle the antecedents from the consequences of the disorder, but cannot establish the direct relevance of any of the biological and psychological differences uncovered, nor the power of these factors. Moreover, its basic procedure for establishing the high

risk group still depends on phenomenological diagnoses with their attendant problems mentioned above. Furthermore, there is still a danger of some false attribution. For example, psychological expression may be due to pathological familial environment (unless an adopted away method is used) instead of biological variables. Moreover, since the focus of the study is on psychological deviance, the normal adaptive correlations between biological variable and psychological expression will be ignored. While there are numerous other problems (Fisher & Jones, 1978), an important practical one is the great cost of finding and following the subjects longitudinally.

The Biological High Risk Method

This recent method (Buchsbbaum, Coursey & Murphy, 1976; Coursey, Buchsbbaum & Murphy, 1979; Buchsbbaum & Reider, 1979) suggests that instead of using psychiatric diagnosis of mother as a criterion for including a subject in the high risk group, and then looking for differences in biological and psychological variables in their offspring, that one can define the subjects in the high risk group solely on the basis of some hypothesized deviant biological variable. This method has a number of further additional attractions over previous forms of high risk research. One, it tests explicit hypotheses about a biological variable and its psychological correlates, as well as gives some indication of how potent that variable is in relationship to other vicissitudes of living. Second, it avoids the problem of biological heterogeneity (Buchsbbaum & Reider, 1979) and lack of reliability of psychiatric diagnoses. Third, it requires the researcher to explore the hypothesized psychological expression of the biological variable in the normal as well as pathological range.

This model also has drawbacks. One, while it allows adequate control groups, much of the adequacy of the controls are based on randomization found in the normal population. For instance, differences in psychopathology found in the Coursey, Buchsbbaum & Murphy study (1979) were attributed to MAO activities rather than to family variables, because the random selection procedure was presumed to have eliminated systematic biases in the families. But child-parent correlations of MAO activity levels as well as increased levels of psychopathology among parents of children possessing extremes of MAO activity (Puchall, et al., 1980) make disentangling biology from modeling and family interactions still impossible without adopted-away studies. Thus far, control groups have been established primarily on the basis of their MAO values, although age, education, socio-economic variables, and sex were equivalent. This last variable is particularly important in light of the common reports of different mean MAO values between males and females (Buchsbbaum, Coursey, & Murphy, 1976; Robinson, Davis, Nies, Ravaris & Sylvester, 1971; Murphy & Weiss, 1972; Murphy & Wyatt, 1975; Murphy, Belmaker, Buchsbbaum, Martin,

Ciaranello & Wyatt, 1977, etc.) and the reversal of some effects on the basis of sex (Coursey, Buchsbbaum & Murphy, 1980; Murphy, Belmaker, Buchsbbaum, et al., 1977; Puchall et al., 1980). The first control group used in this type of research had extremely high MAO levels. They were selected in the hope that extreme deviations would enhance the likelihood of finding MAO effects and because psychopathology had not been found related to high MAO activity. Subsequent reports have found predictable adjustment problems in these high MAO groups (Puchall, et al., 1980; Irving et al., in preparation; Schildkraut, Orsulak, Schatzberg, Cole, Gudeman & Rohde, 1978). Consequently, a middle MAO group is now being used for comparison with the extremes (Irving et al., in preparation). As other factors affecting MAO levels (e.g., possibly cigarette smoking) and psychopathology (e.g., stress, Irving et al., in preparation) become established, the need for more carefully matched control groups will emerge.

As with earlier longitudinal high risk research, here too, the biological high risk approach has high costs. First, the large scale screening of subjects (10% of the screening group has been used) to find the biologically deviant group is expensive, though probably not more so than careful psychiatric work-ups, since the screening does not require high level professionals. Second, the differences that are sought consist of two types (hypothesized differences within the normal range and incidences of abnormal functioning). The cost of the first is that extensive testing for more subtle differences must be undertaken. The cost of the second is that the rate of disturbance is quite low, perhaps even lower than that found among children coming from families with one or two disturbed parents. In our work, we have tried to contain this expense in several ways. First, we have generally worked with subjects who are closer to the age of highest risk and more accessible than have most previous studies (college-aged students rather than children). Second, in another study (Haier, Buchsbbaum, Murphy, Gottesman, & Coursey, 1980), we tried to improve the incidence of disturbance by screening subjects who were biologically extreme and fell in a group of subjects possessing broadly-defined psychopathology (T-scores above 90 on any of the MMPI clinical scales). The reports on these two longitudinal studies are presented below.

A final note, the biological high risk method, because of its expense, will probably not ordinarily be undertaken unless there is some empirical support for the biology-psychiatry relationship from traditional research methods. At present, we are using this method on two further investigations. The biological high risk criteria are deviant eyetracking in one study and attentional deficits in processing in the other. Venables (1977) is also using this method with GSR selection criteria for the high risk group in the Mauritian study.

Longitudinal Research Using Biological Selection Criteria

Probably the first such longitudinal research using a biologically defined high risk sample is the Mauritian study (Bell, Mednick, Raman, Schulsinger, Sutton-Smith & Venables, 1975; Mednick, Shulsinger & Venables, 1970; Venables, 1977) which uses hyperresponsivity and hypo-responsivity on GSR to define the high risk groups. From the initial screening of 1800 children, four matched groups of 50 were selected. In each group of 50, 26 were hyperresponsive, 10 were GSR nonresponders, and the rest were normal. The children were selected at age three and the first follow-up took place at age six where the children were observed in a standardized play situation. There were interactions between GSR responding and nursery-community placement on "constructive play", "watching", and "positive interaction" in accord with the patterns that might be expected to be associated with the psychophysiological types of behavior measured. The children are now 10 years old and further studies are reported elsewhere in this volume.

Two further longitudinal high risk studies using chemical criteria (MA) have been reported. The first study used college-age probands selected solely on the basis of extremely high and low MAO blood platelet activities. The initial results from this approach have been reported elsewhere (personal and family history, Coursey, Buchsbaum & Murphy, 1979; psychophysiological data, Buchsbaum, Murphy, Coursey, Lake & Zeigler, 1980; and a study of their parents, Puchall, Coursey, Buchsbaum & Murphy, 1980). Contact with the subjects was made by phone approximately two years after completion of the initial assessment. The initial phone contact was not made by the interviewer, so the interviewer was blind to the person's identity and MAO level. The structured interview examined changes in the subject's social state, progress in college, work history, medical problems, legal problems, their own mental health and that of their families. Using this procedure, we were able to contact and interview 16/19 low MAO males, 14/17 high MAO males, 17/18 low MAO females, and 16/17 high MAO females from the original group.

The results were as follows. There were few significant changes in the subject's social state (e.g., marriages, pregnancies, children); the overwhelming majority of each group was still single.

The low MAO males have fallen about a half a year behind their high MAO counterparts in school (M low = 2.1 years of college versus M high = 2.8 years; $t = 1.76$, one-tailed $p < 0.05$). This finding of more job changes among the high-risk group was not due simply to the larger number of low MAO subjects working full time since the subjects who began the most new jobs, or were fired were generally those who worked part time or were not working at all at the time of the interview, who were not in school, and who were psychiatrically more disturbed.

The low MAO males and females both reported more major and minor medical problems requiring a physician than did the high MAO groups, though the differences were not significant. However, significantly more low MAO subjects smoked cigarettes than the high MAO subjects (Fisher exact test, $p < .004$). Low MAO males alone smoked significantly more than high MAO males (Fisher exact test, $p = .001$).

Five low MAO subjects had been arrested for violating the law (possession of drugs, destruction of property, intoxication, etc.) versus three high MAO subjects (p not significant). As expected, seven of the eight arrested subjects were male ($p < .05$, Fisher exact test).

In the area of reported mental health, no significant differences were found during this two year period, although more low MAO males had seen a psychiatrist (two versus zero), more (five versus two) experienced a serious depression (two weeks duration meeting four or more of the Research Diagnostic Criteria (RDC); Spitzer, Endicott, & Robins, 1978), and more thought about suicide (five versus two).

Finally, we inquired about the mental health of their families. Since there were three adoptions (two low MAO males and one low MAO female), the adoptee's family reports were excluded from the comparisons. We first asked them if there were in the past two years or currently significant mental health problems in their families. Eight low MAO males answered yes versus one high MAO male ($p = .006$, Fisher exact test). There were no significant difference for the females. Based on previous studies (Nies, Robinson, Davis & Ravaris, 1971; Nies, Robinson, Harris & Lamborn, 1974; Sullivan, Cavenar, Maltbie, Lister & Zung, 1979; Buchsbaum, Haier & Murphy, 1977), we then inquired about three specific problems (depression, alcoholism, and suicide attempts). Among the males, there were seven low MAO probands who had one or more of these problems among their relatives, versus one family of a high MAO male proband ($p = .026$, Fisher exact test). There were no significant differences among the females (five versus eight). We then inquired about three specific mental health interventions (psychiatric visits, prescription of psychotropic medication, and psychiatric hospitalization). Ten low MAO subjects reported one or more of the above treatments had occurred in their families while there were only two reports by the high MAO subjects ($p = .01$, Fisher exact test). Among the males, five families of low MAO probands had received at least one of the treatments versus no families of the high MAO probands ($p = .026$, Fisher exact test). Among the females, five low MAO families and two high had been treated (not significant).

Overall, the low MAO subjects, and especially the males, continue to show signs of poorer functioning in almost all of the areas of their life that we examined, although only one serious psychiatric breakdown (low MAO female) had occurred during this period.

In the second longitudinal study involving MAO, 385 males were screened on the MMPI. All those scoring above 90 on any clinical scale (indicating psychological difficulties) and a group scoring within normal MMPI limits were selected for further biological screening. Sixty one subjects (34 with aberrant MMPI's and 27 with normal MMPI's) were chosen based on their platelet MAO activity assays (Murphy, Wright, Buchsbaum, Nichols, Costa & Wyatt, 1976) and the recording of amplitude/intensity slope of their visual evoked potentials (Buchsbaum, Haier & Murphy, 1977). Of these 61 subjects, 38 were included as either having high platelet MAO activity levels (<11.5 nanomoles/ 10^8 platelets/hr.) or low platelet MAO (<8.5) and divided further into augmenters (<1.00 Microvolts/log footlambert) or reducers (<1.00). The platelet MAO criterion was the same cutoff used in the previous biological risk study and the evoked potential cutoff was the median amplitude/intensity slope in the same study (Coursey, Buchsbaum & Murphy, 1980). The results of this initial screening in terms of MMPI patterns and a clinical interview with a psychiatrist using the Schedule for Affective Disorders and Schizophrenia, the life history version (SADS-L), and the Research Diagnostic Criteria (RDC) are reported elsewhere, (Haier, Rieder, Khouri & Buchsbaum, 1979; Haier, Buchsbaum, Murphy, Gottesman & Coursey, 1980).

Thirty two of these 38 subjects have completed their first wave of follow-up interviews (Haier, Buchsbaum & Murphy, 1979) between 17 and 20 months after their original interview.

A research assistant telephoned each subject and scheduled the follow-up interview. All these interviews were conducted over the telephone by one psychologist who had not participated in the original interviewing and was blind to biological measures. Each interview lasted between 15 and 45 minutes and included the latest SADS-L probes and RDC for major depression, hypomania, and schizotypal features.

At the original interview, the distribution of the 37 subjects meeting the RDC for any affective disorders was: low MAO augmenters 50% (6/12), low MAO reducers 0% (0/12), high MAO augmenters 20% (1/5) and high MAO reducers 55% (5/9). Overall, 52% of the low MAO augmenters and high MAO reducer groups ($n=21$) met the RDC for any diagnostic category whereas only 5% of the low MAO reducing and high augmenting groups are so characterized ($p<0.005$, chi-square). In the same group at follow-up the distribution of subjects meeting the RDC for a new episode of major depression or hypomania during the interval followed a similar pattern of increased incidence of low MAO augmenters and high MAO reducers; low MAO augmenters 83% (10/12), low MAO reducers 28% (2/7), high MAO augmenters 0% (0/4), and high MAO reducers 33% (3/9). The distribution of episodes in the four biological groups again differed significantly from chance whether tested in a four cell comparison or tested by combining the cells containing subjects found to be at higher risk (low MAO augmenting and high MAO reducing) in the previous studies and comparing them with the two remaining cells.

A second analysis concerned the 27 cases that did not meet the RDC for major depression at the original interview. Of these cases, six did meet the RDC for major depression at the follow-up interview. All six had low MAO (Fisher exact test, $p<0.027$) and four were augmenters ($p<0.05$ by chi-square, again combining cells as above). Note, however, that all subjects in this case were low MAO.

Summary

In summary, we have proposed five criteria that a biological factor should satisfy before being used as a diagnostic tool for longitudinal high risk research. We believe that the present evidence surrounding these five criteria (though much remains to be researched) well justifies further use of blood platelet MAO activity in such longitudinal research. In this chapter we have particularly focused on a review of the dependent variables and methodological problems that we believe are of significance in carrying out this research on MAO. Finally, we presented some preliminary data from several studies which lend support to our belief in the fruitfulness of this research strategy and this biological marker.

References

- Bannister, D. The logical requirements of research into schizophrenia. Schizophrenia Bulletin, 1971, 1 (Experimental issue No. 4), 72-77.
- Becker, R.E. & Shaskan, E.G. Platelet monoamine oxidase activity in schizophrenic patients. American Journal of Psychiatry, 1977, 134, 512-517.
- Bell, B., Mednick, S.A., Raman, A.C., Schulsinger, F., Sutton-Smith, B., & Venables, P.H. A longitudinal psychophysiological study of three-year-old Mauritian Children: Preliminary report. Developmental Medicine and Child Neurology, 1975, 17, 320-324.
- Bellak, L. (Ed.) Disorders of the Schizophrenia Syndrome. New York: Basic Books, 1979.
- Belmaker, R.H., Murphy, D.L., Wyatt, R.J., & Loriaux, D.L. Human platelet monoamine oxidase changes during the menstrual cycle. Archives of General Psychiatry, 1974, 31, 553-556.
- Berger, P.A., Ginsburg, R.A., Barchas, J.D., Murphy, D.L., & Wyatt, R.J. Platelet monoamine oxidase activity in chronic schizophrenic patients. American Journal of Psychiatry, 1978, 135, 95-99.
- Brockington, I., Crow, T.J., Johnstone, E.C., & Owen, F. An investigation of platelet monoamine oxidase activity in schizophrenia and schizoaffective psychosis. In G. E. Wolstenhome & J. Knight (Eds.) Monoamine Oxidase and its Inhibition: CIBA Foundation Symposium, 39, New Series. New York: Elsevier, 1976.
- Buchsbaum, M.S., Coursey, R.D., & Murphy, D.L. The biochemical high risk paradigm: Behavioral and familial correlates of low platelet monoamine oxidase activity. Science, 1976, 339-341.
- Buchsbaum, M.S. & Haier, R.J. Biological homogeneity, symptom heterogeneity, and the diagnosis of schizophrenia. Schizophrenia Bulletin, 1978, 4, 473-475.
- Buchsbaum, M.S., Haier, R.J. & Murphy, D.L. Suicide attempts, platelet monoamine oxidase and the average evoked response. Acta Psychiatrica Scandinavica, 1977, 56, 69-79.
- Buchsbaum, M.S., Murphy, D.L., Coursey, R.D., Lake, C.R., & Zeigler, M.G. Platelet monoamine oxidase, plasma dopamine-beta-hydroxylase and attention in a "biochemical high-risk" sample. Journal of Psychiatric Research, 1978, 14, 215-224.
- Buchsbaum, M.S. & Rieder, R.O. Biologic heterogeneity and psychiatric research. Archives of General Psychiatry, 1979, 36, 1163-1169.
- Carpenter, W.T., Murphy, D.L., & Wyatt, R.J. Platelet monoamine oxidase activity in acute schizophrenia. American Journal of Psychiatry, 1975, 132, 438-441.
- Coursey, R.D., Buchsbaum, M.S., & Murphy, D.L. Platelet MAO activity and evoked potentials in the identification of subjects biologically at risk for psychiatric disorders. British Journal of Psychiatry, 1979, 134, 372-381.
- Coursey, R.D., Buchsbaum, M.S. & Murphy, D.L. Psychological characteristics of subjects identified by platelet MAO activity and evoked potentials as biologically at risk for psychopathology. Journal of Abnormal Psychology, 1980, 89, 151-164.
- Coursey, R.D., Buchsbaum, M.S., & Murphy, D.L. Two-year follow-up of subjects defined as biologically at risk on the basis of platelet MAO activities. Submitted.
- Demisch, L., Muhlen, H., Bochnik, H., Seiler, N. Substrate typic changes of platelet monoamine oxidase activity in subtypes of schizophrenics. Arch. Psychiatrica Nervenkr., 1977, 224, 319-329.
- Donnelly, C.H. & Murphy, D.L. Substrate and inhibitor-related characteristics of human platelet monoamine oxidase. Biochemical Pharmacology, 1977, 26, 853-858.
- Dorus, E., Pandey, G.N., Shaughnessy, R., & Davis, J.M. Low platelet monoamine oxidase activity, high red blood cell lithium ration, and affective disorders: A multivariate assessment of genetic vulnerability to affective disorders. Biological Psychiatry, 1979, 14, 989-992.
- Edwards, D.J. & Burns, M.D. Effects of tricyclic antidepressants upon human platelet monoamine oxidase. Life Sciences, 1974, 15, 2045-2058.
- Falek, A., & Moser, H.M. Classification in schizophrenia. Archives of General Psychiatry, 1975, 32, 59-67.
- Feldman, J.M. & Roche, J. Effect of oral contraceptives on platelet monoamine oxidase, monoamine excretion, and adrenocortical function. Clinical Pharmacology and Therapeutics, 1976, 20, 670-675.
- Fisher, L. & Jones, F.H. Planning for the next generation of risk studies. Schizophrenia Bulletin, 1978, 4, 223-235.

- Fontana, A.F. Familial etiology of schizophrenia: Is a scientific methodology possible? Psychological Bulletin, 1966, 66, 214-227.
- Gentil, V., Greenwood, M.H. & Lader, M.H. The effect of adrenaline on human platelet MAO. Psychopharmacologia, 1975, 44, 187-194.
- Goldstein, G. Cognitive and perceptual differences between schizophrenics and organics. Schizophrenia Bulletin, 1978, 4, 160-185.
- Haier, R.J., Buchsbaum, M.S. & Murphy, D.L. Platelet monoamine oxidase and cortical evoked potentials as predictors of psychiatric vulnerability. In E. Usdin, I.J. Kopin, & J. Barchas (eds.) Catecholamines: Basic and Clinical Frontiers, Vol. 2 N.Y.: Pergamon Press, 1979.
- Haier, R.J., Buchsbaum, M.S., Murphy, D.L., Gottesman, I.I., & Coursey, R.D. Psychiatric vulnerability, monoamine oxidase, and the average evoked potential. Archives of General Psychiatry, 1980, 37, 340-345.
- Haier, R.J., Murphy, D.L., & Buchsbaum, M.S. Paranoia and platelet MAO in normals and nonschizophrenic psychiatric groups. American Journal of Psychiatry, 1979, 136, 308-310.
- Haier, R.J., Rieder, R., Khouri, P., & Buchsbaum, M.S. Extreme MMPI scores and the research diagnostic criteria: Screening college males for psychopathology. Archives of General Psychiatry, 1978, 36, 528-534.
- Irving, J., Coursey, R.D., Buchsbaum, M.S. & Murphy, D.L. The psychiatric effects of MAO activities and stress in adult males. (In preparation).
- Kety, S.S. Biochemical theories of schizophrenia, part II. Science, 1959, 129, 1590-1596.
- Kraepelin, E. Lectures on clinical psychiatry. Authorized translation from the German by Thomas Johnston. London: Hafner, 1968.
- Leckman, J.F., Gershon, E.S., Nichols, A.S., & Murphy, D.L. Reduced monoamine oxidase activity in first degree relatives of bipolar affective disorder - a preliminary report. Archives of General Psychiatry, 1977, 34, 601-606.
- Luine, V.N., Knychevskaya, R.I., & McEwen, B.S. Effect of gonadal steroids on activities of monoamine oxidase and choline acetylase in rat brains. Brain Research, 1975, 86, 293-306.
- Luine, V.N. & McEwen, B.S. Effect of oestradiol on turnover of type A monoamine oxidase in brain. Journal of Neurochemistry, 1977, 28, 1221-1227.
- Major, L.F., & Murphy, D.L. Platelet and plasma amine oxidase activity in alcoholic individuals. British Journal of Psychiatry, 1978, 132, 548-554.
- Mednick, S.A. & McNeil, T.F. Current methodology in research on schizophrenia: Serious difficulties which suggest the use of the high-risk group method. Psychological Bulletin, 1968, 70, 681-693.
- Mednick, S.A., Schulsinger, F., & Venables, P.H. Risk research and primary prevention of mental illness. International Journal of Mental Health, 1979, 7, 150-164.
- Meltzer, H.Y. & Stahl, S. M. Platelet monoamine oxidase activity and substrate preferences in schizophrenic patients. Res. Commun. Chem. Pharmacol., 1974, 7, 419-431.
- Mirsky, A.F. Neuropsychological bases of schizophrenia. In P.H. Mussen & M.R. Rosenzweig (eds.) Annual Review of Psychology, Palo Alto California: Annual Reviews, 1969. pp. 321-348.
- Murphy, D.L. Technical strategies for the study of catecholamines in man. In E. Usdin & S. Snyder (Eds.) Frontiers in Catecholamine Research. Oxford: Pergamon Press, 1973.
- Murphy, D.L., Belmaker, R., Buchsbaum, M.S., Martin, N.F., Ciaranello, R., & Wyatt, R.J. Biogenic amine-related enzymes and personality variations in normals. Psychological Medicine, 1977, 7, 149-157.
- Murphy, D.L., Belmaker, R., Carpenter, W.T., & Wyatt, R.J. Monoamine oxidase in chronic schizophrenia. Studies of hormonal and other affecting enzyme activity. British Journal of Psychiatry, 1977, 130, 151-158.
- Murphy, D.L., Balmaker, R., & Wyatt, R. Monoamine oxidase in schizophrenia and other behavioral disorders. Journal of Psychiatric Research, 1974, 11, 221-247.
- Murphy, D.L. & Buchsbaum, M.S. Neurotransmitter-related enzymes and psychiatric diagnostic entities. In R.L. Spitzer & D.F. Klein (eds.) Critical Issues in Psychiatric Diagnosis, N.Y.: Raven Press, 1978.
- Murphy, D.L., Costa, J.L., Shafer, B., and Corash, L. Monoamine oxidase activity in different density gradient fractions of human platelets. Psychopharmacology, 1978, 59, 193-197.
- Murphy, D. L., and Donnelly, C. H. Monoamine oxidase in man: Enzyme characteristics in platelets, plasma, and other human tissues. In E. Usdin (ed.) Neuropsychopharmacology of Monoamines and Their Regulatory Enzymes. New York: Raven Press, 1974.

Murphy, D.L., Redmond, D.E., Baulu, J., & Donnelly, C.H. Platelet monoamine oxidase activity in 116 normal rhesus monkeys: Relations between enzyme activity and age, sex, and genetic factors. Comp. Biochem. Physiol., 1978, 60, 105-108.

Murphy, D.L., Redmond, D.E., Garrick, N., & Baulu, J. Brain region differences and some characteristics of monoamine oxidase type A and B activities in the vervet monkey. Neurochemical Research, 1979, 4, 53-62.

Murphy, D.L. & Weiss, R. Reduced monoamine oxidase activity in blood platelets from bipolar depressed patients. American Journal of Psychiatry, 1972, 128, 1351-1357.

Murphy, D.L., Wright, C., Buchsbaum, M.S., Nichols, A., Costa, J.L. & Wyatt, R.J. Platelet and plasma amine oxidase activity in 680 normals: Sex and age differences and stability over time. Biochemical Medicine, 1976, 16, 254-265.

Murphy, D.L. & Wyatt, R.J. Neurotransmitter-related enzymes in the major psychiatric disorders: I. Catechol-O-Methyl transferase, monoamine oxidase in the affective disorders, and factors affecting some behaviorally correlated enzyme activity. In D.X. Freedman (ed.) Biology of the Major Psychoses, N.Y.: Raven Press, 1975, pp. 277-297.

Nies, A., Robinson, D.S., Davis, J.M., & Ravaris, C.L. Amines and monoamine oxidase in relation to aging and depression in man. Psychosomatic Medicine, 1971, 33, 470.

Nies, A., Robinson, D.S., Harris, L.S., & Lamborn, K.R. Comparison of monoamine oxidase substrate activities in twins, schizophrenics, depressives, and controls. In E. Usdin (ed.) Neuropsychopharmacology of Monoamines and Their Regulatory Enzymes. N.Y.: Raven Press, 1974.

Nies, A., Robinson, D.S., Lamborn, K.R., & Lampert, R.P. Genetic control of platelet and plasma monoamine oxidase activity. Archives of General Psychiatry, 1973, 28, 834-838.

Paasonen, M.K. & Solatunturi, E. Monoamine oxidase in mammalian blood platelets. Annals Med. Exp. Biol. Fenn., 1964, 43, 98-100.

Pintar, Barbose, Hawkins, Castiglione, Francke, C., & Breakefield, X.O. Manuscript in preparation, 1980.

Post, R.M. & Murphy, D.L. Research planning: Studies of cerebrospinal fluid metabolites and blood cell enzymes. In H.M. Van Praag (ed.) Handbook of Biological Psychiatry, N.Y.: Marcel Dekker, 1979.

Potkin, S.G., Cannon, H.C., Murphy, D.L., & Wyatt, R.J. Are paranoid schizophrenics biologically different from other schizophrenics? New England Journal of Medicine, 1978, 298, 61-66.

Puchall, L.B., Coursey, R.D., Buchsbaum, M.S., Murphy, D.L. The parents of high risk subjects biologically defined by high and low platelet MAO activity. Schizophrenia Bulletin, 1980, 6, 338-346.

Redmond, D.E., Murphy, D.L., & Baulu, J. Platelet monoamine oxidase activity correlates with social affiliative and agonistic behaviors in normal rhesus monkeys. Psychosomatic Medicine, 1979, 41, 87-100.

Robinson, D.S., Davis, J.M., Nies, A., Colburn, R.W., Davis, J.N., Bourne, H.R., Bunney, W.E., Shaw, S.M. & Coppen, A.J. Aging monoamines, and monoamine-oxidase levels. Lancet, 1972, 1, 290-291.

Robinson, D.S., Davis, J.M., Nies, A., Ravaris, C.L., & Sylwester, D. Relation of sex and aging to monoamine oxidase activity of human brain, plasma, and platelets. Archives of General Psychiatry, 1971, 24, 536-539.

Salzman, L.F., & Klein, R.H. Habituation and conditioning of electrodermal responses in high risk children. Schizophrenia Bulletin, 1978, 4, 210-222.

Schildkraut, J., Orsulak, P., Schatzberg, A., Cole, J., Gudeman, J., & Rhode, W. Elevated platelet MAO activity in schizophrenia-related depression disorders. American Journal of Psychiatry, 1978, 135, 110-112.

Schooler, C., Zahn, T.P., Murphy, D.L., & Buchsbaum, M.S. Psychological correlates of monoamine oxidase activity in normals. Journal of Nervous & Mental Disease, 1978, 166, 177-186.

Shaughnessy, R., Dorus, E., Pandey, G.N., & Davis, J.M. Personality correlates of platelet monoamine oxidase activity and red cell lithium transport. Psychiatry Research, 1980, 2, 63-74.

Southgate, J., Collins, G.G., Pryse-Davis, J. & Sandler, M. Effectiveness of contraceptive steroids on monoamine oxidase activity. Journal of Clinical Pathology, 1970, 3, 43-48.

Spitzer, R., Endicott, J., & Rubins, E. Research diagnostic criteria (RDC) for a selected group of functional disorders. (3rd Ed.) New York: New York State Psychiatric Institute, 1978.

Stillman, R.C., Wyatt, R.S., Murphy, D.L., & Rauscher, F.P. Low platelet monoamine oxidase activity and chronic marijuana use. Life Sciences, 1978, 23, 1955-1582.

Sullivan, H.S. Conceptions of modern psychiatry. Washington, D.C.: The William Allison White Psychiatric Foundation, 1947.

Sullivan, J.L., Cavenar, J.O., Maltbie, A.A., Lister, P., & Zung, W.W. Familial biochemical and clinical correlates of alcoholics with low platelet monoamine oxidase activity. Biological Psychiatry, 1979, 14, 385-394.

Sullivan, J.L., Dackis, C., & Stanfield, C. In vitro inhibition of platelet MAO activity by tricyclic antidepressants. American Journal of Psychiatry, 1977, 134, 188-190.

Sullivan, J.L., Stanfield, C.N., Maltbie, A.A., Hammett, E. & Cavenar, J.O. Stability of low blood platelet monoamine oxidase activity in human alcoholics. Biological Psychiatry, 1978, 13, 391.

Van Valkenburg, C., & Crowe, R. Monoamine oxidase in schizophrenia. New England Journal of Medicine, 1978, 298, 1150.

Venables, P.H. The electrodermal psychophysiology of schizophrenics and children at risk for schizophrenia: Controversies and developments. Schizophrenia Bulletin, 1977, 3, 28-48.

Wieberg, A., Gottfried, C.C. & Orelund, L. Low platelet monoamine oxidase activity in human alcoholics. Medical Biology, 1977, 55,

Winter, H., Herschel, M., Propping, P., Friedl, W., & Vogel, F. A twin study on three enzymes (DBH, COMT, MAO) of catecholamine metabolism: Correlations with MMPI. Psychopharmacology, 1978, 57, 63-69.

Wyatt, R.J., & Murphy, D.L. Low platelet monoamine oxidase activity and schizophrenia. Schizophrenia Bulletin, 1976, 2, 77-89.

Wyatt, R.J., Murphy, D.L., Belmaker, R., Cohen, S., Donnelly, C.H. & Pollin, W. Reduced monoamine oxidase activity in platelets: A possible genetic marker for vulnerability to schizophrenia. Science, 1973, 179, 916-918.

Wyatt, R.J., Potkin, S.G., & Murphy, D.L. Platelet monoamine oxidase activity in schizophrenia: A review of the data. American Journal of Psychiatry, 1979, 136, 377-385.

Youdin, M.B., Grahame-Smith, D.C., & Woods, H.F. Some properties of human platelet monoamine oxidase in iron-deficiency anemia. Clin. Sci. Mol. Med., 1976, 50, 479-485.

Table 1

Platelet MAO Assay Comparisons in Normal Subjects

	N	Correlation coefficient (r)
Same individual, same sample, same assay day	52	.98
Same individual, same sample, different assay date	57	.99
Same individual 2 to 8 days apart	26	.94
Same individual, 3 months apart	42	.86
Same sonicated sample, tryptamine versus benzylamine as substrate	75	.89

Table 2
Genetic Influences on Platelet MAO Activity

	Intraclass correlation coefficients		
	Murphy (1973)	Nies et al. (1973)	Winter (1978)
Normals			
Monozygotic twins	0.88	0.76	0.86-0.92*
Dizygotic twins	0.45	0.39	0.47-0.70*
Sib pairs	0.28	-	-
Random pairs matched for age and sex	0.12	-0.16	-
Monozygotic twins discordant for schizophrenia (Wyatt et al., 1973)	0.65		
Monozygotic twins concordant for bipolar manic-depression illness (Murphy, 1973)	0.83		

*Depending on whether tryptamine, phenylethylamine, or tyramine was used as substrate.

A Longitudinal Study of Personality Development*

Joan McCord

*This study was supported by U.S. Public Health Service Research Grant No. ROI MH26779, National Institute of Mental Health (Center for Studies of Crime and Delinquency). It was conducted jointly with the Department of Probation of the Commonwealth of Massachusetts. The author wishes to express appreciation to the Division of Criminal Justice Services of the State of New York, to the Maine State Bureau of Identification, and to the states of Florida, Michigan, New Jersey, and Washington for supplemental data about the men, though only the author is responsible for statistical analyses and for the conclusions drawn from this research. Richard Parente, whose inventiveness and tenacity were so largely responsible for the participation of subjects, deserves special appreciation.

A Longitudinal Study of Personality Development

Joan McCord

In the midst of the Great Depression, the late Richard Clark Cabot, a Professor of Clinical Medicine and Social Ethics at Harvard, grafted scientific methodology onto a social action program. With vision rare among reformers, Dr. Cabot insisted that his project, known as The Cambridge-Somerville Youth Study, incorporate a control group against which results of the social action program could be measured.

Developed in hopes of preventing delinquency by providing long-term, friendly assistance, the Youth Study had several features which made it unusual: 1) Participants were selected because they lived in designated neighborhoods and because each could be matched to another child of similar background. Thus, biases associated with self-selection were not introduced. 2) Both "difficult" and "average" boys were included. 3) Boys were assigned to treatment by chance, with a match assigned to a control group. Therefore, the treatment program could be evaluated by comparing men of similar background of whom half had been randomly selected for treatment. 4) The project maintained excellent records, enabling investigators at a later date to use the information for research purposes. The present longitudinal study of personality development traces former members of this project over a period of approximately 40 years.

Method

Between 1935 and 1939, a Selection Committee gathered information from elementary school teachers, the courts, physicians, and parents of boys between the ages of 5 and 11 who lived in the designated working-class areas of Cambridge and Somerville, Massachusetts. On the basis of this information, the Committee identified 325 pairs of children; each pair was similar in age, family background, home environment, and delinquency-prone histories. By flip of a coin, one member in each pair was assigned to a group which would receive treatment. Eight cases were matched after treatment began; the assignment to the treatment group for these 8 was not random. All brothers were assigned to that group which was the assignment of the first brother matched.

Description of the Data

The longitudinal study provides four perspectives on the lives of the 506 men: that of teachers perceiving young children, that of social workers perceiving adolescents, that of formal agency records reflecting behavior of the men, and that of the subjects themselves as middle-aged men. The earliest perspective is that of teachers. In 1936-1937, teachers had been asked to describe the boys on "Trait Record Cards." The cards contained a list of descriptive words or phrases (e.g., "lies," "blames others for his mistakes," "bullies or teases," "has few friends") and a place for comments. At the time of the rating, the boys ranged in age from 4 to 11. Their mean age was 7.6, with a standard deviation of 1.6. Over 200 teachers had rated the boys, reducing reliability for the scales while also decreasing the possibility that perceptual biases would account for obtained relationships between the ratings and subsequent behavior.

The case records provide the data for the second perspective on the 506 subjects. The files of the Cambridge-Somerville Youth Study contained different types of information for the treatment and the control groups. For the control group, records had been kept from the original home visits, from physical and psychological examinations prior to assignment to the control group, from teachers' reports, and from three interviews with someone in the family. For the treatment group, the records contained detailed reports of each encounter between a Youth Study staff person and either the boy or his family. Thus, there were running records covering the years of treatment, 1939-1945. These records included precise descriptions of behavior and verbatim reports of conversations.

In 1957, six trained readers coded the case histories. The codes included information about the socio-economic and cultural milieu of a family, about the parents (e.g., their aggressiveness, conflicts, affection for one another, drinking behavior), about parental interaction with the boy (e.g., disciplinary techniques, expectations for performance, affect toward one another), and about the boy (e.g., relations with siblings, self-confidence, dependency, aggressiveness). The coders also recorded information about the services provided by The Cambridge-Somerville Youth Study; the latter included information about frequency of visits and focus of counselor-boy interaction. A second coder read a randomly selected sample of cases to estimate the reliability of coding. Reliabilities ranged from 67% agreement for a six-category scale describing the father's techniques of discipline to 100% agreement on such variables as the family's cultural background, the boy's relative position in the family, and the age of the mother. Since none of the readers knew what had happened to the boys after termination of the program in 1945, retrospective biases were prevented.

The third perspective is based on the information from formal records. These include records from the Crime Prevention Bureau, established in 1938 to help prevent juvenile delinquency. In 1975, the names and pseudonyms of each person in the study were submitted to the Massachusetts Department of Probation. The Probation Department maintained continuous records

Ten social workers, a psychologist, tutors, a shop instructor, and consulting psychiatric and pediatric doctors formed the treatment staff. By May 1939, each boy in the treatment group had been assigned to a social worker who was expected to build a close relationship with the boy and be available to assist the boy and his family. Counselor turnover and recognition that case loads were too heavy led to a decision to drop some of the boys from the program. When a boy was dropped from the treatment program, his "matched mate" was dropped from the control group. A comparison of the pairs remaining after this selective "retirement" in 1940 and 1941 indicated that treatment and control groups were strikingly similar. No reliable differences were found through comparisons of age, IQ, referral to the Youth Study as "average" or "difficult," or in the delinquency prediction scores assigned by the Selection Committee on the basis of descriptions of the boys' family histories and home environments. No reliable differences appeared in comparisons regarding the boys' physical or mental health, social adjustment, acceptance of authority, or social aggressiveness. Nor were reliable differences found in ratings of delinquency in the home, adequacy of discipline, standard of living, status of the occupation of the father, "social status level" of the elementary school attended (as measured by sampled occupational levels of the parents), or neighborhoods as likely or unlikely to produce delinquency. (See Powers & Witmer, 1951, Chapter VI.)

The present study includes 253 pairs of men born between 1925 and 1934 ($\bar{X}=1928$). These men were the boys who remained in the Youth Study after January 1942, minus 3 pairs dropped because the boy in the treatment group had died prior to December 31, 1945. (No member of the control group was found to have died prior to December 31, 1945.) The 506 subjects came from 466 families. Although the fathers of 12% held white collar positions, the majority of fathers were skilled or unskilled workers.

In 1945, when the program terminated, the 253 boys in the treatment group had been visited approximately twice a month for five and a half years. On any particular visit, a counselor might have talked with a boy about problems with a girlfriend, helped him with school work, given him a driving lesson, or gone for a walk. Sometimes counselors helped parents find jobs or work out family problems. Over a hundred of the boys had been provided medical or psychiatric assistance; over half had been tutored; almost half had been sent to summer camps; and most of the boys had participated in such activities as swimming, local athletic competitions, and work in the project's woodshop. Social workers from the Youth Study were specifically prohibited from working with boys in the control group.

Records from the Youth Study files included information used for identifying pairs of boys similar to one another, detailed reports by the case workers assigned to members of the treatment group, and records collected in 1948 as part of an evaluation of the program. Information from the Youth Study files has been augmented through a follow-up begun in 1975.

showing convictions in the courts of Massachusetts. Together, these sources provided information about both unofficial and official delinquency for juveniles and official criminality for the subjects as adults. Dates, charges, courts, and sanctions have been included in the data bank for the longitudinal study. Massachusetts records have been supplemented by records from criminal histories agencies of several states in which some of the men are known to have resided.

In 1975, centralized records of the Massachusetts Department of Mental Health were searched to identify men from the study who had received diagnoses or treatment in public or private mental hospitals. Records of the Division of Alcohol were similarly searched to identify men from the study who had been treated for problems with alcohol. In 1977 and 1978, records on alcoholism were brought up to date through the cooperation of over 100 treatment clinics. From these records, histories of encounters with the law and of treatment for alcoholism or mental illness were constructed for the 506 men.

No attempt to find former members of the Youth Study had been made since 1948, when a sample had been interviewed. Between 1975 and 1979, 494 men (98%) were located. Almost a third were found through a telephone listing of the man, his wife, or a son living in the Boston area. Almost a third were found through locating a known relative who was living in the Boston area. For each of the remaining men, special search procedures were outlined. Death records were used to locate members of a "missing" family. City business directories were used to identify neighbors of prior addresses. Marriage records were used to retrace siblings; voter registration lists and driver's license applications were searched carefully. Telephone calls were made to people who shared the same surname as the subject or the maiden name of his mother.

Fifty-two men were reported dead. Death records confirmed the deaths of 50, 25 from the treatment group and 25 from the control group. Place, cause, and date of death have been coded.

Among the 494 men who were located, 377 (76%) were found in Massachusetts. These were almost equally divided between treatment and control group (189 of the former and 188 of the latter). The remaining men were found in 24 states, including Alaska, and in Germany. California, with 32 men, had attracted more migrants than had any other state.

Once contact was established, the men were asked to respond to a questionnaire. Questionnaires were sent to 421 men. Thirty-nine percent responded to the first request. Up to 8 requests were sent ($\bar{x}=2.2$), resulting in the return of completed questionnaires from 254 men (60%).

Some time after receiving a questionnaire (the length of time varied from a couple of weeks to several months), each man was asked to consent to an interview. Men were offered \$20 as payment for their time. Three hundred and seven men completed the interview, 73% of those whose addresses were known.

Questionnaires and interviews provided the data for the fourth perspective: that of the men themselves. Three hundred and forty-three men--75% of those not known to be dead actively contributed to the follow-up by responding to the questionnaire or completing the interview. Participants were almost equally divided between treatment and control groups, with 178 of the former and 165 of the latter contributing to the study.

The questionnaire requested information about marital status, occupation, geographic mobility, job satisfaction, political preference, smoking, and health. It included a short form of the F-scale measuring authoritarianism (from Sanford & Older, 1950), a scale measuring self-competence (from Douvan & Walker, 1956), questions on the use of leisure time and on the period of life considered best. The form sent to members of the treatment group asked how, if at all, the Cambridge-Somerville Youth Study had been helpful.

Like the questionnaire, the interview included information about marital status, occupation, and geographic mobility. The interview included modified versions of scales designed to measure differential responses to emotions (Tokar, Brunse, Steffler, & Napier, 1973), self-confidence (Rosenberg, 1965), and orientation toward achievement, affiliation, and sex roles. The interview also requested information about education, children, child-rearing behavior of the men, and memories of their own childhood. Through the interview, the men provided descriptions of their parents, their responses to frustrations, and their ideals and values. They also supplied brief reports on their own delinquency.

In the questionnaire and during the interview, men were asked to respond to four questions about drinking. The four questions, known as the CAGE test (Ewing & Rouse, 1970), asked whether the respondent had ever taken a morning eye-opener, felt the need to cut down on drinking, felt annoyed by criticism of this drinking, or felt guilty about drinking. A response which included at least three affirmative answers was considered evidence that a man was an alcoholic.

Information about occupation was available through several sources: interviews, questionnaires, voter registration lists, city business directories, and telephone contact with the men or their families. Three measures of occupational status were used to describe the 484 men (96%) whose occupations could be ascertained: a NORC score from June 1963 ratings (Hodge, Siegel, & Rossi, 1966), a Hollingshead Occupational Index rating (Hollingshead & Redlich, 1958), and an Occupational Status Score (Nam, LaRocque, Powers, & Holmberg, 1975). The marital status of 471 men (93%) was also recorded.

The Youth Study files included ratings of the neighborhoods in which the boys resided. These ratings had been made prior to the beginning of treatment and took into account delinquency rates, availability of recreational facilities, and proximity to bars, railroads, and junkyards. The files also included criminal records of the boys' parents and siblings; the latter had been collected through 1945. This information, too, was included with the longitudinal data.

Table 1 presents an outline of the information available.

Table 1 about here

Various indices of social problems indicated that the study consisted of a high-risk population. Over half (56%) had been convicted for at least one non-traffic offense; 30% had been convicted for a crime indexed by the FBI; 29% could be identified as alcoholics through their public records or their responses to the CAGE test; and 10% had been treated in mental hospitals for problems other than alcoholism.

On the other hand, the subjects included men who are successful in their own eyes or those of society. As revealed through interviews, 141 men completed high school. In responding to the questionnaire, 225 men reported that they were satisfied with their work and 208 reported satisfaction with how their lives have turned out. As compared with their fathers, more than three times the proportion of subjects are employed as white collar workers.

Results and Discussion

Data analyses have yielded evidence in three areas of investigation: effects of the treatment program, patterns of criminality, and the etiology of criminal behavior.

Since members of the treatment group had been selected by chance, with a matched "mate" assigned to the control group, the study provided an ideal model for assessing effects of the types of treatment provided between 1939 and 1945. Analyses of differences between the two groups led to a conclusion that the treatment program produced negative results (McCord, 1978). Two-thirds of the treatment group responding to the questionnaire claimed that the program had been helpful to them. Yet, as compared with men from the control group, those in the treatment group had a higher probability of showing signs of stress or dissatisfaction.

Several explanations for the adverse impact of the treatment program seemed plausible. These were checked by comparing treatment and control group responses to the questionnaire and interview. There was no evidence to support a view that counselors had created conflicts through introducing middle-class values into a lower-class environment: as compared with the control group, the treatment group was not more likely to have sought education, to emphasize education, or to use success as grounds for admiration. There was no evidence that the treatment program increased dependency: as compared with the control group, the treatment group was not more likely to spend leisure time in the company of others to ask for the opinion of others when making difficult decisions, to belong to clubs, or to visit their parents frequently. Nor was there evidence that treatment increased the likelihood of self-perception as sick or in need of help: as compared

with the control group, the treatment group did not lack self-confidence, have a lower perception of self-competence, or report greater use of medicines. There was, however, some evidence that the treatment program had introduced unrealistically high expectations, thereby producing conditions for disillusion. As compared with men from the control group, those from the treatment group appeared to be less satisfied with their jobs and with their marriages.

An anomaly led to another study of crime patterns in this urban population: prior literature seemed to show that parental rejection and conflict led to juvenile delinquency and that juvenile delinquency presaged adult criminality; yet studies of adult criminals failed to show that they had been rejected by their parents or exposed to parental conflict. If rejection and parental conflict cause delinquency and delinquents become adult criminals, there ought to be a strong relationship between rejection/conflict and adult criminality.

Evidence from the histories of antisocial behavior among the 506 former members of the Youth Study supported this explanation: Despite a statistically reliable association between juvenile delinquency and adult criminality, juvenile delinquents account for only a minority of men convicted as adults for crimes indexed by the FBI. Since a majority of juvenile delinquents had been convicted for indexed crimes as adults, "juvenile delinquents" and "adult criminals" refer to different pools of men. The findings suggest that there may be important differences between these groups in terms of the nature and meaning of criminal behavior. (See McCord, 1980)

The impact of parental behavior during childhood on subsequent criminality seems to be strong (McCord, 1979b). Use of six variables describing home atmosphere (mother's affection, supervision, parental conflict, parental aggression, mother's self-confidence, and father's deviance) resulted in a 21% improvement over chance identification and an 8% improvement over predictions of adult criminality based on the subject's juvenile criminal histories.

In another study of the etiology of criminal behavior, data from case records for the treatment group were used to discover the impact of being reared in a broken home (McCord, 1978b). Comparisons of criminal records indicated that the proportions of sons subsequently convicted for indexed crimes among those raised in broken homes by affectionate mothers and those raised by two parents in the absence of conflict were the same, whereas there was twice the proportion among boys raised in the midst of parental conflict.

Data from teachers' records, too, have been analyzed in relationship to subsequent criminality (McCord, 1977b). This study provided no support for "compensatory" theories about the development of antisocial behavior: early dependency, lack of self-confidence, difficulties with school work were not predictive of subsequent criminality.

Yet the descriptions of behavior collected in 1936-1937 suggest the possibility for distinguishing a "type" of child who could be considered at risk: boys who later were convicted for crimes had been described as blaming others for their mistakes, as resentful of authority, and as aggressive.

In sum, the data from this study of former members of the Cambridge-Somerville Youth Study provide evidence that there are identifiable conditions which are associated with increases in the probability of developing antisocial behavior. Perhaps more importantly, however, the data which span forty years provide a basis for critical evaluation of developmental and therapeutic assumptions.

References

- Douvan, E., & Walker, A. M. The sense of effectiveness in public affairs. Psychological Monographs, 1956, 70 (22): (Whole No. 429).
- Ewing, J. A., & Rouse, B. A. Identifying the "hidden alcoholic." Paper presented at the 29th International Congress on Alcohol and Drug Dependence, Sydney, New South Wales, Australia, 1970.
- Hodge, R. W., Siegel, P. M., & Rossi, P. Occupational prestige in the United States: 1925-1963. In Class, Status, and Power, edited by Reinhard Bendix and Seymour Lipset. New York: The Free Press, 1966.
- Hollingshead, A. B., & Redlich, F. C. Social Class and Mental Illness. New York: John Wiley, 1958.
- McCord, J. A comparative study of two generations of Native Americans. In Theoretical Concerns in Criminology, edited by R. F. Meier. Beverly Hills, Calif.: Sage Publications, 1977A.
- McCord, J. A life-history approach to criminal behavior. Paper presented at the American Society of Criminology Annual Meeting, Atlanta, Georgia, 1977.
- McCord, J. A thirty-year follow-up of treatment effects. American Psychologist, 1978, 33 (3).
- McCord, J. A longitudinal view of the relationship between parental absence and crime. Paper presented at the American Society of Criminology Annual Meeting, Dallas, Texas, 1978.
- McCord, J. Consideration of some effects of a counseling program. Paper presented at a meeting of the National Academy of Sciences, Woods Hole, Massachusetts, June 1979.
- McCord, J. Some child-rearing antecedents of criminal behavior in adult men. Journal of Personality and Social Psychology (in press), 1979.
- McCord, J. Patterns of deviance. In S. B. Sells, R. Crandall, M. Roff, J. D. Strauss, and W. Pollin, (eds.) Human Functioning in Longitudinal Perspective: Studies of Normal and Psychopathological Populations. Baltimore: Williams & Wilkins, 1980.
- Nam, C. B., LaRocque, J., Powers M. G., & Holmberg, J. Occupational status scores: stability and change. Proceedings of the American Statistical Association, 1975, pp. 570-575.
- Powers, E., & Witmer, H. An Experiment in the Prevention of Delinquency: The Cambridge-Somerville Youth Study. New York: Columbia University Press, 1951.
- Rosenberg, M. Society and the Adolescent Self-Image. Princeton: Princeton University Press, 1965.

Sanford, F. H., & Older, J. J. A short authoritarian-equalitarian scale. Institute for Research in Human Relations, Progress Report No. 6, Series A, 1950.

Tokar, J. T., Brunse, A. J., Steffler, V. J., & Napier, D. A. Emotional states and behavior patterns in alcoholics and non-alcoholics. Quarterly Journal of Studies on Alcohol, 1973, 34A (March): 133-134.

Table 1

Description of Data

<u>Type of Data</u>	<u>Source - When Collected</u>	<u>Who</u>
Description of boys	Teachers - 1936-1937	All subjects
Ratings of neighborhoods	Social Workers - 1939	All subjects
Delinquency prediction score	Selection Committee- 1939	All subjects
Descriptions of family life, culture, socio-economic factors, parents, and boys as teenagers	Case records - 1939-1945	Treatment Group
Description of treatment services and counselor-family interaction	Case records - 1939-1945	Treatment Group
Unofficial records of delinquencies of boys	Crime Prevention Bureau - 1948	All subjects
Criminal records of parents and siblings	Massachusetts Department of Probation - 1948	All subjects
Records of convictions, mental hospital and clinic diagnoses and placements	Massachusetts Department of Probation, Division of Alcohol, Department of Mental Health - 1975-1978	All subjects
History of treatment for alcohol problems since 1972	Alcohol Treatment Centers - 1976-1978	All subjects
Date and cause of death	Vital Statistics - 1975-1979	56 subjects
Self-competence scale, authoritarianism scale. Report on marital status, occupation, health, drinking, political attitudes, use of leisure time	Questionnaires - 1976-1979	254 subjects
Self-confidence scale. Response to emotions scale, Orientation toward achievement, affiliation, sex roles, family of orientation, family of procreation. Report on health, drinking, occupational and residential history, education	Interviews - 1976-1979	307 subjects

Table 1 (cont'd.)

Description of Data

<u>Type of Data</u>	<u>Source - when Collected</u>	<u>Who</u>
Supplementary data for occupation	Telephone calls, voter registration, and miscellaneous records - 1978-1979	484 subjects (141 additional)
Supplementary data for marital status	Telephone calls, voter registration, and miscellaneous records - 1978-1979	471 subjects (128 additional)

Longitudinal Data from the Seattle and Denver

Income Maintenance Experiments: Design and Uses*

Michael C. Keeley
SRI International

*This paper draws from previous papers by Keeley, Spiegelman and West (in press), Keeley (in press), and Spiegelman and Yaeger (in press). Arlene Waksberg provided valuable comments. The Seattle and Denver Income Maintenance Experiments were performed pursuant to contracts with the states of Washington and Colorado, prime contractors for the Department of Health, Education and Welfare under contract numbers HEW-100-79-0068 and HEW-100-78-004, respectively. The opinions expressed in the paper are those of the author and should not be construed as representing the opinions or policies of the states of Washington or Colorado or any agency of the United States Government.

Longitudinal Data from the Seattle and Denver
Income Maintenance Experiments: Design and Uses

Michael C. Keeley

SRI International

For many years there has been interest in replacing the existing complex transfer system in the United States with a nationwide negative income tax (NIT) program. The Seattle and Denver Income Maintenance Experiments (SIME/DIME) are the fourth in a series of income maintenance experiments that have been funded by the federal government in order to test alternative NIT programs. The first experiment, the New Jersey experiment, was conducted in New Jersey and Pennsylvania from 1968 to 1972. Subsequent experiments have taken place in Gary, Indiana from 1970 to 1974 and in rural areas of Iowa and North Carolina in 1969 to 1973. The Seattle and Denver experiments began in 1971 in Seattle, Washington and Denver, Colorado and came to a close in 1978, although analysis of the data is still ongoing.

These experiments were launched in order to provide data that would be used to analyze the behavioral effects and related costs of alternative nationwide negative income tax programs. Behavioral responses to the change in environment by the NIT are particularly important from a policy point of view because behavioral responses may have substantial effects on cost. Furthermore, behavioral responses are often of policy importance in their own right.

In the remainder of this paper we discuss both the design of the SIME/DIME experiments and the proper way in which the experimental data should be used. SIME/DIME data are a very rich source of information on individual behavior over time. The potential of these data, both for the analysis of behavioral responses to alternative NIT programs and for other types of micro-analysis is only now beginning to be realized.

The plan of the paper is as follows. In the next section the design of the experiment is discussed and the instrument and data collected

are described in the following section. Finally, a summary and conclusions are presented.

The Design of the Seattle and Denver Income Maintenance Experiments

Sample Design

SIME/DIME is the largest of all the NIT experiments with an original sample of approximately 4,800 urban families. This sample size exceeds the combined samples of all the other NIT experiments. The 4,800 families were selected from a screening of about 60,000 households in Seattle and Denver on the basis of characteristics required for the experiment (described later in this paper). These 4,800 families were assigned to one of eleven NIT programs, one of the manpower programs, or were assigned to the null or control program. To determine whether the duration of the experiment affects response, some families were enrolled in an experiment for 3 years, others for 5 years, and a very small sample were enrolled later for 20 years. About 60% of the sample families were enrolled in one of 11 NIT programs and about 40% were financial controls. In addition, about 60% of the families received one of the three manpower treatments. In Table 1, the composition of the sample at enrollment by site, financial/control status, and family type is described. Details of the design are discussed in Muraka and Spiegelman (1978).

Table 1

Numbers of Financial and Control Families

By Family Structure and Site

	One-Parent Families	Two-Parent Families	Total	
Seattle				
Financials	492	615	1,107	54%
Controls	378	557	935	46%
TOTAL	870	1,172	2,042	100%
Denver				
Financials	726	914	1,640	59%
Controls	435	683	1,118	41%
TOTAL	1,161	1,597	2,758	100%

The SIME/DIME sample is not intended to be representative of the United States population as a whole. Instead, it is limited to two sites (Seattle and Denver); families with two heads or single headed families with at least one dependent; male or female heads, between 18 and 58 years of age; family income, adjusted by family size, less than \$9,000 per family of four with one working head and less than \$11,000 per family of four with two working heads (in 1971); and race or ethnic groups of white, black and Chicano. In addition to being stratified by race and family type, the sample is stratified by income level adjusted by family (7 different income strata are used) in order to insure that adequate numbers of low income families are enrolled. These income strata are supposed to approximate permanent income adjusted for family size.

The Treatments

Eleven different NIT treatments or combinations of tax rates and support levels are being tested (see Kurz & Spiegelman, 1972, for a detailed discussion of the treatments). The support level is the grant the family receives if it has no other income and the tax rate is the rate at which the grant declines as other income increases. Programs with support levels of \$3,800; \$4,800; and \$5,600 normalized for a family of four persons at 1970-71 prices are being tested. The real values (adjusted for inflation) of the support levels have been kept approximately even over time by the use of an automatic cost of living escalator which is adjusted every quarter. The following figures show the nominal support levels for a family of four at the end of 3 and 5-year programs in Seattle.

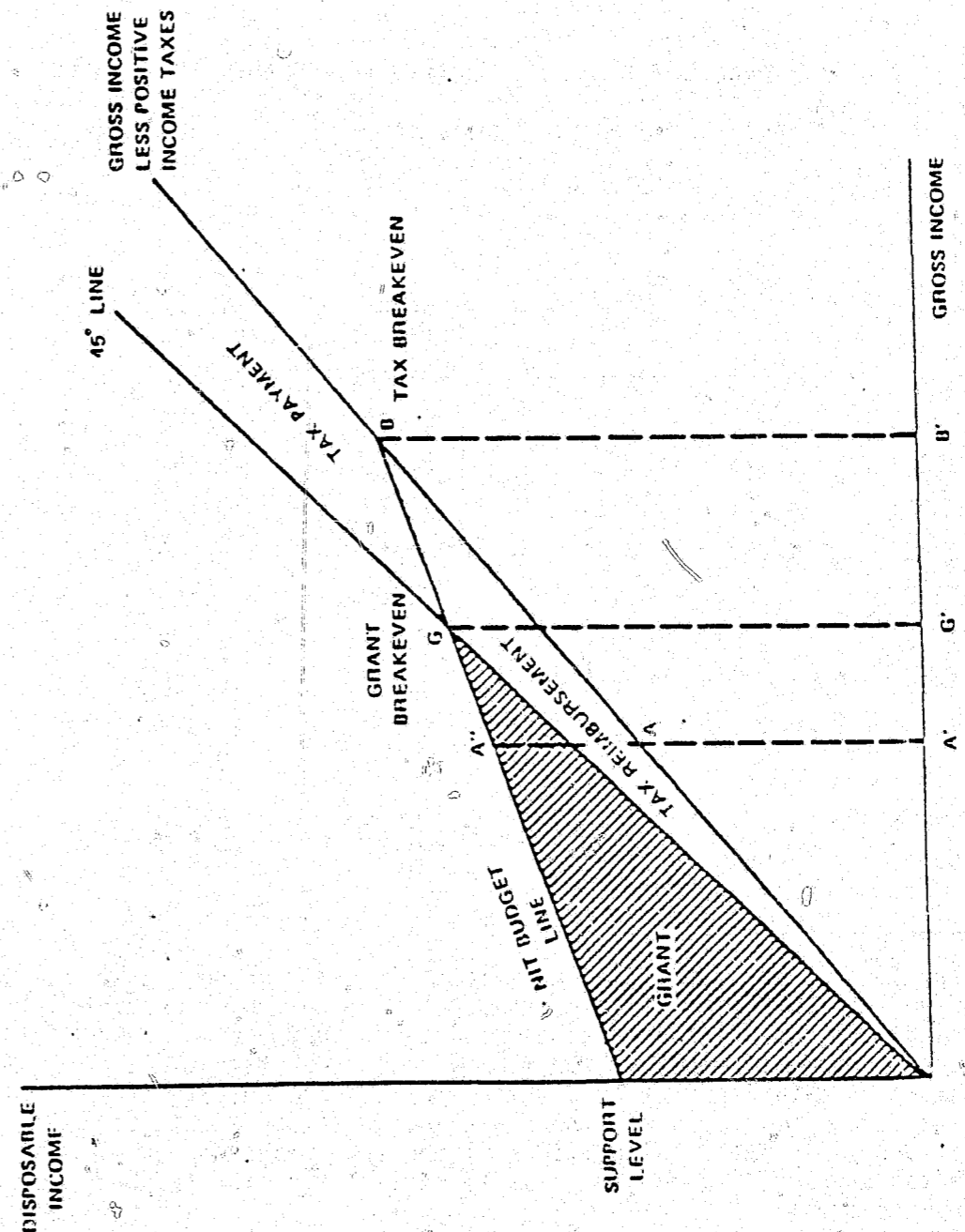
Beginning of the Experiment April, 1971	April 1974	April 1976
\$3,800	\$4,320	\$5,200
\$4,800	\$5,460	\$6,560
\$5,600	\$6,360	\$7,650

In addition, four different NIT tax systems are being tested; two with tax rates of 50 and 70 percent of earned income, and two with declining tax systems in which the tax rate declines as income increases. A declining tax is one in which the marginal and average tax rate on income declines as income increases. A declining tax system (relative to a constant tax system) leads to corner solutions. That is, if persons tend to work very little they are apt to be induced to drop out of the

labor force and if they tend to work a lot then the declining tax system provides less work disincentive than the constant tax system (with the same breakeven level). In addition, the declining tax system can be easily integrated into the positive income tax system without discontinuity in the marginal tax rate at the point where an individual goes from negative to the positive tax system. However, with a declining tax system the tax rate is an endogenous (choice) variable that depends on income. This endogeneity has important implications for the research. Namely, simultaneous equation techniques must be used to analyze the structural effects of these programs.

In addition, experimental training and education subsidies are being tested through the assignment of three different manpower programs. The manpower programs are intended to improve job search and increase wage rates, thereby increasing the value of market work. Manpower programs are conducted at three levels: M1, M2, and M3. M1 provides only counseling services. M2 provides counseling services and subsidizes 50% of the direct costs of any training taken over the life of the experiment, and M3 provides counseling plus 100% subsidy of the direct costs of any training over the life of the experiment.

Figure 1 shows how the experimental NIT program functions. The horizontal axis shows gross income (income before taxes or transfer payments such as welfare); the vertical axis shows disposable income (income after taxes and with transfer payments). At zero gross income, the NIT grant is equal to the support level. Without the NIT the family with gross income A' would have disposable income A after paying positive income taxes; with the NIT payments that family would have disposable income of A'. Notice that the NIT payment has two components, a grant and a reimbursement of positive income taxes. At gross income level G' the NIT grant has declined to zero. The family still receives benefits in the program and is receiving reimbursement of positive income taxes. Positive income taxes are reimbursed under the NIT program to eliminate at other than those imposed by the NIT. Between the gross income levels of G' and B' the family still benefits from the NIT program by receiving partial reimbursements of positive income taxes. Families with incomes above the tax breakeven level B' do not receive any benefits from the NIT program. Table 2 presents the NIT plans being tested in SIME/DIME along with the grant and tax breakeven levels of the programs.



NOTE: Figure assumes no income outside of earnings and a linear positive income tax system.

FIGURE 1 A NEGATIVE INCOME TAX PROGRAM WITH POSITIVE TAX REIMBURSEMENT

Table 2
Plan Breakeven Level for the Seattle and Denver
Income Maintenance Experiments
(1971 Dollars)

Plan	Grant Breakeven Level	Tax Breakeven Level
F1 (S = 3800, $t_e = .5$, $r = 0$)	\$ 7,600	\$10,250
F2 (S = 3800, $t_e = .7$, $r = 0$)	5,429	6,350
F3 (S = 3800, $t_e = .7$, $r = .025$)	7,367	10,850
F4 (S = 3800, $t_e = .8$, $r = .025$)	5,802	7,800
F5 (S = 4800, $t_e = .5$, $r = 0$)	9,600	13,150
F6 (S = 4800, $t_e = .7$, $r = 0$)	6,867	8,520
F7 (S = 4800, $t_e = .7$, $r = .025$)	12,000	19,700
F8 (S = 4800, $t_e = .8$, $r = .025$)	8,000	11,510
F9 (S = 5600, $t_e = .5$, $r = 0$)	11,200	15,700
F10 (S = 5600, $t_e = .7$, $r = 0$)	8,000	9,780
F11 (S = 5600, $t_e = .8$, $r = .025$)	10,360	16,230

Note: These figures are for a family of four with only one earner and no income outside of earnings. Positive tax reimbursements include the federal income tax and Social Security taxes. The federal income tax assumes the family takes the standard deduction. State income taxes, which are relevant only for the Denver Experiment (there is no state income tax in Washington), are ignored in these calculations of the tax breakeven levels. The tax breakeven levels are thus slightly higher for the Denver Experiment.

Key: S = NIT annual support level

t_e = initial NIT tax rate

r = rate of decline of the average NIT tax rate per thousand dollars of income.

Assignment of Treatment

In all of the NIT experiments, including SIME/DIME, assignment to the experimental treatment plans or to control status is not done on a simple random basis. Instead a mathematical model known as the Conlisk-Watts (Conlisk-Kurz in the case of SIME/DIME) allocation model is used to determine the allocation of experimental treatments (including the control or null treatment) based on the family's race, normal income level, and number of heads.

The intention of this model is to increase the number of observations possible with a given budget so as to maximize the informational content of the sample by accounting for the fact that the cost of an observation varies systematically with both household and treatment characteristics.

In principle, the informational content of an experimental sample selected with a given budget can be increased (compared to a simple random assignment) by accounting for the fact that different treatments have different costs. For example, if the NIT treatment observations are more expensive than control observations, then statistical considerations indicate that the ratio of treatment to control observations should equal the square root of the ratio of control to treatment costs. If treatment observations are four times as expensive as controls, two-thirds of the sample should be assigned to control status and one-third to experimental treatment. This basic idea is generalized in the NIT experiments to the case of many treatments and to the case where treatment cost depends on family characteristics.

In particular, the higher the income of the family, the less costly the expected NIT payment for a given treatment. In fact, a family with a high enough income to put it sufficiently above the breakeven level, has an NIT payment of zero. Similarly, for a family with a given income, NIT programs with higher support levels and lower tax rates are more expensive depending on response.

The use of this assignment model results in an experimental sample in which experimental treatments (including the control treatment) are not independent of pre-experimental household characteristics (in particular normal income). This non-independence between pre-experimental household characteristics and experimental treatment has several important implications for measuring and interpreting experimental response which are discussed below.

The assignment model used in SIME/DIME results in a sample in which families with low normal income levels are more likely to be assigned to the programs with low support (and breakeven) levels. Persons with high normal income levels are more likely to be assigned to programs with high support levels. With regard to control status, low normal income families are less likely and high normal income

families are more likely to be assigned to control status. Thus, the distribution of income differs in different treatment categories and the mean income of controls differs from the mean income of persons assigned to the experimental programs.

Design of the Instrument and Data Available

All families enrolled in the experiment were given a series of in depth interviews in the respondents' home (see Davis & Waksberg, forthcoming for further discussion of data collection). All families received a pre-enrollment interview conducted prior to the family's actual enrollment in the experiment. An enrollment interview was given at the time of the enrollment in the experiment (which retrospectively covered the year prior to enrollment) and periodic interviews were administered approximately every four months throughout the duration of the experiment. All families in the 3-year group received periodic interviews for 3 experimental years and 1 post-experimental year, and half of these families were interviewed for 2 post-experimental years. Thus, for all families in the 3-year group, there are at least 5 years of longitudinal data, and for some, 6 years not including the pre-enrollment interview. All families in the 5-year group have at least 7 years of longitudinal data available, not including the pre-enrollment interview. Currently, data from the pre-enrollment interviews are available for the Denver subjects only, although there are plans to make the Seattle pre-enrollment interview available. In addition, the small sample of 20-year families were all interviewed for 6 experimental years.

Data collection was intended to support a broad range of economic and sociological research studies. Longitudinal data were collected on such topics as family composition, including changes in marital status, fertility, and movements in and out of the family, employment history, including wage rates, hours of work and other characteristics of each job held; nonwage income; assets; training and education undertaken; and other social, demographic and economic data. A unique feature of the data is its continuous time nature. That is, an attempt was made to record these longitudinal variables for each individual at each moment in time. This greatly facilitates new kinds of research that model the rates of transition from one type of activity to another, such as from employment to unemployment or from one type of job to another or from one marriage to another.

In addition to the set of longitudinal continuous-time data, point in time questions were asked about such diverse topics as health, job satisfaction, education of children, mental health status, community ties, ethnicity, work attitudes, and family roles. Each periodic interview consists of a basic core section, in which longitudinal continuous

time data were collected primarily on employment and earnings, and a series of non-core modules.

The longitudinal employment and earnings data are probably the most comprehensive and detailed data on work force behavior ever collected. Each instant in time for each individual is accounted for in these data. Wage rates, hours of work, type of job, and detailed data on each non-work period and each period of looking for work are obtained.

Each non-core module consists of a series of questions on a specific topic. Each was designed to be repeated either within a single interview with more than one individual within a family or in later interviews. Some modules were repeated at every interview to cover dynamic situations and other modules covering topics that changed less rapidly were repeated less frequently, generally yearly. The aim of this module was to provide efficiency in constructing the interviews and consistency in the data. A list of the general topics covered in the non-core materials is presented in Table 3.

Table 3

Non-core Data Modules

Economic Variables

Net Worth
Real Property
Job Training History
Housing/Rent
Point in Time Work History
Employment History
Sale of Real Property
Non-durable Consumption
Perceived Wage Rate of Non-workers
Children's Paying Jobs
Job Satisfaction and Stability
Job Counseling

Education

Education History
Educational and Occupational Expectations
for Children
Educational Status of Children
Occupational/Education of Parents
Children's School Attendance

Attitudes

Family Opinions
Female Work Attitudes
Attitudes Toward Training
Attitudes Toward Use of Credit
Husband's Attitude Toward Wife's Working
Attitude Toward Welfare
Social Ethics of Work
Time Horizon and Planning
Budget Attitudes
Self Perception of Socioeconomic Status

Health/Intelligence

Mental Health Status
Health Status
Verbal Awareness
Family Arguments

Demographic Variables

Ethnicity
Mobility History
Fertility History
Marital Status and History

Home Activities/The Family

Family Decisions
Family Roles
Family Leisure Time
Utilization of Child Care
Household Chores
Home Environment
Community Ties

In addition to the pre-enrollment interview, the enrollment interview and the periodic interviews, there was a post-move interview that was used to collect data for families that had moved away from Seattle and Denver. Due to cost considerations these families were interviewed only once a year. However, in principle the post-move interview collected all the information that would have been collected in regular periodic interviews, if the family had not moved. The post-move interview was also used as a catch-up interview for families who missed one or more regular periodic interviews but resided within the site area.

In addition to the actual data available in the interviews, a master file system was created to track families. The master file is a key source of analytical information on such variables as divorce, marriage and remarriage, fertility, and changes in family composition. The master file system contains master records for each family and for each individual known to the experiment. These records contain treatment information about the family (support level, labor force participation level, enrollment date), and descriptive information about each individual (birthdate, sex, race, etc.). The master file also contains information documenting each interview, each family composition change and the relationships of family members in predecessor and successor families. Thus the complete history of families and individuals in the experiment can be readily traced.

The SIME/DIME data are an enormous body of data containing a variety of detailed information about each family and each individual within a family over a relatively long period of time. Periodic interviews were occasionally as long as 300 pages and were administered approximately three times a year to each family. This is probably the most comprehensive set of longitudinal data ever collected on so large a number of families. The possibilities for analysis are virtually unlimited. However, researchers must be aware of the limitations of experimental data, namely the nonrepresentativeness of the samples and the non-random assignment of actual treatments which are two important features that are not characteristic of most other longitudinal data sets, such as the National Longitudinal Survey or the Michigan Panel Study.

Research Completed Using Data from the Seattle and Denver Income Maintenance Experiments

There have been over 65 research memoranda that contain analyses of SIME/DIME. A list of research memoranda follows this chapter. These research memoranda are available upon request (for a nominal fee) from SRI International. In addition there have been a large number of professional journal articles reporting results of these experiments. The topics include studies of labor supply response, effects on the rate of unemployment and employment, child care utilization, wage rates, job satisfaction, migration, schooling and training, housing, fertility, and family stability. As is evident from the list of research memoranda, the range of research is far too broad for a comprehensive review in the limited space available here.

Summary and Conclusions

Social experimentation is a powerful research tool that is being used to provide estimates of the behavioral responses and cost of alternative NIT programs. The Seattle and Denver Longitudinal data provide a particularly rich source of information regarding a wide range of behavioral responses. The potential of these data is just now beginning to be realized. Although experimental data have many advantages over nonexperimental data, a different methodology must be used to analyze the response to experimental programs when using longitudinal experimental data for nonexperimental research purposes. This is largely because the experimental samples are truncated and are highly stratified and because experimental treatment is assigned only on a stratified random basis.

The primary advantages of experimental longitudinal data are that (1) they allow the researcher to distinguish inter-individual difference and changes in a given individual's behavior over time, and (2) there are exogenous measurable changes in individuals' budget constraints. Thus, permanent and transitory components of behavioral response can be isolated and the casual effects of the treatments can be separated from the effects of other variables.

Overall, the longitudinal data from the Seattle and Denver Income Maintenance Experiment represent a valuable resource. These data provide the opportunity to analyze a variety of behaviors for a very large group of people over a relatively long period of time. Furthermore, the exogenously determined treatments make it possible to address an entirely new set of research questions. It is hoped that this paper will contribute to the use of these data.

SRI SIME/DIME Research Memoranda*

Research Memorandum Number	Title and Authors
15	<u>The Assignment Model of the Seattle and Denver Income Maintenance Experiments.</u> J. Conlisk and M. Kurz, July, 1972.
18	<u>The Design of the Seattle and Denver Income Maintenance Experiments.</u> M. Kurz and R.G. Spiegelman, May 1972.
19	<u>The Payment System for the Seattle and Denver Income Maintenance Experiments.</u> M. Kurz, R.G. Spiegelman, and J.A. Brewster, June, 1973.
21	<u>The Experimental Horizon and the Rate of Time Preference for the Seattle and Denver Income Maintenance Experiments: A Preliminary Study.</u> M. Kurz, R.G. Spiegelman, and R.W. West, November, 1973.
22	<u>Social Experimentation: A New Tool in Economic and Policy Research.</u> M. Kurz and R.G. Spiegelman, November, 1973.
23	<u>Measurement of Unobservable Variables Describing Families.</u> N.B. Tuma, R. Cronkite, D.K. Miller, and M. Hannan, May, 1974.
24	<u>A Cross Section Estimation of Labor Supply for Families in Denver 1970.</u> M. Kurz, P. Robins, R.G. Spiegelman, R.W. West, and H. Halsey, November, 1974.
25	<u>Job Search: An Empirical Analysis of the Search Behavior of Low Income Workers.</u> H.E. Felder, May, 1975.
26	<u>Measurement Errors in the Estimation of Home Value.</u> P. Robins and R.W. West. June, 1975.

*Research Memoranda 1 through 14, 16, 17 and 20 are obsolete and are not available for distribution. Other Research Memoranda, through No. 65 are available upon request for \$3.00 charge per copy from the Socioeconomic Research Center, SRI International, 333 Ravenswood Avenue, Menlo Park, California 94025

- 27 A Study of the Demand for Child Care by Working Mothers. M. Kurz, P. Robins and R.G. Spiegelman, August, 1975.
- 28 The Impact of Income Maintenance on the Making and Breaking of Marital Unions: Interim Report. M. Hanna, N.B. Tuma and L.P. Groeneveld, June, 1976.
- 29 The Estimation of Labor Supply Models Using Experimental Data: Evidence from the Seattle and Denver Income Maintenance Experiments. M.C. Keeley, P.K. Robins, R.G. Spiegelman and R.W. West. August, 1976.
- 30 Determinants and Changes in Normative Preferences of Spouses. R.C. Cronkite, May, 1977.
- 31 Homogamy, Normative Consensus, and Marital Adjustment. R.C. Cronkite, May, 1977.
- 32 The Determinants of Participation of Single-Headed Families in the AFDC Program. Arden Hall, May, 1977.
- 33 The Supply of Day Care Services in Denver and Seattle. Arden Hall and Sam Weiner, June, 1977.
- 34 The Impact of Income Maintenance and Manpower Subsidies on the Decision to Invest in Human Capital: Interim Results from the Seattle and Denver Income Maintenance Experiments. H.E. Felder, A. Hall, Y. Weiss, October, 1977.
- 35 First Dissolutions and Marriages: Impacts on 24 Months of the Seattle and Denver Income Maintenance Experiments. N.B. Tuma, L.P. Groeneveld and M.T. Hanna, August, 1976.
- 36 The Estimation of Nonlinear Labor Supply Functions with Taxes from a Truncated Sample. Michael Hurd, November, 1976.
- 37 The Welfare Implications of the Unemployment Rate. Michael Hurd, November, 1976.
- 38 The Labor Supply Effects and Costs of Alternative Negative Income Tax Programs: Evidence from the Seattle and Denver Income Maintenance Experiments, Part I: The Labor Supply Response Function. M.C. Keeley, P.K. Robins, R.G. Spiegelman and R.W. West, May, 1977.

- 39 The Labor Supply Effects and Costs of Alternative Negative Income Tax Programs: Evidence from the Seattle and Denver Income Maintenance Experiments, Part II: National Predictions Using the Labor Supply Response Function. M.C. Keeley, P.K. Robins, R.G. Spiegelman and R.W. West, May, 1977.
- 40 Cost of Compliance with Federal Day Care Standards in Seattle and Denver. Sam Weiner, May, 1977.
- 41 An Interim Report on the Work Effort Effects and Costs of a Negative Tax Using Results of the Seattle and Denver Income Maintenance Experiments: A Summary. M.C. Keeley, P.K. Robins, R.G. Spiegelman and R.W. West, June, 1977.
- 42 The Reporting of Income to Welfare: A Study in Accuracy of Income Reporting. H.I. Halsey, M. Kurz, R.G. Spiegelman and A. Waksberg, August, 1977.
- 43 Variation Over Time in the Impact of the Seattle and Denver Income Maintenance Experiments on the Making and Breaking of Marriages. N.B. Tuma, M.T. Hannan and L.P. Groeneveld, February, 1977.
- 44 A Model of the Effect of Income Maintenance on Rates of Marriage Dissolution: Evidence from the Seattle and Denver Income Maintenance Experiments. M.T. Hannan, T.B. Tuma, and L.P. Groeneveld, February, 1977.
- 45 Job Satisfaction and Income Maintenance: Evidence from the Seattle and Denver Income Maintenance Experiments. P.K. Robins. October, 1977.
- 46 Unemployment Insurance, Wage Changes and Search Behavior: An Analysis. H.E. Felder, October, 1977.
- 47 Impact of Income Maintenance on Geographical Mobility: Preliminary Analysis and Empirical Results from the Seattle and Denver Income Maintenance Experiments. M.C. Keeley, October, 1977.
- 48 Changes in Rates of Entering and Leaving Employment Under a Negative Income Tax Program: Evidence from the Seattle and Denver Income Maintenance Experiments. P.K. Robins and N.B. Tuma, March, 1977.

9
0
1
The Impact of Income Maintenance on Fertility: Preliminary Findings from the Seattle and Denver Income Maintenance Experiments. P.K. Robins and N.B. Tuma, March, 1977.

Income and Psychological Distress: Evidence from the Seattle and Denver Income Maintenance Experiments. P. Thoits and M. Hannan, February, 1978.

The Rate of Time Preference of Families in the Seattle and Denver Income Maintenance Experiments. R.W. West, March, 1978.

AFDC, Food Stamp, and Public Housing Taxes in Seattle and Denver in 1970-1971. H.I. Halsey, March, 1978.

Participation in the Seattle and Denver Income Maintenance Experiments and its Effects on Labor Supply. P.K. Robins and R.W. West, March, 1978.

The Impact of Income Maintenance on the Utilization of Subsidized Housing. M. Avrin, July 1978.

The Effective Income Tax: Evidence from the Seattle and Denver Income Maintenance Experiments. H.I. Halsey, July, 1978.

The Effects of Negative Income Tax Programs on Marital Stability: A Summary and Discussion of Results from the Seattle and Denver Income Maintenance Experiments. M.T. Hannan, N.B. Tuma and L.P. Groeneveld, October, 1978.

The Design of Social Experiments: A Critique of the Conlisk-Watts Assignment Model. M.C. Keeley and P.K. Robins, December, 1978.

Dynamic Analysis of Marital Stability. N.B. Tuma, M.T. Hannan and L.P. Groeneveld, December, 1978.

A Longitudinal Analysis of the Labor Supply Response to a Negative Income Tax Program: Evidence from the Seattle and Denver Income Maintenance Experiments. Philip K. Robins and Richard W. West, December, 1978.

The Effects of the Seattle and Denver Income Maintenance Experiments of the Labor Supply of Young Nonheads. Richard W. West, May, 1979.

-834-

61 The Impact of the Seattle and Denver Income Maintenance Experiments on Wage Rates: An Interim Analysis. Richard W. West, May, 1979.

62 A Preliminary Analysis of the Effects of the Seattle and Denver Income Maintenance Experiments on the Choice of Occupation. R.W. West, May, 1979.

63 Income and Independence Effects on Marital Dissolution: Results from the First Three Years of SIME/DIME. M.T. Hannan, N.B. Tuma and L.P. Groeneveld, May 1979.

64 The Destination Choices and Earnings of Migrants: The Impact of Alternative Negative Income Tax Programs. M.C. Keeley, May, 1979.

65 Taxes, Transfers, and Subsidies and the Demand for Children: The Impact of Alternative Negative Income Tax Programs. M.C. Keeley, May, 1979.

Technical
Memorandum
Number

1

Sample Selection in the Seattle and Denver Income Maintenance Experiments. B.A. Murarka and R.G. Spiegelman, July, 1978.

-835-

END

References

Davis, V. & Waksberg, A. Data collection and processing. In R.C. Spiegelman, S. Weiner & J.G. Bell (Eds.), A guaranteed annual income: Results from an income maintenance experiment, in press.

Hannan, M.T., Tuma, N.B. & Groeneveld, L.P. Income and marital events: Evidence from an income-maintenance experiment. The American Journal of Sociology, 1977, 4, 65-70.

Hausman, J.A. & Wise, D.A. The evaluation of results from truncated samples: The New Jersey income maintenance experiment. Annals of Economic and Social Measurement, 1976, 5,(4), 421-445.

Keeley, M.C. Labor supply and public policy: A critical review. Academic Press, in press.

Keeley, M.C. Migration as consumption: The impact of alternative negative income tax programs. In J. Simon & J. DaVanzo (Eds.), Research in population economics, Volume 2, JAI Press, 1979.

Keeley, M.C. The effects of an NIT on fertility. The Journal of Human Resources, manuscript submitted for publication.

Keeley, M.C. & Robins, P.K. The design of social experiments: A critique of the Conlisk-Watts assignment model. In R.G. Enrenberg (Ed.), Research in labor economics, Volume 3, JAI Press, in press.

Keeley, M.C., Robins, P.K., Spiegelman, R.G. & West, R.W. The estimation of labor supply models using experimental data. The American Economic Review, 1978, 68(5), 873-887.

Keeley, M.C., Spiegelman, R.G. & West, R.W. Design of the income maintenance experiments and an overview of the results. In R.G. Spiegelman, S. Weiner & J.G. Bell (Eds.), A guaranteed annual income: Results from an income maintenance experiment, in press.

Kurz, M. & Spiegelman, R.G. The design of the Seattle and Denver income maintenance experiments, (SRI International Research Memorandum No. 18), Stanford, CA.: July, 1978.

Murarka, B.A. & Spiegelman, R.G. Sample selection in the Seattle and Denver income maintenance experiments, (SRI International Technical Memorandum No. 1), Stanford, CA.: July, 1978.

Robins, K., Tuma, N.B. & Yaeger, K.E. Effects on job turnover. The Journal of Human Resources, in press.

-836-

Spiegelman, G. & Yaeger, K.E. An overview of the Seattle and Denver income maintenance experiments. The Journal of Human Resources, in press.

Weiss, Y., Hall, A. & Doug, F. Schooling effects. The Journal of Human Resources, in press.

West, R. Effects on wage rates. The Journal of Human Resources, in press.

83