THE IMPACT OF CRIMINAL COURT SENTENCING DECISIONS AND STRUCTURAL CHARACTERISTICS

Martin A. Levin

Brandeis University

March 1973

This project was supported by NI-70-065-PG-19 awarded by the Law Enforcement Assistance Administration, U.S. Department of Justice, under the Omnibus Crime Control and Safe Streets Act of 1968, as amended. Points of view or opinions stated 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.

Short Term Stay Only (contd)

Name	Age On Entry	Court	Offense	Alternative If PORT Were Not Available	Date of Entry To PORT	Present Progress & Prognosis
19. VH	14	Juv.	Home Incorrigibility	Continue Probation	4/72	Short term stay at own request; returned home with parents' consent.
20. KT	15	Juv.	Home Incorrigibility, Run-away	State Commitment	7/72	Withdrew satisfactory 8/20; AWOL. Referred himself to residential treatment center.
21. TM	16	Juv.	Burgl.	State Commitment	8/72	Withdrew AWOL 9/72
22. KH	16	Juv.	Run-away	Continue Probation	9/72	Self referral; returned home with parents' agreement.

ABSTRACT

THE IMPACT OF CRIMINAL COURT SENTENCING DECISIONS AND STRUCTURAL CHARACTERISTICS

This is a preliminary effort to analyze some problems in the evaluation of policy impact in a general framework and in the specific context of the evaluation of the impact of alternative criminal court policies, especially on recidivism.

The nature of policy evaluation presents major problems: A particular policy output is often only one of a range of causal factors affecting the behavior at which the policy is directed and many of them are highly correlated. Because it is difficult to measure the precise contribution of each factor one cannot determine precisely for what reasons a particular program has the effect that it does. An alternative response to these problems of policy evaluation is controlled experimentation which surmounts several of these difficulties, but it has its own disadvantages including experimenter effects such as the "Hawthorne" effect and those producing self-fulfilling prophecies and labeling.

Critical analysis of non-experimental evaluations of the factors shaping recidivism, including alternative sentencing policies, indicates that on the whole the type of treatment has a major impact. Those offenders who are granted probation generally have significantly lower rates of recidivism than those who have been incarcerated. This pattern tends to persist when offender characteristics and type of offense are controlled. The analysis also indicates that other factors such as type of offense, prior record, race, age, and narcotics history have a major impact. However, with a few exceptions the type of treatment prescribed by the judge seems to have a greater impact than these characteristics or the type of offense.

However, ~~lower recidivism rates for~~ those granted probation, even when other facts are controlled, does not necessarily indicate that lower rates are a specific function of this type of treatment. This relationship may be largely an artifact of the court's decision-making. Judges may grant probation to persons who they determine have favorable characteristics on the basis of the lack of seriousness of the crime and the judge's perception of the offender. Thus it is possible that it is the offender's characteristics rather than the particular treatment which is the primary influence on recidivism.

Nevertheless, for the purpose of policy evaluation and prescription, this possibility may not be fully relevant. It is insufficient to simply ascertain which factor is the "best predictor". Policy-makers need information about the explanatory factors over which they have some control. These factors will probably give the policy-maker greater ability to affect the outcome than he might otherwise have. A judge, or any other policy-maker, can do little to change an offender's age or his number of prior convictions, but he can prescribe the precise type of treatment (probation or incarceration) which he will receive. The requirements of analysis are different for pure social science and applied or policy social science.

An alternative method of analyzing the relationship between type of treatment and recidivism is conducting a controlled experiment. The California Youth Authority's "Community Treatment Project" experiment randomly assigns, after an initial screening, convicted juveniles either to an experimental group which receives probation and intensive counselling or to a control group which is incarcerated. After a follow-up period of twenty-four months the "failure" rate for the experimental group was 38% and 61% for the control group.

However, the CTP experiment was flawed in several respects. For example, initial screening eliminated about 25 percent of the convicted male juveniles (and 10 percent of the females) for whom institutionalization was deemed mandatory because they were involved in serious cases or because there was community objection. This limits the breadth of the conclusions that can be drawn. The data indicate that recidivism is less likely if the offenders receive probation, but we do not know if this applies to the most serious offenders.

Also, because of the nature of the intensive and special attention and supervision of the experimental group it seems very possible that its lower "failure" rate is to some degree a function of experimenter effects such as the Hawthorne effect and the effects of a positive self-fulfilling prophecy and positive labeling.

Nevertheless, since the requirements of analysis seem to be different for pure social science and applied or policy social science, these scientifically flawed aspects of the CTP offer promising possibilities for the goal of policy evaluation and prescription: (1) Screen out cases with assaultive backgrounds, and then grant probation to all other juvenile first offenders. The CTP indicates that probation leads to less recidivism; thus probation for all but those screened out by the above criteria should significantly lower the present "failure" rate. (2) The CTP's low "failure" rate for the experimental group can be utilized as an explicit and intentional positive policy. After an initial screening out of assaultive cases, perhaps all juvenile first offenders should be granted probation and assigned to an explicitly and intentionally "Hawthorne" and "positive self-fulfilling prophecy" program. The community agents would intentionally have expectations of "success" for these youths, who also would be positively labeled (as being an "experimental" participant). The CTP findings indicate that this could significantly lower the present overall "failure" rate (i.e., the combined rate for both offenders who are granted probation without a special program and those who are incarcerated).

In a preliminary fashion these findings and prescriptions are applied to the criminal courts of Pittsburgh and Minneapolis. This effort indicates that to convert these findings into policy guidance for criminal court judges, the many goals of the criminal court, in addition to reduced recidivism, must be considered. There is, however, a great deal of tension among these goals, and there are many factors which affect them which probably are beyond the reach of the courts.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

I wish to thank the following individuals for many helpful suggestions during the preliminary stages of this study: James Q. Wilson, Martin Shapiro, Eugene Bardach, Frank Levy, Sheen Kassouf. Michael Furstenberg and James Wexler provided invaluable aid as research assistants during the final stages of the study.

SUMMARY

Until recently most public discussion of crime prevention has emphasized the highly visible and dramatic role of the police. The tragedy of Attica, the disturbances in many other prisons, and the high rate of recidivism seem to indicate that prisons are imperfectly fulfilling their function. In all probability then, the police have less impact on the extent of crime than do the decisions of judges to sentence convicted criminals to prison or to let them remain free on probation. This then is a preliminary effort to analyze some problems in the evaluation of policy impact both in a general framework and in the specific context of the evaluation of the impact of alternative criminal court policies, especially on recidivism.

The nature of policy evaluation presents major problems: A particular policy output is often only one of a broad range of causal factors affecting the behavior at which the policy is directed and many of them are highly correlated. (For example, sentencing decisions are at the most, just one of the many causal factors of recidivism. One must consider what job opportunities are made available, counseling, home life, etc). It is thus difficult to measure the precise contribution of each factor because of the problem of multicollinearity. Evaluations that conclude that the programs make no difference, may be largely the result of the problem of multicollinearity: The program may have an effect, but it may not be possible to precisely measure it as distinguished from the effect of other interrelated factors. However, the frequent absence of a satisfactory solution to the multicollinearity should lead one to be more cautious in rejecting a policy program.

There are two major responses to this problem in policy evaluation. The first is the use of regression analysis, which is designed to separate

effects in multivariate situations. Regression analysis holds several variables constant to ascertain the independent effect of another variable. If, however, the variables are highly correlated then it is not useful. This problem is multicollinearity in its extreme.

An alternative response to these problems of policy evaluation is controlled experimentation which surmounts several of these difficulties, but it has its own disadvantages. They suffer from such experimenter effects as the "Hawthorne" effect (the subject acts differently solely because they are in an experiment) and those producing self-fulfilling prophecies and labeling are demonstrated in Robert Rosenthal's laboratory and educational investigations. A third methodological and interpretive problem is the initial selection of the population from which the control and experimental groups are chosen. This problem of selection becomes a constraint on the generality of the conclusions drawn from the experiment.

Critical analysis of non-experimental evaluations of the factors shaping recidivism, including alternative sentencing policies, indicates that on the whole the type of treatment has a major impact. On the whole those offenders who are granted probation generally have significantly lower rates of recidivism than those who have been incarcerated. This pattern generally tends to persist when offender characteristics and type of offense are controlled. The analysis also indicates that other factors such as type of offense, prior record, race, age, and narcotics history also have a major impact. However, with a few exceptions the type of treatment prescribed by the judge seems to have a greater impact than these characteristics or the type of offense.

However, lower recidivism rates for those granted probation, even when other factors are controlled, does not necessarily indicate that the lower rates are a specific function of this type of treatment. Instead, this relationship may be largely an artifact of the court's decision-making process. It is possible that those granted probation have lower recidivism rates because (a) those individuals with "favorable" characteristics and offenses (e.g., the absence of a prior record) are generally granted probation by the courts, and (b) those individuals with these "favorable" characteristics and offenses are most likely to have lower recidivism rates. Thus it is possible that it is the offender's characteristics rather than anything inherent in the type of treatment, or anything inherent in being given one's freedom when probation is granted. That is the primary influence on recidivism.

Nevertheless, for the purpose of policy evaluation and prescription, this possibility may not be fully relevant. For this purpose it is insufficient to simply ascertain which factor is the "best predictor". Policy-makers need information about the explanatory factors over which they have some control. These factors may predict an outcome less perfectly, but they will probably give the policy-maker greater ability to affect the outcome. A judge, or any other policy-maker, can do little to change an offender's age or his number of prior convictions, but he can prescribe the precise type of treatment (probation or incarceration) which he will receive. <u>The requirements of analysis are different for pure social science and applied or policy social science.</u>

An alternative method of analyzing the relationship between type of treatment and recidivism is conducting a controlled experiment. The California Youth Authority's "Community Treatment Project" experiment randomly assigns, after an initial screening, convicted juveniles either to an experimental group which receives probation and intensive counselling or to a control group which is incarcerated. After a follow-up period of twenty-four months the "failure" rate for the experimental group was 38% and 61% for the control group.

However, the CTP experiment seems to have been flawed in several significant respects. For example, initial screening eliminated about 25 percent of the convicted male juveniles (and 10 percent of the females) for whom institutionalization was deemed mandatory because they were involved in serious assaultive cases or because there was community objection. This clearly limits the generality of the conclusions that can be drawn: These data indicate that recidivism is less likely if offenders receive probation, but we do not know if this applies to the most serious offenders. Also, because of the nature of the intensive and special attention and supervision of the experimental group it seems very possible that its lower "failure" rate is to some degree a function of experimenter effects such as the "Hawthorne" effect and the effects of a positive self-fulfilling prophecy and positive labeling. An additional flaw in the CTP seems to have been the initial selection of the population from which the experimental and control groups were selected: the subjects were chosen from areas which did not represent a full cross-section of characteristics. For example, there were few blacks and few subjects from large industrialized areas.

Nevertheless, since the requirements of analysis seem to be different for pure social science and applied or policy social science, these scientifically flawed aspects of the CTP offer quite promising possibilities for the goal of policy evaluation and prescription: (1) Perhaps, the initial screening points to a general policy prescription: Screen out cases with assaultive backgrounds, and then grant probation to all other juvenile first offenders. The CTP indicates that probation leads to less recidivism; thus probation for all but those screened out by the above criteria should significantly lower the present "failure" rate. (2) Perhaps the possibility that the CTP's low "failure" rate for the experimental group is to some degree the function of experimenter effects can be utilized as an explicit and intentional positive policy. Indeed, though it is a rather bold policy, after an initial screening out of assaultive cases, perhaps all juvenile first offenders should be granted probation and assigned to an explicitly and intentionally "Hawthorne" and "positive self-fulfilling prophecy" program. The community agents would intentionally have expectations of "success" for these youths, who also would be positively labeled (as being an "experimental" participant). The CTP findings indicate that this could significantly lower the present <u>overall</u> "failure" rate (i.e., the combined rate for both offenders who are granted probation without a special program and those who are incarcerated).

In a preliminary fashion these findings and prescriptions are applied to the criminal courts of Pittsburgh and Minneapolis. This effort indicates that to convert these findings into policy guidance for criminal court judges, the many goals of the criminal court, in addition to reduced recidivism, must be considered.

A summary of the major findings and some of the implied recommendations of the analysis and experiments reviewed can be stated: For many persons, especially those with certain "favorable" characteristics (e.g., the absence of a prior record), probation can reduce the recidivism rate to approximately 33 per cent. The experiment indicates that for most persons, probation along with intensive and special attention can reduce recidivism to 28 per cent. For most convicted felons, therefore, the type of treatment makes a significant difference. For other felons, however, personal characteristics and the nature of their offense seem more important. The influence of probation on recidivism is thus far from total; but it is clear that knowledge of recidivism rates associated with specific offenses and particular offender characteristics could be of considerable practical value to judges in sentencing.

On the other hand, it must also be emphasized that even among probationers there is a recidivism rate of approximately 33 per cent. This figure represents a very large number of individuals and crimes, and it, as well as the fact that probationers recidivate less than those who are incarcerated, must be taken into account in designing sentencing policies based upon the findings of social science.

There is considerable tension among the goals of the criminal courts, as usually is the case with basic institutional goals and values. Indeed, few important goals and values in society can be simultaneously maximized. It is this tension which makes a consideration of these goals and values so fascinating and perplexing. However, in terms of the single goal of reduced recidivism, this study has attempted to offer more empirical guidance to decision makers and policy evaluators. Yet to achieve this goal, policy makers must also look beyond the criminal courts. As this study has indicated, factors other than court decisions also have a major impact on recidivism. The courts cannot and probably should not affect these factors.

what factors

CHAPTER 1: NON-EXPERIMENTAL ANALYSIS AND EVALUATION

During the past few years there has been a refreshing wave of studies of the relationship between policy inputs and outputs in American urban and state politics.* The reader is referred to articles in the bibliography by Wilson (1968), Jacob and Vines (1965), Jacob and Lipsky (1968), Levin (1970), and Fry and Winters (1970) which cite and discuss many of these recent studies. In addition, for an analysis of the impact of recent Supreme Court decisions, the reader should examine the articles by Muir (1970), Wasby (1970) and Becks (1969). These studies have attempted to go beyond the analysis of the political processes of a unit of government, to analyze its relationship and that of other factors such as socio-economic characteristics to the policy outputs of that unit. They have gone beyond the analysis of "who governs?" to the analysis of "what difference does it make who governs?" and "what difference do certain socio-economic characteristics make?" (Wilson, 1964: 133). In other words, what are the consequences of these inputs for the life of the average citizen? These consequences have been analyzed in terms of the policy outputs and services of these governments in areas such as education, welfare, criminal justice, planning programs, and general social welfare measures. (For a more complete description of this approach, and its theoretical interpretations, see David Easton's Political Systems and Levin, forthcoming.)

As the logical conclusion of these studies and the input-output framework, policy analysts ought to evaluate the impact or outcomes of these policy outputs and thus attempt to discover their ultimate consequences for society. (For example, in the area of input-output analysis of comparative politics, Pennock (1966) argues that an approach that focuses on "outcomes" (the ultimate consequences for society of policy outputs) "deserves a certain

---

*Portions of this report appeared in "Policy Evaluation and Recidivism", Vol. 6, <u>Law and Society Review</u> (August 1971), pp. 17-47; "Crime and Punishment and Social Science", <u>The Public Interest</u> (Spring, 1972).

priority" because "the test of anything in terms of what produces seems to make sense." However, the nature of policy evaluation presents major problems: First, a particular policy output is often only one of a broad range of causal factors affecting the behavior at which the policy is directed. Second, the knowledge of the precise degree to which each of the causal factors affect this dependent variable typically is imperfect. Policy-makers need information about the explanatory factors over which they have some control. These factors may predict an outcomes less perfectly, but they will probably give the policy-maker greater ability to affect the outcome. For example, in evaluating the impact of alternative criminal court sentencing policies on recidivism rates, one is faced with these problems: Sentencing decisions are, at the most, only one of the many causal factors of recidivism. A judge, or any other policy-maker, can do little to change an offender's age or his number of prior convictions, but he can prescribe the precise type of treatment (probation or incarceration) which he will receive. Also, it is difficult to ascertain the precise degree to which each of these causal factors affects recidivism rates. It is especially difficult to ascertain the precise degree to which criminal court sentencing decisions affect them.

The third problem in the nature of policy evaluation is that the range of causal factors is typically broad and very complex. Many of them are highly correlated. That is, real world policy analysis is often confronted with the problem of multicollinearity. This is the name given to the general problem which arises when some or all of the explanatory variables in a relation are highly correlated. It then becomes very difficult, if not impossible, to distinguish and assess their precise relative effects on the dependent variable.[1]

There have been two major responses to these problems in the evaluation of policy impact. The first has been evaluation by the use of regression analysis which is designed to predict effects in a multivariate situation. Controlled experimentation has been the second response.

In principle, regression analysis can hold several explanatory variables constant to ascertain the independent effect of another variable. In practice, however, it is only as effective as the nature of the data allows. If the explanatory variables are highly correlated (i.e., multicollinear), then it is very difficult for regression analysis to assess the precise contribution of each of these variables.

The Coleman Report (1966) is a classic illustration of this problem in policy evaluation. It found that differences in family backgrounds of students account for much more variation in achievement than do school differences. However, this analysis has been criticized for greatly under-estimating the contribution of school quality.[2] In part this resulted from the difficulty in assessing the relative contributions of family background and school quality. This difficulty is a product of the high correlation between the explanatory variables: Good homes and good schools tend to occur together and weak homes and weak schools tend to occur together.

In short, regression analysis is only effective in assessing the precise contribution of several explanatory variables if the data are "internally controlled" (i.e., if there is a good deal of independent variation among the explanatory variables). Significantly, multicollinearity also weakens inferences based on cross-tabulations.[3] Thus, multicollinearity puts analysis and evaluation "in the statistical position of not being able to make bricks without straw" (Johnston 1963: 207).

It must be noted, however, that multicollinearity is a statistical rather than a mathematical condition. Thus, one should think in terms of the problem's severity rather than its existence or non-existence. Also, despite its frequent presence, especially in policy data, multicollinearity is neither always severe nor always present. Farrar and Glaubner (1967: 94) suggest that the problem becomes severe when the explanatory variables are not just correlated but are also highly correlated (e.g., greater than .75), when it is difficult or impossible to obtain the additional information to mitigate this high intercorrelation, and when in addition to these two conditions there are less than twenty data points.

The response of social scientists to the problem of multicollinearity generally is unsatisfactory, especially for the purposes of policy evaluation. They almost exclusively suggest obtaining additional data which hopefully will lessen the degree of correlation between the explanatory variables. This is the standard response given by econometricians such as Johnston, Farrar, and Glaubner.[4] This strategy has led them to frequently use cross-sectional over time-series data which has a high infomational content, and they have apparently been successful at times in surmounting the problem of multicollinearity. See for example Prais and Houthakker (1955), Meyer and Kuh (1957), Orcutt (1961), and Stone (1954). More information may be a suitable abstract solution to the problem, but it is often unsuitable for real-world policy evaluation. Indeed, problems in non-experimental data such as multicollinearity were long overlooked because the source of statistical analyses, such as regression and cross-tabulation, was the controlled world of the laboratory experiment. There, unlike the real world, variables can be manipulated so that the major explanatory factors under study operate independently of one another (Blalock, 1963: 233). The world simply seems to be more complex than additive models suggest.[5] Policy data, especially, often seem to be so complex and intercorrelated that additional data will not mitigate multicollinearity. Often sufficient independent variation among the explanatory variables simply does not exist. Again, the Coleman Report (1966) is illustrative. It was based on an extremely large national sample (3,155 schools and 569,000 students), but this did not mitigate the correlation between the explanatory variables. There are simply very few cases of good homes and weak schools or weak homes and good schools, even in a national survey.

Indeed, even some of those who suggest additional data to mitigate the multicollinearity problem admit that frequently it is not a possible solution. For example, Farrar and Glaubner admit "Admonitions that new data, or additional _a priori_ information, are required to break the multicollinearity deadlock are hardly reassuring, for the gap between information on hand and information required to estimate a model fully is so often immense." Farrar and Glaubner, (1967: 96). Similarly J. Johnston cautions "the remedy lies essentially in the acquisition, _if possible_, of new data which will break the multicollinearity deadlock". Johnston, (1963: 207) (emphasis added).

Farrar and Glaubner do, however, go on to suggest some mild palliatives: They suggest a "specification" and diagnostic approach which (a) focuses on the "critical variables", (b) tolerates multicollinearity among "non-critical" variables, (c) seeks additional information for the "critical variables" if they are affected by multicollinearity, and (d) utilizes diagnostics to develop the additional information: "Structural integrity over an entire set, admittedly, requires both complete specification and internal orthogonality. One cannot obtain reliable estimates for an entire n-dimensional hypotheses, with fewer than n significant dimensions of independent variation. Yet all variables are seldom equally important... Theoretical questions ordinarily focus on a relatively small portion of an independent variable set...Only one - or at most two or three - strategically important variables are ordinarily present in a regression equation. With complete specification and detailed insight into the location and pattern of interdependence in X, structural instability within the critical subset can be evaluated and if necessary, corrected. Multicollinearity among non-critical variables can be tolerated. Should critical variables also be affected, additional information to provide coefficient estimates either for the essential variables directly, or for those members of the set on which they are principally dependent, is required. Detailed diagnostics for the pattern of interdependence that undermines the experimental quality of X permits such information to be developed and applied both frugally and effectively." Farrar and Glaubner, (1967: 106, 95, 107).

Nevertheless, these suggestions do not represent anything approaching a solution of the problem. Farrar and Glaubner are still required to rely on "additional information" for their critical variables. As this paper has argued, and as Farrar and Glaubner indirectly state in the preceding paragraph, such information frequently does not exist in policy data. The diagnostics that they suggest do not obviate this problem and in a research situation they are at best only a place to begin to deal with the problem. Indeed, Farrar and Glaubner are forced to conclude: "It would be pleasant to conclude on a note of triumph that the problem has been solved...Such a feeling, clearly, would be misleading. Diagnosis, although a necessary first step, does not insure cure...The diagnostics described here offer the econometrician a place to begin...<u>No miraculous</u> '<u>instant orthogonalization</u>'

can be offered.

Their assumption that theoretical questions ordinarily focus on a small portion of the independent variable set is another significant weakness in Farrar and Glaubner's suggestions. Indeed, in the evaluation of court policies that is carried out in this paper, one of the major problems of analysis is the large number of "critical variables" involved.

The explanations given for the heretofore inadequate attention given to the problem of multicollinearity seem to indicate that the problem is precisely a function of the complexity of the real world. Most suggest that problems presented by nonexperimental data were long overlooked because the source of statistical analyses, such as regression and cross-tabulation, was "the controlled world of the laboratory experiment. For example, "Theoretical statisticians, drawing their training, experience, and data from the controlled world of the laboratory experiment, are noticeably uninterested in the problem of multicollinearity altogether." (Farrar and Glaubner, p. 95. H. Blalock, p. 233.)

Indeed, as Blalock points out, by contrast to nonexperimental or real world researchers, those working with experiments are able to manipulate their variables so that the major explanatory factors under study operate independently of one another.

"For example, they may make use of two-way analysis of variance or some more complex design involving equal numbers of replications in all subcells. Randomization may then be used to provide assurance that at least some of the additional but unmeasured variables operating will affect the dependent variable independently of the factors under study." (Blalock, 1963: 233).

Therefore, policy evaluators are faced with what Johnston called a multicollinearity deadlock which in practice cannot be either by additional information or by the use of different statistical methods (Johnston's conclusion is less pessimistic and far-reaching than the one reached here).[6] Tukey's conclusion based on a discussion of both regression and path coefficients is "The problem is highly complex and perhaps not capable of yielding any satisfactory solution" (Tukey, 1954). Moreover, multicollinearity seems to have special consequences for policy evaluation which have been overlooked by even the most careful analysts. For example, Blalock's excellent analysis of the problem of multicollinearity is one of the few that analyzes both its nature and some of its consequences for non-experimental social scientists. Nevertheless, he does not discuss its consequences for policy analysis and evaluation. (Blalock, 1963: 234). Similarly Tufte's (1969) excellent analysis of methods of improving data analysis in political science points out some consequences of multicollinearity but not for policy evaluation.

In policy evaluation, the program in question is often only one of several factors affecting the behavior at which it is directed, and often these factors are highly correlated. A program may have an effect, but because of multicollinearity it may not be possible to measure it precisely as distinguished from the effect of other intercorrelated factors. Thus, when evaluations of specific programs conclude that the programs make no difference (e.g., crime and delinquency reduction programs. Operation Head-start, and other compensatory education programs). This conclusion may be largely the result of the problem of multicollinearity.

For example, this is the conclusion of Miller (forthcoming) concerning the effect of almost all programs to control and prevent delinquency, including detached worker programs. Similarly, after analyzing the various crime prevention and rehabilitation programs undertaken thus far, Stanton Wheeler and his associates conclude: "As of now, there are no demonstrable and proven methods for reducing the incidence of serious delinquent acts through preventive or rehabilitative procedures. Either the descriptive knowledge has not been translated into feasible action programs, or the programs have not been successfully implemented; or if implemented, they have lacked evaluation; or if evaluated, the results have usually been negative; and in the few cases of reported positive results, replications have been lacking" (Wilson, 1967: 73).

After surveying various efforts at compensatory education, the U.S. Civil Rights Commission (1967: 138) said "none of the programs appear to have raised significantly the achievement of participating pupils." There have been similar studies of Operation Headstart, which have had similar conclusions. (Westinghouse Learning Corporation Study, 1969; Evans, 1969). Most of these studies have been said to have serious methodological limitations; I am not referring to these questions but only to the problem of multicollinearity. Also see Cohen (1970: 8, 23-24) for an analysis of more recent evaluations of "Title I" programs which reach similar conclusions.

The difficulties in construing policies to reduce the problem of recidivism are typical of the dilemma which both the social scientist and the policy-maker often face in developing the "best" and most "rational" program. One difficulty is the cognitive limitations of pure social science. If we could know perfectly what the best and most rational policy is, if social science could produce knowledge as unambiguous, clear and precise as that of the natural sciences, then policy-making would become more of a technical and administrative task and less of a political one. But, as our analysis of recidivism shows, pure social science findings are often far from ambiguous. Pure social science is an essential but usually imperfect foundation for policy making. Yet in the face of pressing social problems, the policy analyst--like the political actor--cannot afford to remain agnostic or passive. This is the first basis of our distinction between pure social science and policy social science.

A second element in this distinction is policy social science's emphasis more on explanatory factors over which they have some control rather than simply those which are the "best predictors". Pure social science ideally emphasizes certainty and truth. In parsimonious and minimalistic manner, it attempts to isolate the necessary and sufficient conditions for the occurrence of the phenomenon under investigation. (As put in "Ockham's razor", explanatory variables are not to be multiplied unnecessarily.) But the application of parsemonious methods of proof to the causal analysis of recidivism would lead to agnostic conclusions and thus passivity. The policy analyst, however, need not be limited by these requirements of pure social science. However, for policy analysis it is insufficient to simply ascertain which factor is the "best predictor".

For example, the problem of multicollinearity occurred when the President's Commission on Law Enforcement and Criminal Justice (1967) analyzed the policy of improved street lighting. The proponents of this policy suggest that adequate — and particularly above adequate — street lighting will, first deter certain types of street crimes by increasing the offender's risk of being detected and, second, enhance the probability of apprehending the offender. A study of Flint, Michigan's major improvement of its central business district lighting found that over a six-month period there was a 60% reduction in the number of all felonies and misdemeanors and an 80% reduction in larcenies. However, at the same time there was an increase in police surveillance of the area. Therefore, it is not possible to ascertain the precise effect of street lighting alone. In other words, two possible causal factors — improved lighting and increased police surveillance — were perfectly correlated, and it is impossible to assess the precise contribution of each. Both of these possible causal factors often are likely to occur together because they are the product of the same general force — the desire to reduce crime. Such a complex and collinear pattern is probably typical of policy programs and the factors surrounding them. Nevertheless, the findings of the Flint study and other studies of improved street lighting led the President's Commission to conclude that "there is no evidence that improved lighting would have a lasting or significant impact on crime rates," though they did add that "there is a strong suggestion that it might."[7]

In summary, the frequent absence of a satisfactory solution to the multicollinearity deadlock should lead one to be more cautious in rejecting a policy program as "making no difference." If such a finding seems to be indicated, one should then investigate the interrelationships among the independent variables to ascertain whether multicollinearity does in fact exist and, if so, to what degree. If it does exist, however, policy evaluators should not despair completely. First, as will be discussed in the following section, there are methods of policy evaluation other than regression analysis and cross-tabulation. Second, the experience of physicists is perhaps instructive. Heisenberg's uncertainty principle has caused difficulties in the field of subatomic theory, but on the whole physicists have made major theoretical strides despite this principle. (The uncertainty principle states that it is impossible to specify or determine simultaneously both the position and velocity of a particle with full accuracy. It is possible to fix either of these quantities as precisely as desired, but the more exactness in one, the increasing uncertainty in the other. This lack of precision results from the effect of the observation on the observed particle.) More importantly, the precision of applied science (e.g., sending a man to the moon or pinpointing an ICBM target 3,000 miles away) has not been deterred significantly. The requirements of analysis seem to be different for pure science and applied science. Similarly, social scientists ought to be able to make strides in both pure and applied fields despite the complexity and frequent multicollinearity of the real world. Perhaps they ought to develop their own uncertainty principle: The closer one gets to the facts, the more difficult it is to offer confident generalizations. Moreover, they ought to become aware that the requirements of analysis are different for pure social science and applied or policy social science. The second part of this paper will attempt to indicate the fruitfulness of this distinction in a concrete case of policy evaluation and prescription.

CHAPTER 2: CONTROLLED EXPERIMENTAL ANALYSIS AND EVALUATION

Controlled experimentation is an alternative response to some of the problems of evaluating policy impact. In randomly applying a program or treatment to a population, the various possible independent variables other than the program in question are controlled. The data are thus controlled by randomization at the outset rather than in ex post facto manner, as in regression analysis. Also, in controlled experimentation the data can be manipulated so that they do not present problems for statistical analysis, such as multicollinearity. Finally, in controlled experimentation all variables which are present in that population are included, and it thus avoids the problem of failing to include them in the regression equation. This is especially important in policy evaluation (e.g., often we may not be fully aware of all the possible major causes of a social problem such as recidivism.)

In the criminal justice area, the possibility that the relationship between type of treatment and recidivism rates may be an artifact of the court's decision-making process may not be fully relevant for the purpose of policy evaluation and prescription. It is possible that offenders' characteristics, rather than anything inherent in the type of treatment, is the primary influence on recidivism. However, for policy analysis it is insufficient to simply ascertain which factor is the "best predictor". Policy-makers need information about the explanatory factors over which they have some control. These factors may predict an outcome less perfectly, but they will probably give the policy-maker greater ability to affect the outcome. A judge, or any other policy-maker, can do little to change an offender's age or his number of prior convictions, but he can prescribe the precise type of treatment (probation or incarceration) which he will receive.

In the education area, even if a student's family background is the best predictor of educational achievement, it is almost impossible for policy-makers to influence this factor. However, they do have some control over school quality, though in the Coleman Report analysis this factor seems to have been a less important predictor of educational achievement.

Recently policy analysts have had a few opportunities to conduct genuine controlled experimentation with social policies. (Rosenthal and Jacobson (1968a) provide for descriptions of the experiments they conducted which focus on educational achievement and experimenter-subject interaction and bias. They also list several other controlled experiments focusing on social policies, primarily in the field of education.)[8] However, the difficulty with controlled experimental evaluations of policy impact seems to be less a problem of conducting them than a problem of the nature of their methodology and experimentation.

The first of these problems is a well-known concept to social scientists, and therefore will be discussed briefly. In controlled experimentation there is the danger of the "Hawthorne effect" occurring. The name comes from the intensive series of experiments conducted at the Western Electric Company's Hawthorne Works in Chicago in the 1920's to determine how various changes in working conditions would affect the performance of female workers. Some of the experiments, for example, involved changes in lighting. The researchers found it was not significant whether the worker had more or less light but merely that she was the subject of attention. Any changes that involved her, and even actions that she only thought were changes, were likely to improve her performance. Thus, the improved performance seems to have been largely a function of the workers being part of an experiment.[9] The term Hawthorne effect is usually applied to changes brought about in this or a similar manner such as a subject simply becoming the focus of a special effort or attention. In the medical sciences a similar phenomenon is the "placebo effect" which is the introduction of a new treatment accompanied by improvement regardless of the nature of that treatment.[10]

A second, and until recently a less frequently analyzed set of problems in controlled experimentation, results from experimenter effects which produce a self-fulfilling prophecy and the special cases of experimenter bias and labeling. During the last ten years Robert Rosenthal and various associates have investigated experimenter effects and experimenter-subject interaction in various contexts.[11] One investigation focused on the effect of teacher expectations with experiments in which teachers were led to believe at the beginning of a school year that, on the basis of tests that had been administered toward the end of the preceding school year, certain of their pupils could be expected to show considerable academic improvement during the year. In actuality the children designated as potential "spurters" had been chosen at random and not on the basis of testing. Nonetheless, intelligence tests given after the experiment had been in progress for several months indicated that on the whole the randomly chosen children had improved more than the rest (Rosenthal and Jacobson, 1968b: 19-20).

Specifically, they investigated the effect of teacher expectations with experiments in which teachers were led to believe at the beginning of a school year that certain of their pupils could be expected to show considerable academic improvement during the year. The teachers were told that the predictions were based on tests that had been administered to the student body toward the end of the preceding school year. In actuality the children designated as potential "spurters" had been chosen at random and not on the basis of testing. Nonetheless, intelligence tests given after the experiment had been in progress for several months indicated that on the whole the randomly chosen children had improved more than the rest.

Rosenthal and Jacobson had taken steps to make certain that the predictions about the children were not based on judgments derived from previously observed behavior. They thus explain this greater improvement as a function of a self-fulfilling prophecy which, in this case, was the teacher's positive expectations for these children. "The essence of the concept of the self-fulfilling prophecy," Rosenthal and Jacobson explain, is that one person's prediction of another person's behavior somehow comes to be realized. The prediction may, of course, be realized only in the perception of the predictor. It is also possible, however, that the predictor's expectation is communicated to the other person, perhaps in quite subtle and unintended ways, and so has an influence on his actual behavior. The general phenomenon that they suggest is that in some instance a prediction about subsequent behavior has an affect on that behavior independent of (and sometimes greater than) other factors.

This explanation of a pattern of experimenter effects in the form of a self-fulfilling prophecy was developed earlier in Rosenthal's laboratory experiments. Here the experimenter effect that was focused on was the special case of experimenter bias. Rosenthal's experiments used "rats that were said to be either bright or dull. In one experiment 12 students in psychology were each given five laboratory rats of the same strain. Six of the students were told that their rats had been bred for brightness in running a maze; the other six students were told that their rats could be expected for genetic reasons to be poor at running a maze. The assignment given the students was to teach the rats to run the maze. From the outset the rats believed to have the higher potential proved to be the better performers. The rats thought to be dull made poor progress and sometimes would not even budge from the starting position in the maze". Rosenthal and Jacobson, (1968b).

The problem of experimenter bias is often dealt with in natural science and medical science by using double-blind trials--withholding from the experimenter the knowledge of both the recipient of the treatment and the exact treatment in an individual case. However, even in these sciences these precautions are frequently not executed successfully. (Rosenthal and Jacobson, 1968a). More importantly, such precautions are usually difficult to even institute and then execute successfully in policy evaluation as indicated by Rosenthal's own investigations and those that he describes.

Rosenthal and Jacobson also tested the alternative explanation that these intelligence test results were a function of a Hawthorne effect rather than of a self-fulfilling prophecy. Perhaps the fact that researchers supported by federal funds were interested in this school led to a general improvement of morale and a greater effort on the part of the teachers. They are able to reject this alternative explanation because "a Hawthorne effect might account for the gains shown by the children in the control group, but it would not account for the greater gains made by the children in the experimental group" (Rosenthal and Jacobson: 1968b: 23).[12]

Rosenthal and Jacobson also analyze negative experimenter effects caused by negative prophecies,[13] and they cite several examples of such prophecies from "everyday life" and the medical sciences. Rosenthal and Jacobsen (1968a). As will be noted in the discussion of criminal justice policies, the alternatives of positive and negative prophecies and their effects are quite significant for the special case of a self-fulfilling prophecy termed "labeling". In this process an individual is named or given a "label which then seems to often create a self-fulfilling identity of personal definition of his behavior.

Rosenthal and his associates do not use the term "labeling" in any of the studies nor do they explicitly discuss this special case. Other studies offer little hard data concerning this phenomenon, though it is dealt with in interpretive and descriptive terms in many fascinating studies.

The lack of precision indicated by the Heisenberg uncertainty principle is another difficulty resulting from experimenter effects. This lack of precision specifically results from the effect of observation on the observed particle. In observing a system it is necessary to exchange energy and momentum with it. This exchange alters the original properties of the system.

The third of these methodological and interpretive problems in controlled experimentation is the initial selection of the population or universe from which the control and experimental groups will be randomly selected. The nature and characteristics of this population becomes a constraint on the generality of the conclusions drawn from the experiment. For example, if this population is not typical of the more general population at which the policy is to be directed or if it differs in even one or two major characteristics, the applicability of the experiment's conclusions for this more general population is clearly questionable. (For example, in looking at the question of recidivism in the Community Treatment Program examined below, the most serious felons were screened out from the program. It is impossible to ascertain what effect that program would have on them.) This situation seems to be a real possibility. In practice there seems to be a tendency in policy evaluation to select the population or universe on criteria of convenience and non-controversy. This often means that close at hand and those in political jurisdictions whose elected officials are willing to allow a policy experiment to occur. For example, in some instances small urban areas have been the source of experimental populations because of convenience--they are close to the state capital in which the governmental unit conducting the experiment is located, and they are considered to be "more manageable". Such a population may not be typical of a larger urban area toward which the policy is generally directed. In other instances small urban and rural areas have been used because the elected officials of the more relevant larger urban areas have been unwilling to allow a policy experiment to take place there. This seems to have occurred in some welfare and income maintenance experiments. Also, in several policy experiments the source of the population has been individuals in a university town or individuals in the university itself, and this has obvious shortcomings. In a few instances the source of the population has been typical of the more general policy target population, but it has been small in order to keep the study "more manageable". This unfortunately has meant that there have sometimes been too few cases for valid conclusions in certain categories of the population in the study.

CHAPTER 3: THE EVALUATION OF ALTERNATIVE CRIMINAL COURT SENTENCING POLICIES

Non-experimental Analysis and Evaluation

This part of the study will evaluate the impact of alternative criminal court sentencing policies on reducing recidivism. Criminal court judges have a very high degree of discretion in sentencing decisions. Criminal statutes in most states allow the judge the choice of incarcerating a convicted defendant or of granting probation in common felonies. If the judge decides to imprison him, the statutes also allow him freedom to set the term in prison within certain prescribed limits. Courts in some areas generally tend to incarcerate convicted defendants more frequently than they grant probation. Courts in other areas generally tend to do the opposite, and some courts choose each alternative with about the same frequency. That is, the sentencing decisions of some criminal courts are generally lenient, while others are generally severe. (Of course, there may be a good deal of variance among the decisions of the individual judges in a single jurisdiction. I am referring here to the overall statistical pattern found in that jurisdiction as indicated by the percentage of convicted offenders that receive probation.) For example, in 1966, in the state of California as a whole 32.0% of the convicted defendants in superior court received probation. The range of frequency of probation among the state's thirteen largest counties was from 7.2 (Fresno) to 40.7 (Alameda i.e., the Oakland area). In Los Angeles County the frequency of probation was 37.0%, but in Orange County it was 12.5%; in San Francisco County it was 35.6%, but in Sacramento County it was 9.1% (Beattie and Bridges, 1970). Similarly an earlier study that I conducted indicated during the mid-1960's approximately 49% of the convicted common felons received probation in Pittsburgh, while approximately 37% of them received it in Minneapolis. Moreover, this difference between these two cities is even greater when controls are introduced for factors such as race and prior record (Levin, 1970, Chapter V).

This part of the paper will evaluate the impact of alternative criminal court sentencing policies. It will focus almost exclusively on the impact of these policies on reducing recidivism. This is one of the major goals of the criminal court, and it also contributes to the attainment of two other closely related goals of the court: greater protection of society and greater rehabilitation of defendants. Moreover, reduced recidivism can be quantified and measured with some precision. Thus, evaluation of the impact of sentencing policies on reducing recidivism aids in making a policy choice between courts that frequently tend to incarcerate and those that frequently tend to grant probation. The impact on recidivism of factors other than court policies will also be evaluated. This analysis also aids in prescribing general policies to achieve reduced recidivism. (By recidivism I simply mean an individual who is convicted of an offense after he has been convicted of a previous offense. The use of this term in no way implies the opposite or rehabilitation. For the sake of brevity, recidivism rates will almost always be stated in the short-hand terms of "success" rates or "failure" rates (which in no way imply any "existential" state). Since they are shorthand terms, the reader should note their precise operational definition which often varies among the studies described here.

In analyzing these studies formal probation with supervision and suspended sentences which do not involve supervision are considered together under the shorthand category of "probation." In most of the studies, almost all of the cases in this category involve formal probation with supervision. No cases which involve probation plus some term of incarceration are included in the category "probation," although in the official data of some states, such as California, the term "probation," includes such cases. In this analysis these latter cases are included in the category of "incarceration."

Ultimately, this study is primarily concerned with the impact of criminal court policies on recidivism, but there seems to be no a priori reason to suspect that court policies would be the only factor, or even the predominant factor, shaping recidivism rates. The studies of recidivism that will be analyzed here are therefore those that deal with the impact of several variables, and the relative impact of each of these variables will be analyzed.

The studies of factors affecting recidivism all indicate that offenders who have received probation generally have significantly lower rates of recidivism than those who have been incarcerated. They also indicate that of those incarcerated, the offenders who have received a shorter term of incarceration generally have a somewhat lower recidivism rate than those who receive longer terms. With a few exceptions, these differences persist when one controls for factors such as type of offense, type of community, the offender's age, race, and number of previous convictions. That is, the difference in recidivism rates for the two treatments generally remains the same for all types of offenders. However, for those with certain characteristics (e.g., youthfulness, previous record) there are some significant variations in the overall recidivism rates when type of treatment is controlled (e.g., for all those who receive probation the recidivism rates are highest for the youngest and for those with the greatest prior record).

Beattie and Bridges' analysis in 1970 of recidivism rates of offenders who were either granted probation or were incarcerated by the incarceration are included in the category "probation," although in the official data of some states, such as California, the term "probation," includes such cases. In this analysis these latter cases are included in the category of "incarceration."

Ultimately, this study is primarily concerned with the impact of criminal court policies on recidivism, but there seems to be no a priori reason to suspect that court policies would be the only factor, or even the predominant factor, shaping recidivism rates. The studies of recidivism that will be analyzed here are therefore those that deal with the impact of several variables, and the relative impact of each of these variables will be analyzed.

The studies of factors affecting recidivism all indicate that offenders who have received probation generally have significantly lower rates of recidivism than those who have been incarcerated. They also indicate that of those incarcerated, the offenders who have received a shorter term of incarceration generally have a somewhat lower recidivism rate than those who receive longer terms. With a few exceptions, these differences persist when one controls for factors such as type of offense, type of community, the offender's age, race, and number of previous convictions. That is, the difference in recidivism rates for the two treatments generally remains the same for all types of offenders. However, for those with certain characteristics (e.g., youthfulness, previous record) there are some significant variations in the overall recidivism rates when type of treatment is controlled (e.g., for all those who receive probation the recidivism rates are highest for the youngest and for those with the greatest prior record).

Beattie and Bridges' analysis in 1970 of recidivism rates of offenders who were either granted probation or were incarcerated by the

Superior Courts of California's thirteen largest counties is the most comprehensive study to date of factors affecting recidivism (Beattie and Bridges, 1970). (The Superior Court is the county trial court in California; its criminal jurisdiction includes all serious offenses (i.e., all felonies and several major misdemeanors). The offenses included in their study are homicide, robbery, assaults, forged checks, auto theft, "other theft," sex offenses, drug law violations, and "other offenses."

The data in this study include all the Superior Court probation and jail cases for the first six months of 1966 for twelve of the thirteen counties and 30% of those cases from Los Angeles.) It simultaneously analyzes recidivism for both those incarcerated and those granted probation, with controls for many factors other than type of treatment. It indicates that the "success" rate for those granted probation was 65.8% (2,148) after a one-year follow-up and 48.6% (2,561) for those sentenced to jail. ("Jail" refers here to a term of incarceration of no more than one year, which is served in a city or county jail. In California all terms of incarceration greater than one year are served in a state prison. The Beattie and Bridges study did not include offenders sentenced by the Superior Courts to state prison, but they are analyzed in studies described below.

The follow-up period in this study was twelve months from the time of the individual's release to the street on probation or following incarceration. This is a limitation only in assessing the general degree of recidivism. (Other studies have indicated that while most recidivism occurs during the first year following release, a significant degree does occur in the next year.) This does not seem to be a limitation for assessing the differences, if any, in recidivism rates for different types of treatment. There is no evidence in other studies that the recidivism rates for different types of treatment would vary significantly from the first to second year. Nor is there any substantive reason to entertain such a hypothesis.) The "success" rate cited here is Beattie and Bridges' "none" category which signifies no known arrest either for a new crime or for technical violation of probation or parole during the one-year follow-up period. This difference between "success" rates for the probation and jail groups persists when the following factors are controlled: county, sex, age, race, prior record, offense; and when the following factors are controlled simultaneously: offense and age, offense and race, offense and prior record (Beattie and Bridges, 1970: 11-200).

George Davis' earlier study indicates, after a four to seven-year follow-up period, a "success" rate (i.e., no subsequent probation violations or arrests) of 67.1% (6,268) for all those granted probation in fifty-six of California's fifty-eight counties. All defendants granted probation or "probation plus jail" in California during 1956 to 1958 were included in the analysis, except those in Los Angeles and Alameda (Oakland area) counties for which "there was inadequate information at that time" (Davis, 1960). This study also included offenders incarcerated under the sentence--"probation plus jail." Their "success" rate was 56.7% (5,400); it did not include offenders incarcerated under the sentence--"straight jail". (See note above for an explanation of these categories.) These overall rates for each type of treatment were not controlled for factors such as offense, age, race, and prior record. (Davis only presents percentages for the combined categories probation and "probation plus jail" categories; I have recalculated his raw data to ascertain percentages for these categories separately.)

Ralph England's study indicates, after a six to eleven-year follow-up period, a "success" rate of 82.3% (490) for a sample of adult probationers sentenced in the federal district court of the Eastern District of Pennsylvania from 1939 to 1944 (England, 1957). This study does not include any recidivism data on a comparable group of offenders who were incarcerated by this court. Also, the "success" rate that it indicates is not strictly comparable to those in the studies cited above. England used a less stringent criteria of "success" than did the other studies, and the offenses committed by those in his sample are generally less serious than those in the other studies. England explicitly states only a precise criterion of "failure"--

if a probationer is subsequently convicted of a misdemeanor or felony. Therefore, it is likely that included in his "success" group are some individuals who were arrested but not convicted, or who committed a technical violation of probation but were not convicted of a new offense. Thus, in comparison to these other studies, the "success" rate that England indicates is probably somewhat of an overestimation. Despite these limitations, the relatively unique characteristics of the offenders and offenses in England's study makes it of special interest to this analysis. These unique characteristics present a good opportunity to test some hypotheses concerning the relative impact of offender characteristics on recidivism as opposed to the impact of different types of treatment, and this will be done below.

England's study also summarizes the findings of eleven follow-up studies of recidivism rates of individuals placed on probation. In nine of the eleven studies there was a "success" rate of 70 to 90 percent and in the other two it was between 60 and 70 percent. Again, the criteria of "success" used in most of these studies is less stringent than that of the Beattie and Bridges or Davis studies. Therefore, in comparison to these studies, the "success" rates are probably somewhat of an overestimation. However, aside from this, the validity of these findings is greatly bolstered by their uniformity and their breadth--they were carried out in five states and one European country over a thirty-year period (1921 to 1954).

Data from the California Department of Corrections (CDC) for individuals released after incarceration in California state prisons indicate, after a one-year follow-up period, a "failure" rate ranging from 24.7% to 34.2% with a median "failure" rate 30.5% (9,226) for each year from 1958 to 1968.[17] (These data include no information on any type of "success" rate.) The criterion for "failure" used by the CDC is returned to prison either with a new felony conviction or without one (i.e., a technical violation). By contrast, when a similar criterion is applied to the Beattie and Bridges' data, the "failure" rate for those granted probation then the sample is only 10.9% (2,148). (Because of the differences in the categories used by Beattie and Bridges, this "failure" rate is probably somewhat of an underestimation in comparison to the CDC data.) Since these two sets of data are both from California, they also enable us to examine possible differences in recidivism rates according to length of incarceration. All individuals in the Beattie and Bridges "jail" group were incarcerated for twelve months or less and their "failure" rate is 21.1% (2,561). (Again this percentage is probably somewhat of an underestimation.) By contrast, all individuals in the CDC data were incarcerated for more than twelve months. (The median term of incarceration for this group ranged from twenty-four to thirty-six months during 1960 to 1968.) As noted, their failure rate for these years ranged from 24.7% to 34.2% with a median of 30.5% (9,266).

A detailed 1970 study by Public Systems Incorporated (PSI) based on California Department of Corrections data for individuals released from state prisons in 1964 and 1966 indicates, after a three-year follow-up period, a "success" rate of 32.8% (1,423) and 33.6% (1,208) respectively, or about half that of the "success" rate of those California offenders granted probation in the Beattie and Bridges analysis (Kolodney, 1970: Vol. 2, III-7). (The definitions of "success" were exactly identical in both studies--no subsequent arrests. However, the follow-up period in the PSI study was three years and in the Beattie and Bridges study it was only one year. This should not have significantly lowered the "success" rate in the PSI study because most studies indicate that the preponderance of recidivism occurs during the first twelve months. Indeed, the PSI data themselves indicate almost 70% of the recidivism of those in its study occurred during that period.) Also, a comparison of the PSI and the Beattie and Bridges data again indicates lower recidivism rates for shorter terms of incarceration: All of the Beattie-Bridges "jail" group had terms of twelve months or less and, as noted, their "success" rate was 4816% (2,561); all in the PSI group had terms for more than twelve months, with the median term of incarceration of 30 months in 1964 and 36 months in 1961 and, as noted, their "success" rate for 1964 and 1966 was 32.8% (1,423) and 33.6% (1,208).

Charles Eichman's study of two groups of incarcerated offenders indicates a lower "failure" rate for those with shorter terms of incarceration (Eichman, 1966). The Gideon v. Wainwright "right to counsel" decision by the U.S. Supreme Court required the state of Florida to discharge 1,252 prisoners well before their normal release dates. These were indigents who had been tried for felonies without counsel. Eichman analyzed the post-release experience of a group of 110 of these Gideon early releases and a control group of 110 full-term releases. The two groups were carefully matched for similar characteristics such as prior convictions, type of offense, age, and occupational skill level. (The small final sample was the result of rigorous selection among 406 prisoners for true matches.) Upon release the Gideon early releases had been incarcerated for significantly less time than the full-term releases. (60% of the Gideon early releases had been incarcerated for less than eighteen months and only 46.5% of the full-term releases had been incarcerated for less than that time. (Eichman's analysis of the statistical significance of the difference in length of incarceration indicates a P of less than .001 for a Chi Square=53.6321, with 6 degrees of freedom.) Eichman found that after a twenty-eight month follow-up period the "failure" rate for the Gideon early release group was 13.6% (110). Eichman's "failure" rate is based on subsequent incarceration. Eichman's analysis of the statistical significance of this difference in "failure" rates indicates a P less than .05 for a Chi-Square=4.1624, with one degree of freedom.) For the full-term releases it was 25.4% (110).

Daniel Glaser's monumental study of the Federal prison and parole system indicates, after a four-year follow-up, a "success" rate of 52.2% (1,015) for individuals who had been incarcerated in the federal prison system (Glaser, 1967). It should be noted that because Glaser's sample was from Federal prisons it includes offenses that are generally less serious than those in the other studies, which are based on state prison and probation populations. Thus, in comparison to these other studies the "success" rate of the Glaser study, like that of the Ralph England study which covers Federal probationers, is probably somewhat of an overestimation.

Glaser also describes three studies similar to his own which cover state prisons in California (1946 to 1949), Washington State (1957 to 1959), and Pennsylvania (1956 to 1958). They indicate that after follow-up periods of thirty-six months, six to thirty months, and approximately twenty-eight months, there were "success" rates of 28%, 49%, and 52% respectively (Glaser, 1964: 21-24).

Some of these studies analyzed the impact on recidivism factors other than the type of treatment prescribed by the court. Beattie and Bridges found that the younger the defendant, the more likely he was to repeat. For both those who received probation and those incarcerated, the youngest offenders had the lowest "success" rates and these rates increased for each age category (Beattie and Bridges, 1970: 14-15). They also found that Negro offenders have lower "success" rates than whites, for both offenders granted probation and those incarcerated (Beattie and Bridges, 1970: 15-28). The greater an offender's prior record, the more likely he is to repeat. For both those granted probation and those incarcerated, those with no prior record had the highest "success rates, and these rates decreased for each level of a prior record (Beattie and Bridges, 1970: 16-29). They also found significant variation in the recidivism rates according to the type of offense. For both those granted probation and those incarcerated, those who had committed sex offenses and crimes against persons (homicide, robbery, and assaults) had the highest "success" rates respectively; those that had committed auto theft, burglary, and drug law violations had the lowest "success" rates respecitvely (Beattie and Bridges, 1970: 13, 24-25). The studies by George Davis, PSI, and Daniel Glaser have similar findings.

In a few instances in these other studies, the impact of these other factors is not as pronounced as it is in Beattie and Bridges' data. However, in a later analysis we will indicate in detail that this usually has been either due to the insufficiency in these studies of detailed data on these various factors or the unwillingness of investigators in government sponsored research to draw apparently controversial conclusions, such as higher recidivism rates for certain racial groups.

Thus, these factors clearly have an impact on recidivism, but in almost all instances there is still a significant difference in recidivism rates for those individuals with these characteristics who receive probation and those who are incarcerated. In these instances the <u>type of treatment</u> prescribed by the judge seems to have a greater impact than these characteristics or the type of offense. However, one characteristic--the absence of a prior record, and the two offenses--auto theft and drug law violations--seem to have a greater impact on recidivism than does the type of treatment prescribed by the judge. In another instance the combination of a particular offense with two other characteristics, has a greater impact on recidivism than does the type of treatment.

Specifically, for offenders with no prior record, Beattie and Bridges found that for those granted probation the "success" rate is 78.2% (687) and for those incarcerated it is 72.8% (377). Thus in this instance the type of treatment has less of an impact on the "success" rate than does the characteristic of having no prior record; if an offender has no prior record, he will have a very high "success" rate no matter which type of treatment is prescribed by the court. Similarly, if an offender commits auto theft or a drug law violation he will have a low "success" rate no matter which type of treatment he receives. For auto theft, for those granted probation the "success" rate is 44.1% (118) and for those incarcerated it is 39.0% (241). For drug law violations these figures are 58.7% (339) and 52.0% (321) respectively.

However, for all other offenses, including those that the data indicate have a major impact on recidivism, the type of treatment has a greater impact than does the type of offense. For burglary the "success" rates are low for both those granted probation and those incarcerated--56.3% (304) and 43.8% (526) respectively, but it is significantly lower for those incarcerated. There are similar patterns for sex offenses and crimes against persons; these offenders have high "success" rates for both treatment categories, but they are significantly higher for those granted probation. Though when offense and age are simultaneously controlled, for sex offenses committed by individuals over thirty years old the simultaneous impact of these factors is greater than the type of treatment which they receive. The "success" rate for these offenders is 86.9% (114) for those granted probation and 84.2% (32) for those incarcerated. When offense and prior record are simultaneously controlled, there are similar patterns of a greater impact of these simultaneous factors for sex offenses committed by whites (almost identically high "success" rates for both types of treatment), for sex offenses committed by individuals with no prior record (almost identically high), and for burglary committed by individuals with no prior record (almost identically moderate "success" rates.)

Beattie and Bridges data also indicate that for offenders with all other characteristics, including those that the data indicate have a major impact on recidivism, the type of treatment has a greater impact than do their characteristics (either individually or simultaneously). For example, as noted, youthfulness has a major impact on recidivism, but for offenders under twenty years old and for those twenty to twenty-four years old the "success" rates are higher for those granted probation--54.0% (176) and 58.0% (712) respectively--than for those incarcerated--44.4% (180) and

42.7% (924) respectively. Similarly, as noted, whether an offender is a Negro has a significant impact on recidivism, but for Negro offenders the "success" rates are much higher for those granted probation. The degree of prior record also has a major impact on recidivism, but for those with the greatest degrees of prior record, the "success" rates are significantly higher for those granted probation (Beattie and Bridges, 1970: 21-35).

The regression analysis of the PSI study is the most sophisticated and careful effort thus far to assess the relative impact on recidivism of type of treatment, type of offense and offender characteristics. Before describing the PSI findings, it should be noted that two major shortcomings in the analysis limit its application. First, the PSI study only analyzed an incarcerated population; it has no data whatsoever on individuals who received probation. Second, as will be indicated below, even within this limited population, some of the analysis is plagued by multicollinearity.

The PSI study concluded that "at the 90% level of confidence, the variables which are associated with the response [i.e., no recidivism] are, in order of their contribution, prior record, class, narcotic history, ethnic [i.e., racial] group, base expectancy, and age. Prior record, class and narcotic history are by far the most important variables.... The variables of primary interest, time served [in incarceration], 'fell out' of the model. This variable has no effect or is not associated with the probability that an individual is clean [i.e., no recidivism]." (Kolodney, 1970: III-27) It should be emphasized that this conclusion only applies to one type of treatment--incarceration. More importantly, there seems to be three reasons to be hesitant in accepting it even with respect to incarceration. First, the cross-tabulation analysis presented in the PSI study itself indicates that when type of offense is controlled, for most offenses there are significant differences in the recidivism rates for those incarcerated for "short" or "long" terms. (Kolodney, 1970: III-18) (Admittedly of course, cross-tabulation is less powerful and less revealing than regression.)

Second, there are reasons to suspect that there was insufficient variation among the data points for the independent variable of "time served in incarceration" to properly assess its potential contribution to recidivism rates. (The Coleman Report had precisely the same difficulty with insufficient variation among the data points for the independent variable of "class size". There was an insufficient number of small classes. Some critics have suggested that this led the Coleman analysis to underestimate the potential impact of class size—especially a small class size—on educational achievement. This type of insufficient variation is common in the analysis of policy data.) Specifically, the terms of incarceration all tend to be rather long. The PSI study uses the labels "short" and "long" terms of incarceration but in fact, there are too few genuinely short terms of incarceration (e.g., twelve months or less or even eighteen months or less) to test whether a short term has any impact on recidivism. Evidence for the latter possibility comes from the comparison noted above of the "jail" group data in the Beattie and Bridges study (those incarcerated twelve months or less) and the PSI sample (all of whom were incarcerated for more than twelve months and for whom the median term was 30 months in one year and 36 in the other). It indicated that the "success" rate for the "jail" group was 48.6% (2,561), while it was only 32.8% (1,423) and 33.6% (1,208) for each of the years in the PSI sample.

Third, the PSI study is plagued by multicollinearity. Several of the independent variables to be tested—such as prior record, length of incarceration, and ethnic group (i.e., race)—appear to be highly intercorrelated. The PSI study does not state the precise correlations among its independent variables, but some of its raw data indicate this degree of intercorrelation (e.g., most individuals—71.1% (1,972)—who have the most serious prior records also received long terms of incarceration, while only about 6% of the entire sample received a long term of incarceration and had no prior record). See Kolodney, et al. (1970: III-20). In an analysis of the PSI data which is planned later these precise correlations will be ascertained and further tests for multicillinearity will be applied. This makes it difficult to assess their relative impact on recidivism with true precision.

Significantly, several of the independent variables in the Beattie and Bridges analysis also appear to be highly intercorrelated. This may weaken some of the conclusions based on their data concerning the relative impact of these variables, especially those other than type of treatment prescribed by the judge. Therefore, the non-experimental analysis of the factors shaping recidivism seem to indicate that on the whole the type of treatment has a major impact. However, they also indicate that other factors such as type of offense, prior record, race, age, and narcotics history also have a major impact. Some aspects of these characteristics such as the absence of a prior record, having a narcotics history and certain offenses such as auto theft and drug law violations seem to have a greater impact on recidivism than does the type of treatment. However, this last point is stated with only a moderate degree of confidence because of the inherent limitations in the data which were discussed above.

Moreover, the goal of policy evaluation which leads to policy prescription, the absolutely precise analysis of the relative impact of all of these variables may not be of primary importance. Instead, the complete range of variables affecting the dependent variable must be ascertained first. It is insufficient to ascertain that a certain type of treatment has a certain effect on recidivism, when all other factors are equal or are controlled. In the real world all other factors are rarely equal; in reaching his decision a criminal court judge is faced with an individual with several characteristics. Knowledge of the general relationship between a type of treatment and recidivism is an insufficient policy guide because the judge's decision is not likely to be a general one; to a significant degree it will be relative to the individual before him. Similarly, for the prupose of policy prescription it is insufficient to simply ascertain which variable is the "best predictor".

In summary, the non-experimental analyses of the factors shaping recidivism seem to indicate that on the whole the type of treatment has a major impact. However, they also indicate that other factors, such as type of offense, prior record, race, age, and narcotics history, also have a major impact. These analyses also indicate that on the whole those offenders who are granted probation generally have significantly lower rates of recidivism than those who have been incarcerated. This pattern generally tends to persist when offender characteristics and type of offense are controlled.

However, this general finding of lower recidivism rates for those granted probation, even when these other factors are controlled, does not necessarily indicate that the lower rates are a specific function of this type of treatment. Instead, this relationship may be largely an artifact of the court's decision-making process. It is possible that those granted probation have lower recidivism rates because, first, those individuals with "favorable" offenses and characteristics (e.g., the absence of a prior record) are generally granted probation and, second, those individuals with these "favorable" offenses and characteristics are most likely to have lower recidivism rates.

In short, the judge's decision concerning type of treatment to prescribe tends to coincide — that is his intention — with the actual correlation between offender characteristics and recidivism. Indeed, the judge usually bases his decision on offender characteristics and type of offense. Thus it is possible that it is the offender's characteristics rather than anything inherent in the type of treatment, or anything inherent in being given one's freedom when probation is granted, that is the primary influence on recidivism. This suggestion would apply in an analogous manner to those incarcerated whose higher recidivism rates may be largely a function of their "unfavorable" characteristics, such as a serious prior record.

The data analyzed above, which tentatively indicate that a few characteristics and types of offenses may have greater impact on recidivism than the type of treatment received, in part tend to support this suggestion. On the whole, however, it does not seem possible to test this suggestion properly because of insufficient variation among the data points for several independent variables. For example, there are very few individuals with no prior record who are incarcerated; or, conversely, there are few Negroes with the following combination of characteristics who are granted probation: a serious prior record and the commission of a drug law violation. Moreover, for the purpose of policy evaluation and prescription, the possibility that the relationship between type of treatment and recidivism rates may be an artifact of the court's decision-making process may not be fully relevant. For this purpose it is insufficient to simply ascertain which factor is the "best predictor."

Policy makers need information about the explanatory factors over which they have some control. These factors may predict an outcome less perfectly, but they will probably give the policy maker greater ability to affect the outcome. A judge, or any other policy maker, can do little to change an offender's age or his number of prior convictions, but he can prescribe the precise type of treatment (probation or incarceration) which he will receive. The factors influencing educational achievement which are analyzed in the Coleman Report are another example of this pattern. Even if a student's family background is the best predictor of educational achievement, it is difficult for policy makers to influence this factor. By contrast, they do have some control over the size of his class in school, which in Coleman's analysis seems to have been a less important predictor of educational achievement. This pattern again seems to indicate that the requirements of analysis are different for pure social science and applied or policy social science.

For example, in the area of state and local governments' outputs, as Levy, Meltsner, and Wildavsky point out, even the best analyses focus on factors that are not useful for policy social science or policy-makers:

> They say little about the allocation process itself and, therefore, do not identify particular levers which might be used to alter policy outcomes. Clark's fine analysis of fifty-one American communities investigates (among other things) the effect of several independent variables on a dependent variable which used general budget expenditures as a measure of policy output. His most "influential" explanatory variable was the percentage of the city's population who were Catholic, particularly Irish Catholic. From a policy perspective, if a community wants to increase its budget, should one suggest that it import Catholics from Ireland and solve two problems with one recommendation? Or, would proselytizing Jews and Protestants help? Many of the more prosaic findings, such as the relationship between city expenditure and

citizen median income, display a similar lack of policy direction. What good does it do for a mayor to know that if his city were richer it could spend more? This is not to suggest that demographic variables are unimportant in determining municipal outcomes, or that it is not essential to learn about the constraints that bind. Rather we say that, for purposes of policy, it is important to study those variables which are under the agency's (or at least someone's) control (Levy, Meltsner and Wildavsky, pp. 13-24).

Consequently, scientifically rigorous principles often will receive less emphasis in policy social science than in pure social science. What are serious cognitive limitations for the latter need not be for the former. Even if in some instances pure social science were able to deleniate the causes of a social problem with greater precision and certainty, often it is beyond the policy-maker's ability to affect these causal factors. To make a significant contribution to policy analysis, social scientists should therefore broaden their focus beyond parsimony and minimalism. This perhaps especially applies to the political scientists who recently have been over-burdened with "scientific" requirements, many of which may be less relevant for policy analysis situations.

The experimental evaluation of recidivism had several scientific flaws and thus could not establish causal relationships between sentencing and recidivism that were unambiguous. Yet this experiment suggests promising policy opportunities. This is a reflection of a third element of the distinction between pure and policy social science: analysis of social phenomena that is scientifically flawed nevertheless can have significant heuristic value to the policy-maker. His goal is action--specifically the ability to alter outcomes successfully. The pure social scientists goal is the attainment of knowledge. Pure and policy social science do not differ in their degree of clear analytical thinking, but rather in these goals. And thus policy social science does not <u>end</u> with the attainment of knowledge, it rather <u>begins</u> with it. Policy social science cannot exist without pure social science. Yet, as the following discussions of the CTP experiment and the later discussion of the limitations of all evaluations suggest, the results of pure social science are a necessary foundation but not a sufficient condition for successful policy-making.

The problems caused by the possibility that an apparent relationship is an artifact of the treatment process being analyzed are endemic to the analysis of non-experimental data. For example, if there is this insufficient variation among the data points of some of the independent variables, any type of statistical controls are of little help. An alternative method of analyzing the relationship between type of treatment and recidivism is conducting a controlled experiment. In this way the decision to grant probation or incarceration is not "contaminated" by a real decision maker.

EXPERIMENTAL EVALUATIONS OF ALTERNATIVE SENTENCING POLICIES

A controlled random experiment can isolate the effect on recidivism of the alternative types of treatment as opposed to the effect of a type of treatment linked to a type of individual — one who has been directed to that type of treatment by a judge. The various possible independent variables other than the program in question are controlled through a random application of that program or treatment to a population. The data are thus controlled by randomization at the outset rather than in an ex post facto manner (e.g., regression analysis) as in non-experimental research. As opposed to evaluation of actual policy decisions, a controlled random experiment of sentencing can do two things. It can "randomize" the offender characteristics of its population in advance—assuring, for instance, that enough blacks with narcotics histories <u>are</u> granted probation so that a researcher can evaluate the effect on recidivism of incarceration or probation <u>per se</u>, rather than only that of incarceration combined with blackness and past narcotics use. And more generally, the experiment can guarantee that the decision on whether to grant probation is not "contaminated" by a real-life decision-maker and his views of the recidivism risk of certain offenders and offenses.

The California Youth Authority has recently been conducting a controlled experiment in the cities of Stockton and Sacramento to evaluate the effectiveness of alternative treatment programs for convicted juveniles. At the level of general strategies for policy evaluation, the results and methods of this experiment — "The Community Treatment Project" (CTP) — seem to be indicative of both the potentialities and some of the drawbacks of experimental methods of evaluation (Warren, 1967). At the level of scientific evaluation of alternative sentencing policies and specific policy strategies for reducing recidivism, the CTP results and methods are very useful and suggestive.

The CTP experiment involves an initial screening of convicted juvenile delinquents. The remainder are then randomly assigned either to an experimental group which is returned to the community (i.e., receive probation) and receives intensive counseling, or to a control group which is assigned to California's regular juvenile penal institutions. Seventy to eighty percent of those in the experimental group resided in their own homes. The remainder were placed in a foster or group home because it appeared to the CTP investigators that they could not live in their own home and remain non-delinquent. These 20-30% usually spend at least part of the time in their own home, but their lives generally are somewhat more constrained. See Warren (1967: 5). However, it does not seem that this constraint is significant enough to suggest that they are no longer experiencing freedom. Their experience is still much like that of those in the experimental group who live at home and it is still radically unlike that of those in the incarcerated or control group.

After a follow-up period of fifteen months the "failure" rate for the experimental group was 28% (134) and 52% (168) for the control group; after twenty-four months the respective "failure" rates were 38% and 61%. ("Failure" was defined here more inclusively than in the studies described above, such as Beattie-Bridges. It consisted of parole revocation which included "serious" violations [e.g., new felony convictions and/or new incarceration] and "technical" violations which did not always involve an arrest. This may explain the lower "failure" rate — 10.9% — for the probation group in Beattie-Bridges, which becomes 34.2% when the Beattie-Bridges data are analyzed according to the CTP definition of "failure.") Personal and attitudinal change as reflected in psychological test scores were also measured during this period. The experimental group was also more "successful" according to this standard. Tests administered both at intake into the Youth Authority and after release (after treatment in the case of the experimental groups and after discharge from institution in the case of the control group) indicated that "although both groups showed improvement from pre-test to post-test, the experimental group showed considerably more positive change than the control group, together with a higher level of personal and social adjustment." See Warren (1967: 7).

However, as is often the case in experimental evaluation, the CTP experiment seems to have been flawed in four significant respects. First, the initial screening eleminated about 25% of the convicted male juveniles (and 10% of the females) for whom institutionalization was deemed mandatory because they were involved in serious assault cases or because there was community objection. This clearly limits the generality of the conclusions that can be drawn: These data indicate that recidivism is less likely if offenders receive probation, but we do not know if this applies to the most serious offenders.

Second, there seems to have been ambiguous specification of the independent variables in the creation of the experimental design. Those in the experimental group receive both probation and intensive counseling. Thus there is no way to ascertain which of these aspects of their treatment is related to their lower failure rates. To do this an additional experimental group should have been created which received probation but no counseling at all.

The second flaw described may seem to be a description of the

Hawthorne effect, but it is not. It referred to ambiguous specification of the independent variables in which the experimental design included in effect two independent variables (counselling and probation) in one type of treatment. This made it difficult to ascertain which aspect of the treatment was affecting the recidivism rate. It is incidental to this flaw that one of the ambiguously specified variables--counselling--is in itself somewhat of a Hawthorne effect.

The CTP investigators seem to have been somewhat aware of this flaw in retrospect. In a proposed new phase of the experiment they have suggested the use of a second experimental group. However, this suggestion does not meet the specific criticism made here. This group will not be incarcerated in the regular California institutions, but it will be sent to a special treatment center in their community and receive intensive counselling there. This group will not be actually released to the community (i.e., it will not receive probation.) Thus this second experimental group only allows greater specification of the independent variables of incarceration or semi-incarceration. For a description of this proposed new phase see T. B. Palmer, A Proposal for Phase 3 of The Community Treatment Project (Sacremento: California Youth Authority, 1969).

Third, because of the nature of the supervision of the experimental group it seems very possible that its lower "failure" rate is to some degree a function of experimenter effects such as the Hawthorne effect and the effects of a positive self-fulfilling prophecy and positive labeling.

Specifically, several aspects of the experimental group clearly have characteristics of the Hawthorne effect which can lower influence the group's "failure" rate. First, the youths in the experimental group receive intensive attention from a "community agent" (i.e., a probation officer) whose entire caseload is twelve youths, compared to a normal caseload of from four to eight times that amount. During the intensive stage of the treatment in the community, the youths see the agent from two to five times weekly, either individually or in group or family meetings. Second, they receive special types of attention in addition to these meetings: group and family therapy sessions, various group activities, and school tutoring by a certificated teacher experienced in working with delinquents. Third, they receive attention with a group-oriented focus. Much of this activity focuses around a program center which resembles a settlement house. (Palmer and Warren, 408). (The center houses the staff, provides a recreation area, classrooms, a music room and outdoor sports.)

Moreover, other aspects of the experimental group clearly have the potential for creating experimenter effects such as a positive self-fulfilling prophecy, positive labeling and even a positive experimenter bias which can lower the group's "failure" rate. First, the youths are not only aware that they are receiving intensive, special and group-oriented attention and that they are part of an experiment, but it seems possible that they are also aware that those in the experimental group are "supposed to act better" because they have had this "extra break of not being incarcerated". This would seem to create a positive labeling effect which could lower the "failure" rate for this group. This would be analogous to the often stated, though rarely systematically proven, view that incarceration and all the official and unofficial stigma attached, creates a negative labeling process which increases the "failure" rate of those labeled "prisoner" and "ex-con". Second, the decision to revoke probation for the experimental group is made by the community agents themselves. This is not unusual in itself because probation officers generally play a large part in this decision when it is usually made by the court.

This is not to suggest that the agents make these decisions in anything but a fair and conscientious manner and that they fail to attempt to uphold the standards of scientific objectivity that are necessary in an experiment. Indeed, the evidence concerning their intentions are clearly the contrary. However, the issue here is the possibility of a more subtle and unconscious factor such as the agent's expectations and their effects. It seems quite likely, first, that the agents expect the experimental group to do better, and second that they convey this expectation to the youths. In short, it seems likely that there is some degree of positive experimenter bias and a positive prophecy for these youths to "do better", and it seems that to some degree this may become self-fulfilling.

The reports of the CTP experiment do not present a great deal of direct evidence to support these speculations. However, the first possibility described above seems highly likely on the face of it, especially because it is quite possible that the youth's own awareness is reinforced by whatever positive expectations the community agents have for the experiment. Similarly, it is this type of expectation on the agent's part that would make the second possibility seem quite likely. There is some indirect evidence that indicates that the agents have this expectation. They are all probation officers who generally believe in probation, especially if it can occur in "ideal" and intensively supervised circumstances such as those in the CTP case.

Moreover, the CTP reports clearly state that in making probation revocation decisions, the agents often do not make revocations for minor misbehavior. They do, however, often "suspend" the probation of the experimental youths for such misbehavior. Suspension generally only involves serious warning. In practice, once the suspension is made, it is rare for revocation to follow. One possibility, although there is no direct evidence, is that the agents are unfairly and unscientifically lenient to the experimental youths in such situations. Moreover, even if they are not lenient in this manner, the agent's behavior may still affect the "failure" rate. For example, they may expect the youths to "come around" and avoid revocation because of the "second chance" they have had in general (and also because of the "second chance" they have had in the form of the suspension rather than revocation). Finally, on the basis of the literature on experimenter effects in experiments and quasi-experiments described above, it seems very probable that at least to some degree the agents convey these expectations to youths.

The fourth flaw in the CTP experiment seems to have been the initial selection of the population or universe from which the experimental and control groups were selected (after the screening out of the assaultive cases). The nature and characteristics of this population--convicted first offenders from Sacramento and Stockton - seem to weaken significantly the generality of the experiments. Neither city is typical of the large and heterogeneous urban areas from which the largest proportion of offenders come. Both cities are relatively small in comparison to Los Angeles, San Francisco, and Oakland; are not heavily industralized, and do not have large Negro populations (though Sacramento has modest numbers of Mexican Americans and Stockton has a sizeable number). The populations of Sacramento and Stockton in 1960 were 191,667 and 86,321 respectively. 7% and 10% of these populations respectively were Negro, 8.1% and 16.8% were Mexican-American and both had rather large portions of their labor force in white collar occupations (54.6% and 46.1% respectively). Indeed, the important and easily obtainable variables of race and ethnicity are not mentioned in the CTP experiment. In an apparent effort to remedy this flaw, phase 2 of the CTP experiment was extended to predominantly Negro areas of Los Angeles and Oakland.[18] However, for reasons that can only be speculated upon, phase 2 does not include random assignment of convicted delinquents. Instead the youths are assigned to the community treatment program after screening by the project staff. Moreover, there is no control group whatsoever. In the absence of a control group, the success of the program is measured by comparing the failure rate of youths assigned to it with equivalent statewide rates for youths of the same middle to older adolescent age range. After a follow-up period of fifteen months of parole exposure, the "failure" rate (defined as parole revocation) for the project's youths is 39% compared to 48% for the statewide group of that age.

The primary goal of both the CTP experiment and this paper is policy evaluation leading to prescription. As I have tried to indicate, the requirements of analysis seem to be different for pure social science and applied or policy social science. Thus, two of the scientifically flawed aspects of the CTP experiment are nevertheless quite promising possibilities for the goal of policy evaluation and prescription. However, let me emphasize that I clearly do not mean that invalid methods or findings should be tolerated when the investigator is primarily interested in policy evaluation and prescription. Accurate analysis and evaluation is the essential foundation of policy analysis. Yet as I will indicate, findings that are the product of somewhat less than perfectly controlled analysis may be of great heuristic value to the policy analyst. (The policy analyst's boldness and tolerance for uncertainty and imperfect findings ought to be tempered, however, by the awareness that his responsibility is even greater than a pure scientist's. The policy analyst's errors are much more costly — especially in immediate terms — than those of the pure social scientist. If a researcher is in error concerning the degree of pluralism in city X, then our understanding of the city's political process is faulty. However, if a researcher is in error concerning the impact of program X on a population and his evaluation is acted upon, then many resources will be misallocated and it is possible that the population may be deprived of a potentially beneficial program.)

First, perhaps, the initial screening out of about 25% of the convicted male juveniles (and about 10% of the females) for whom institutionalization was deemed mandatory because of their assault background, points to a general policy prescription: Screen out such cases, and then grant probation to all other juvenile first offenders. According to the CTP findings, probation leads to less recidivism. Thus probation for all but those screened out by the above criteria should significantly lower the present "failure" rate.

Second, perhaps the possibility that the CTP's low "failure" rate for the experimental group is to some degree the function of experimenter effects such as the Hawthorne effect and the effects of a positive self-fulfilling prophecy, positive labeling and even a positive experimenter bias can be utilized as an explicit and intentional positive policy. If the community agents' expectations of "success" for youths in a CTP type program and if a positive label (such as being a CTP participant) and an unconscious and unintentional experimenter bias toward the youths can lower the "failure" rate, then perhaps a program should be created which focuses explicitly and intentionally on such efforts. Indeed, though it is rather bold, a possible policy prescription flowing from this is that after an initial screening out of assaultive cases, all juvenile first offenders should be granted probation and assigned to this explicitly and intentionally "Hawthorne" and "positive self-fulfilling prophecy" program. According to the above analysis and hypotheses, this could significantly lower the present overall "failure" rate, which is based on both offenders who are granted probation without a special program and those who are incarcerated. The potential of this proposed experimenter effect of positive labeling for lowering the "failure" rate should not be judged in comparison with an alternative of "no labeling". It should be judged in comparison with the negative labeling which offenders receive when they are incarcerated. This label often remains with them for a considerable subsequent period because of the informal and official stigma attached to their previous incarceration. Indeed, the list of official stigma and lost rights of felony offenders who are simply convicted but not necessarily incarcerated is considerable. In 46 states the rights lost include the right to vote, serve on a jury, hold public office or a position of trust or certain other kinds of employment, obtain certain licenses and hold public employment. Furthermore, administrative policies create many bars. For example, offenders are excluded from the Job Corps, the Neighborhood Youth Corps and other OEO projects. (Task Force Report: Corrections, op. cit., pp. 171 and 204.) Investigations in the literature on experimenter effects which range from Rosenthal's "Pygmalion in the classroom" study to the use of "placebo effects" in medical science indicate that positive expectations, prophecies, biases and labeling can be conveyed to a subject and can effect his behavior positively. In the "Pygmalion study" Rosenthal and Jacobson conclude:

"Our experiment rested on the premise that at least some of the deficiencies--and therefore at least some of the remedies--might be in the schools, and particularly in the attitudes of teachers toward disadvantaged children. In our experiment nothing was done directly for the child...The only people affected directly were the teachers: the effect on the children was indirect." By contrast I am not necessarily suggesting that present non-experimental "failure" rates are a function of deficiencies in present probation officers. This may be so to some degree. However, the premise of this policy is simply that whatever the abilities of the present probation officers, an explicitely and intentionally "Hawthorne" and "positive self-fulfilling prophecy" program carried out by probation officers possibly could lower the present "failure" rate.

If this negative labeling which results from these stigma does affect "failure" rates negatively, then another significant policy innovation (though less bold than that proposed above) would be to at least minimize the official stigma. The policy could vary from absolute secrecy concerning an individual's previous conviction (except for release to criminal justice agencies) to an official annulment of this record after the individual has successfully completed a period of probation or parole. Indeed, a few states now have versions of the latter proposal. Again, since our primary goal is policy prescription, it is sufficient to know that they can be conveyed and it is not initially necessary to fully understand how and why they are conveyed. Of course, research should nevertheless seek this latter knowledge.

Charles Tittle has proposed a very different policy for reducing recidivistic and even initial criminal behavior which is also based on the concept of labeling. He suggests that deterrence can occur through the certainty of punishment. He acknowledges that "labeling [an individual] a deviant may result in greater deviant behavior" but he adds that "it may also result in less deviance by those who observe his stigmatized status... Application of sanctions may result in identities which influence the conduct of those not so labeled." (Charles Tittle, "Crime Rates and Legal Sanctions," Social Problems v. 16 (spring 1969) p. 421.) Based on my experience in doing empirical research on police and criminal courts and my knowledge of the empirical literature on the attitudes of offenders, potential offenders, and the lower-class subculture, my own view is that Tittle's reasoning is based on a profound misunderstanding of the values, life styles and attitudes of offenders and potential offenders. It seems generally remote possibility that the behavior of a young, poor, male minority group member with a present-oriented time-horizon could be significantly affected by a high certainty of punishment which will cause him to be negatively labeled sometime in the future.

The suggestion of direct application of experimenter effects to policy programs is not new. The original Hawthorne experiments were directed toward this end, and in fact they greatly changed policies for employee-management relations. Similarly, Rosenthal and Jacobson's investigations have been explicitly directed toward the creation of policy programs. Indeed, at the conclusion of their "Pygmalion study", they specifically suggest methods of converting their findings into policies for teacher training and classroom strategies.

Since policy rather than purely scientific considerations have been emphasized as the primary goal of this paper, the policy prescription suggested here must be considered in terms of the realistic policy constraint of cost. A general evaluation of alternative sentencing on the basis of cost-effectiveness will be described in part III below. A brief discussion of the cost of the prescribed policy without consideration of effectiveness will suffice at this point. At present there are no cost data available for Phase 1 of the CTP experiment. However, cost data for Phase 2 of CTP are available, and they can give us an approximate idea of the costs of CTP Phase 1 and of the costs of the policy suggested here relative to the costs of alternative policies. In Phase 2 the probation officers have caseloads of fifteen youths per officer (the caseload is twelve in Phase 1), and this costs $150 per month per boy which is three to four times as much as regular probation. However, it is still less than half the average monthly cost of institutionalizing an offender. Phase 2 handles a group that is larger than the capacity of one of the new institutions that the Youth Authority is building at a cost of six to eight million dollars. This type of comparison of alternative costs has not been done by those that have criticized Phase 1 of CTP as being impractical for wide application because of its cost.

Of course, this type of comparison does not consider the probabilities of a further cost to society of the CTP experiment; namely, the probability that the offender will recidivate while on probation. By contrast, there is almost a zero probability that an offender will recidivate while incarcerated. There will be a discussion below of these types of trade offs which are involved in these alternative policies.

CHAPTER 4: SOME POLICY IMPLICATIONS

Both the non-experimental and experimental data analyzed above seem to indicate that, on the whole, those convicted individuals who are granted probation have lower recidivism rates than those who have been incarcerated. However, offender characteristics and type of offense committed — especially certain characteristics and certain offenses — also seem to have a significant impact on recidivism.

One can summarize some of the major findings and some of the implied recommendations of the analyses and experiments mentioned above in the following way: For many persons, especially those with certain "favorable" characteristics (e.g., the absence of a prior record), probation can reduce the recidivism rate to approximately 33 per cent. The experiment indicates that for most persons, probation along with intensive and special attention can reduce recidivism to 28 per cent. For most convicted felons, therefore, the type of treatment makes a significant difference. For other felons, however, personal characteristics and the nature of their offense seem more important. The influence of probation on recidivism is thus far from total; but it is clear that knowledge of recidivism rates associated with specific offenses and particular offender characteristics could be of considerable practical value to judges in sentencing.

On the other hand, it must also be emphasized that even among probationers there is a recidivism rate of approximately 33 per cent. This figure represents a very large number of individuals and crimes, and it, as well as the fact that probationers recidivate less than those who are incarcerated, must be taken into account in designing sentencing policies based upon the findings of social science.

In addition, the experimental data seem to suggest that if probation programs can intentionally and explicitly develop "Hawthorne" effects and effects of positive self-fulfilling prophecies and positive labeling, then it may be possible that recidivism rates can be kept relatively low. For example, if there is initial screening of offenders to eliminate the most serious and dangerous offenders, this program of an intentional "Hawthorne effect" may be able to keep the "failure" rate below 30%.

However, one should not be discouraged by the ability of these studies to establish a causal relationship between sentencing and recidivism. Because one is unable to establish such a clear connection does not in itself destroy their value to policy makers. This is because "pure" social science and "policy" social science are different enterprises with different requirements. That is to say what may be serious problems need not be such serious problems to the latter. For example, the CTP's initial screening process may limit the scope of its conclusions, but its finding that probation lowers recidivism for the non-assaultive suggests a sentencing policy that is likely to reduce recidivism: Screen out those with a history of assault, then grant probation to all other juvenile first offenders. Similarly, the probable existence of a "Hawthorne" effect in the CTP experiment, though it is unquestionably a flaw from the standpoint of pure social science, also suggests a strategy for treating juveniles: Juvenile first offenders who are granted probation should be assigned to an intentional "Hawthorne" and "positive self-fulfilling prophecy" program. Admittedly, it is difficult to institutionalize for large numbers of people a feeling of being part of a special experiment and the subject of special attention; but precedents such as the Hawthorne experiment itself indicates that that it is clearly possible, and the CTP experiment indicates that such a program will succeed in reducing recidivism. (It will also be cheaper. The costs per month in the CTP program seem to be about half the per capita cost of incarceration).

However, to convert these findings into policy guidance for criminal court judges and to apply them to the evaluation of a specific set of courts, the question of the goals of the criminal court must be analyzed. In addition to reduced recidivism, these seem to include maintaining order and stability in society, maintaining the freedom of the individual, satisfying a common notion of justice (i.e., equality and consistency of treatment), maintaining an image of the court as a fair institution, maintaining the "declarative" nature of the criminal law (i.e., the criminal law is in large part more intended to be a list of acts that society wishes to "declare" inappropriate rather than a list of acts against which it wishes full enforcement), and maintaining a favorable cost-effectiveness outcome for the courts' decisions. One must then note that many of these goals are by no means fully consistent with the goal of reducing recidivism. There is in fact a great deal of tension between these various goals. The following brief examples will illustrate this tension.

First, lower recidivism rates may be associated with a policy of probation such as the one proposed in the critique of the CTP experiment. This policy probably would satisfy the goal of reduced recidivism more than would increased incarceration. Nevertheless, it also risks significant short-run sacrifices in the goal of order and stability in society because it gives freedom to many convicted individuals who have a reasonably high probability of recidivating. One must remember that almost one-third of the CTP probationers did recidivate despite the special attention and intensive supervision they received. Incarceration may have a small or negative effect on reducing recidivism. However, by denying the freedom of some individuals — especially those with a reasonably high likelihood of recidivating — it does tend to satisfy the goal of maintaining order and stability in society, at least in the short run. There is almost a zero probability of an offender recidivating while incarcerated. (The policy of probation suggested in the critique of CTP would mean a low number of incarcerations, and thus it probably would also involve sacrifices in the achievement of the goal of maintaining the "declarative" nature of the criminal law.)

Second, a policy to reduce recidivism may involve sacrifices in other goals even if it does not involve granting probation more frequently. For instance, from what we know about the type of offenders who are most likely to fall into the recidivating group, one could derive the following policy to reduce recidivism: Incarcerate for the longest terms the youngest offenders, especially if they are black or have a narcotics history. But such a policy, however effective it might be in reducing recidivism, is obviously unacceptable if the court is to remain in our eyes a fair and non-discriminatory institution which exercises a due regard for equality and individual liberties. Conversely, the same findings of social science with regard to reducing recidivism would dictate that judges incarcerate for the shortest terms possible under the law whites over 40 who have committed murder. Since this group has an extremely low recidivism rate, this policy would involve only a small risk of sacrifice in the goal of reduced recidivism. In addition, such a policy would also save the state money in incarceration costs. Or to take some cases that typically face criminal court judges, from these evaluations one could also derive the policy of incarcerating for the shortest terms whites who have committed sex offenses or crimes against persons and are over 40 years old. Again since these offenders have very low recidivism rates, this policy would involve a small risk of recidivism. (It would also maximize the cost-effectiveness of the courts' decisions.) However, both these policies would conflict with the goals of maintaining the "declarative" and condemnatory nature of the criminal law and maintaining the image of the court as a fair and just institution. Even if these policies did not increase recidivism, most of society would feel that the shortest terms for these offenders was somehow wrong. Not because they are vengeful, but because they probably would feel that these policies did not sufficiently express society's condemnation of these offenses. They probably would also feel that different treatment for men who had committed the same offense--especially if one of the criteria were racial--was unjust.

In addition, as another example, from the findings of this paper one could derive the following policy to reduce recidivism: Incarcerate, until they reach the age of 30 or 35, all individuals who commit their second felony offense. It is likely that this would reduce recidivism because the data indicate that after this age there is a sharp reduction in the probability of recidivating. However, this policy probably would contribute to the image of the court as an unfair institution. It would also involve sacrifices in the goal of maintaining a favorable cost-effectiveness outcome for the courts' decisions because of the immense capital and maintenance costs of incarceration. Indeed, the same amount of reduced recidivism achieved by this policy of incarceration until the age of 30 or 35 could probably be achieved by a probation policy at almost one-half the cost.

Third, a brief evaluation of the sentencing decisions of the criminal court judges of Pittsburgh and Minneapolis indicates the difficulty in evaluating the most effective policy to reduce recidivism. An earlier study that I conducted indicated that sentencing decisions are more lenient in Pittsburgh than in Minneapolis. White and Negro defendants receive both a greater percentage of probation and a shorter length of incarceration in Pittsburgh. This pattern persists when the defendants' previous record, plea, and age are also controlled. Although both white and Negro defendants receive more lenient sentences (i.e., more frequent grants of probation) in Pittsburgh in both cities whites receive more lenient sentences than Negroes. However, this difference in the direction of greater leniency for whites is very small in Pittsburgh, while it is large in Minneapolis. Also, in Minneapolis defendants with a prior record receive a much lower percentage of probation and a much longer length of incarceration than do defendants with no prior record. In Pittsburgh, on the other hand, defendants with a prior record (with the exception of Negroes in a few categories) generally receive only a slightly lower percentage of probation and only a slightly longer length of incarceration than defendants with no prior record.

On this basis one might conclude that the Pittsburgh judges' decisions, on the whole, tend to contribute more effectively to reduced recidivism because they grant probation more frequently. However, their frequent grants of probation for individuals with a high probability of recidivating (e.g., those with a prior record and Negroes) probably does not effectively contribute to reduced recidivism. By contrast, the Minneapolis judges' generally severe decisions for these specific individuals may contribute to reduced recidivism more effectively.

(Elsewhere I have attempted to evaluate the decision making of these two courts in terms of the multiple goals of the criminal court [Levin, forthcoming: Ch. 10]).

This effort to systematically evaluate the consequences of alternative sentencing policies for recidivism should clarify the nature, limitations and potentialities of a policy social science. Social science evaluation can serve as a partial guide for the policy-maker, but as the analysis of the tensions in the goals and values of the criminal court indicates, it cannot serve as the definitive and ultimate guide because these tensions cannot be resolved on utilitarian grounds. Social science evaluation in itself cannot give policy guidance; it does not yield self-explanatory policy choices. It can only indicate the consequences of alternative policies, their utility and disutility. The evaluations here indicate that an optimal policy for the reduction of recidivism is to incarcerate for the longest terms the youngest offenders, especially if they are black or have a narcotics history. This probably would reduce recidivism, but it would also cause sacrifices in achieving other goals of the court. Thus for recidivism, and for other policy areas, ultimately we will have to choose our priorities among these multiple goals and values, and the trade-off rate among them that we wish to follow. Social science evaluation cannot do that for us.

In summary, there is considerable tension among the goals of the criminal courts, as usually is the case with basic institutional goals and values. Indeed, few important goals and values in society can be simultaneously maximized. It is this tension which makes a consideration of these goals and values so fascinating and perplexing. However, in terms of the single goal of reduced recidivism, this study has attempted to offer more empirical guidance to decision makers and policy evaluators. Yet to achieve this goal, policy makers must also look beyond the criminal courts. As this study has indicated, factors other than court decisions also have a major impact on recidivism. The courts cannot and probably should not affect these factors.

NOTES

1. For a technical discussion of multicollinearity, see Johnston (1963: 201-207).

2. See Bowles and Levin (1968), which also deals with other methodological and statistical problems beyond the scope of our discussion here.

3. See Tufte (1969: 653) and Blalock (1962) for examples and discussion of multicollinearity in cross-tabulation analysis.

4. See Johnston (1963) and Farrar and Glaubner (1967: 92-106).

5. See Blalock (1963: 233) for additional examples.

6. Johnston (193-207).

7. See the President's Commission on Law Enforcement and the Administration of Criminal Justice (1967: 261).

8. See Campbell (1969) for a description of quasi-experiments with social policies and a bibliography of this field.

9. There are several detailed descriptions of the Western Electric experiments and explanations of the Hawthorne effect. For example, see George Homans, The Human Group (1950) and F. J. Roethlisberger and W. J. Dickson, Management and the Worker (Cambridge: Harvard University Press, 1939).

10. See Arthur Shapiro, "A Contribution to a History of the Placebo Effect", Behavioral Science (1960) pp. 109-135, for review of the history of the placebo effect, and see M. Greenblatt "Controls in Clinical Research" (Tufte's University Medical School, 1964) for a fascinating account of the placebo effect in the use of the controversial drug Krebiozen.

11. See Rosenthal (1966) and Rosenthal and Jacobson (1962). Several other similar investigations by Rosenthal and others are cited in these two works. The highlights of a significant amount of their work in various contexts is described in Rosenthal and Jacobson (1968b).

12. See Rosenthal and Jacobson (1968b: 22-23) for a description of other controls which were used and other tests of alternative hypotheses.

13. R. Rosenthal in Edward Tufte, Quantitative Analysis of Social Problems, pp. 164-165.

14. See Howard S. Becker, Outsiders, New York: Free Press, 1963; John I. Kitsuse, "Societal Reaction to Deviant Behavior: Problems of Theory and Method," Social Problems, 9 (Winter, 1962), pp. 247-256; and Edwin M. Lemert, Human Deviance, Social Problems, and Social Control, Englewood Cliffs, New Jersey: Prentice-Hall, 1967, Chaps. 1 and 3.

15. I wish to thank the following individuals who graciously helped to provide the data which appears in this section: Ronald Beattie, Marie Vida Ryan, Charles Bridges, William Hutchins, Robin Lamson (all of the state government of California); State Assemblyman Craig Biddle of Riverside, California; Steven Kolodney (Public Systems Inc.), Don Gottfredson (National Council of Crime and Delinquency), Richard McGee (American Justice Institute), Charles Eichman (Florida Department of Corrections), H. P. Higgins and Carole Bartholemew (Minnesota state government), and John Yeager (Pennsylvania Department of Justice).

16. The larger project of which this study is a part will analyze the data collected by Beattie and Bridges in more detail than was possible in their own study. For example, additional characteristics will be controlled simultaneously. Regression analysis of their data will also be carried out to assess more precisely the relative effect on recidivism of the various offender characteristics and types of treatment. A preliminary effort at such an assessment is described below.

17. Internal memoranda of the California Department of Corrections, May 1, 1970 and April 20, 1967. I am indebted to Marie Vida Ryan, Senior

Statestician of the CDC for graciously providing these data and many other aids to this study.

18. U. S. Task Force Report on Corrections, p. 42 (1967, Washington, D.C.)

CASES

Gideon v. Wainwright 372 U.S. 335 (1963).

REFERENCES

BEATTIE, Ronald H. and Charles K. BRIDGES (1970) Superior Court Probation and/or Jail Sample. Sacramento: Bureau of Criminal Statistics, Department of Justice.

BECKER, Howard S. (1963) Outsiders, New York: Free Press.

BECKER, Theodore L. (1969) The Impact of Supreme Court Decisions. Oxford: Oxford University Press.

BLALOCK, Herbert Jr. (1963) "Correlated Independent Variables: The Problem of Multicollinearity," 62 Social Forces 233.

BOWLES, S. and H. LEVIN (1968) "The Determinants of Scholastic Achievement," 3 Journal of Human Resources 3.

CAMPBELL, Donald T. (1969) "Reforms as Experiments," 24 American Psychologist 409.

COHEN, David (1970) "Politics and Research: The Evaluation of Social Action Programs in Education." 40 Review of Educational Research 213.

COLEMAN, James S., et al. (1966) Equality of Educational Opportunity. Washington, D.C.: Government Printing Office.

DAVIS, George (1964) "A Study of Adult Probation Violation Rates by Means of the Cohort Approach," 55 Journal of Criminal Law, Criminology and Police Science 70.

DYE, Thomas (1966) Politics, Economics, and the Public. Chicago: Rand McNally.

EICHMAN, Charles J. (1966) The Impact of the Gideon Decision Upon Crime and Sentencing in Florida. Tallahassee: Florida Division of Correction.

ENGLAND, Ralph W. Jr. (1957) "What is Responsible for Satisfactory Probation and Post-Probation Outcome?" 47 Journal of Criminal Law, Criminology and Police Science 667.

EVANS, John (1969) "The Westinghouse Study: Comments on the Criticisms," in David G. HAYS, Britannica Review of American Education, Vol. 1. Chicago: Encyclopaedia Britannica.

FARRAR, Donald E. and Robert R. GLAUBNER (1967) "Multicollinearity in Regression Analysis: The Problem Revisited," 49 Review of Economics and Statistics 92.

FRY, Brian and Richard WINTERS (1970) "The Politics of Redistribution." 64 American Political Science Review 508.

GLASER, Daniel (1964) The Effectiveness of a Prison and Parole System. Indianapolis: Bobbs Merrill.

HOMANS, George (1950) The Human Group, New York: Harcourt, Brace.

JACOB, Herbert and Michael Lipsky (1968) "Outputs, Structure and Power: An Assessment of Changes in the Study of State and Local Politics," 30 Journal of Politics 510.

JACOB, Herbert and Kenneth VINES (1965) Politics in the American States, Boston: Little, Brown.

JOHNSTON, John (1963) Econometric Methods. New York: McGraw-Hill.

KITSUSE, John T. (1962) "Societal Reaction to Deviant Behavior: Problems of Theory and Method," 9 Social Problems 247.

KOLODNEY, Steven, et al. (1970) A Study of the Characteristics and Recidivism Experience of California Prisoners. San Jose: Public Systems Incorporated.

LEMERT, Edwin M. (1967) Human Deviance, Social Problems, and Social Control. Englewood Cliffs, N.J.: Prentice Hall.

LEVIN, Martin A. (forthcoming) Urban Political Systems and Judicial

Behavior: The Criminal Courts. Cambridge: Harvard University Press.

.... (1970 "An Empirical Evaluation of Urban Political Systems: The Criminal Courts," in Sam KILPATRICK and David MORGAN, Urban Politics: A System Analysis. Glencoe: Free Press.

MEYER, John R. and Edwin KUH (1957) The Investment Decision: An Empirical Analysis. Cambridge: Harvard University Press.

MILLER, Walter (forthcoming) City Gangs. New York: John Wiley.

MUIR, William Jr. (1967) Prayer in the Public Schools: Law and Attitude Change. Chicago: University of Chicago Press.

ORCUTT, Guy, et al. (1961) Microanalysis of Socioeconomic Systems: A Simulation Study. New York: Harper.

PENNOCK, J. Roland (1966) "Political Development, Political Systems, and Political Goods," 18 World Politics 415.

PRAIS, S. J. and H. S. HOUTHAKKER (1955) The Analysis of Family Budgets. Cambridge, England: University Press.

President's Commission on Law Enforcement and Administration of Criminal Justice (1967) The Challenge of Crime in a Free Society, Washington, D.C.: Government Printing Office.

ROETHLISBERGER, Fritz J. and William J. DICKSON (1939) Management and the Worker. Cambridge: Harvard University Press.

ROSENTHAL, Robert (1966) Experimenter Effects in Behavioral Research. New York: Appleton-Century-Crofts.

ROSENTHAL, Robert and Lenore JACOBSON (1968a) Pygmalion in the Classroom. New York: Holt, Rinehart and Winston.

.............................(1968b) "Teacher Expectations for the Disadvantaged," 218 Scientific American 19.

SHAPIRO, Arthur (1960) "A Contribution to a History of the Placebo Effect," 5 Behavioral Science 109.

STONE, Richard (1954) The Measurement of Consumers' Expenditure and Behavior in the United Kingdom. Cambridge, England: University Press.

TUFTE, Edward (1969) "Improving Data Analysis in Political Science," 21 World Politics 641.

TUKEY, J. W. (1954) "Causation, Regression, and Path Analysis," in Oscar KEMPTHORNE, et al., Statistics and Mathematics in Biology, Ames, Iowa: Iowa State College Press. Quoted on Page 237 of Herbert Blalock, Jr. (1963) "Correlated Independent Variables: The Problem of Multicollinearity," 62 Social Forces 233.

U.S. Commission on Civil Rights (1967) Racial Isolation in the Public Schools. Washington, D.C.: Government Printing Office.

U.S. Task Force on Corrections (1967) Task Force Report: Corrections. Washington, D.C.: Government Printing Office.

WARREN, M. Q. (1967) The Community Treatment Project After Five Years. Sacramento: California Youth Authority.

WASBY, Stephen L. (1970) The Impact of the U.S. Supreme Court, Homewood, Ill.: Dorsey Press.

Westinghouse Learning Corporation Study (1969) The Impact of Head Start: An Evaluation of the Effects of Head Start on Children's Cognitive and Affective Development.

WILSON, James Q. (1968) City Politics and Public Policy. New York: John Wiley.

.... (1967) "The Crime Commission Reports," 9 The Public Interest 64.

.... (1964) "Problems in the Study of Urban Politics," in E. H. BUEHRIG, Essays in Political Science. Bloomington: Indiana University Press.