12/19/74

INSTITUTE OF GOVERNMENTAL RESEARCH
UNIVERSITY OF WASHINGTON

PUBLIC POLICY PAPER No. 2

THE CAPACITY OF SOCIAL SCIENCE ORGANIZATIONS TO PERFORM LARGE-SCALE EVALUATIVE RESEARCH

by

Walter Williams

April 1971
$2.00

3935 University Way N.E. Seattle, Washington 98105

FOREWORD

In many social policy areas the experiences of the past few years have raised serious doubts concerning the effectiveness of both present public programs and proposed ideas for new policies. It is clear that all levels of governments badly need more and better evaluative information as a guide to improving present public programs and devising new policies with a reasonable probability of producing desirable results. Unfortunately it is also becoming clear that various social science institutions including those located at universities are ill-equipped to carry out the badly needed evaluative research.

This paper investigates the capacity of social science organizations to develop a high level of large-scale evaluative studies in support of public decisionmaking. The issues are critical for both governments and universities. Governments must face the problem of how to stimulate more evaluative research in the social areas. Universities with large numbers of social scientists on their staffs must decide <u>if</u> and <u>how</u> they should participate in these studies. Neither task will be an easy one.

A major objective of the Institute of Governmental Research is to encourage analyses that will lead to the improvement of public policy and to disseminate the results of such investigations. It is in keeping with that objective that this paper is being published by the Institute.

The author is Professor of Public Affairs in the Graduate School of Public Affairs and Director of Research for the Institute of Governmental Research at the University of Washington. Dr. Williams previously was on the staff of the Office of Research, Plans, Programs and Evaluation, Office of Economic Opportunity. He brings rich experiences from both government and university to the complex problems of developing evaluative research for social policy purposes.

<u>The Capacity of Social Science Organizations to Perform Large-Scale Evaluative Research</u> was prepared by Dr. Williams to be presented at the Cost-Effectiveness Conference, sponsored by the International Federation of Operations Research Societies, in Washington, D.C., April 12-15, 1971.

Robert H. Pealy
Director
Institute of Governmental Research

CONTENTS

ACKNOWLEDGMENTS

Several individuals have provided either useful comments on earlier drafts of the paper or discussion on specific points of issue: Brewster Denny, University of Washington; Robert Levine, RAND; Richard Nelson, Yale University; and Peter Rossi, Johns Hopkins University. In addition, I am indebted to the Executive Committee and to Henry David and Stephen Baratz of the Division of Behavioral Sciences, National Academy of Sciences/National Research Council. During the period I was preparing this paper, I was serving as a consultant on federal evaluation policy for the Executive Committee. The latter activity permitted me to discuss many of the issues in this paper with the Executive Committee and staff of the Division, and contributed to my understanding of them. The views expressed in the paper, however, are those of the author, and not necessarily those of the individuals and institutions mentioned above.

The paper was improved by editorial criticism from Lucille Fuller of the Institute of Governmental Research; and Joyce Creighton of the Institute staff aided me by typing several drafts of the paper.

I. INTRODUCTION

The principal concern of this paper is with the capacity of social organizations to perform large-scale evaluative research in support of social policymaking. This type of research presents a number of problems for social science organizations. First, the research generally will need to be multidisciplinary in nature often drawing on both the social and the biological sciences. Second, evaluative studies of major social policies will be complex undertakings requiring a high level of technical, organizational, and administrative skills. Third, the research often will have to conform to a tight time schedule to be useful in policy formulation. Finally, the results if directly relevant to major policy decisions, may involve the organization in heated political controversy.

Evaluative research includes two types of studies: outcome evaluations and field experiments. The former are studies intended to measure the effects of an agency's existing projects or programs on their direct participants, other designated groups, and/or specific institutions (e.g., what is the relationship between benefits and costs). Field experiments are designed to assess the merits of new policy ideas in terms of outcomes in a setting corresponding at least in part to actual field operating conditions.

Few would argue with the normative proposition that a major evaluative effort in the social areas is badly needed, particularly for those programs serving disadvantaged groups.[1] The experience of the 1960's had brought serious doubts about the effectiveness of most major social programs and a lessening of the confidence in what seemed to be an implicit premise of the early War on Poverty years that programs could be launched full scale without testing and yield significant improvements in the lives of the disadvantaged. Bits and pieces of evidence (some from research) hard thinking, and good will did not necessarily combine to produce programs leading to dramatic breakthroughs in the social areas. Over a wide range of social programs, it has been found how difficult and/or expensive remedies for poverty are, and how little is really known about workable techniques for helping the disadvantaged.

Yet few sound evaluations and even fewer rigorous small scale projects have been undertaken by the social agencies serving the disadvantaged. And a corresponding disquiet has set in concerning the existing capacity in the social science research community for large-scale evaluative research in support of social policymaking. In light of the the short supply of competent policy-oriented researchers, the inadequacy of methods and concepts for carrying out evaluative research, and the lack of organizational capacity to undertake large-scale evaluative activities, one can ask legitimately whether or not evaluative results used directly in the social policy process may not cause more

[1]The subsequent discussion concerning the methodological capacity to assess project or program outcomes is probably just as pertinent for all social programs as it is for programs aimed at reducing poverty and/or the barriers to equal opportunity. However, my experience is with the latter so that the remarks in the following pages technically refer to social programs and policies for the disadvantaged. For a fuller discussion of my experience, see section III. B. of this paper.

harm than good. Thus the basic issues of the paper concern:

1. The organizational changes that might be made within the government and the social science community to increase the number of capable policy-oriented researchers, to foster improvements in evaluative concepts and techniques, and to develop more social science organizations capable of carrying out sound, large-scale evaluative investigations and other kinds of research needed for policy formulation; and

2. The possible deleterious consequences for society and for social scientists in performing studies directly relevant to social agency policy, and the measures that might be taken to reduce these dangers.

A number of factors relevant to the consideration of the two questions set out in the previous paragraph will be discussed. It is important to observe at the outset that we will find far more that is unknown than is known. Certainly no definitive answers will be forthcoming -- rather at times the gaps in our knowledge will present barriers even to intelligent discussion. Be that as it may, the questions are too important to ignore. At basic issue is whether or not social science can make a significant contribution to social policymaking while still maintaining its historical role as a relatively independent critic of public policy.

II. FACTORS RELEVANT TO A CONSIDERATION OF SOCIAL SCIENCE'S CONTRIBUTION TO SOCIAL POLICY

The purpose of this section is to discuss in a broad context a number of factors relevant to a consideration of the issue of social science's contribution to social policy. Various factors -- many of which would require at a minimum a lengthy paper for a reasonable treatment -- are discussed briefly to suggest the nature of the topic. These factors include:

1. The social policy areas present complexities beyond those that exist elsewhere.

Rats, pigeons, missiles, genes, and even elementary particles appear to be far easier to understand in a rigorous way than people in complex social situations. At the heart of the matter is the fact that a practical infinity of possible relevant variables exist; all are likely to be related to one another; and it is very difficult to guess on an a priori basis which are going to be most pertinent to any given problem. This dilemma must be faced both by researchers in treating intercorrelation and interaction among variables, and by persons responsible for developing programs. For example, consider the area of education for the disadvantaged. First, we do not understand the process of education and the determinants of educational achievement. Second, available evidence indicates that the relevant factors are not limited to those the school itself may be able to control such as classroom techniques, teachers, school budget, and school organization (and these are complex enough), but include such factors as socioeconomic status, race, community (or neighborhood), and peer group associations. And, of course, policymakers have little or no control over these variables.

The fact that social programs must operate in a complex political and bureaucratic setting complicates matters even more. For example, Title I of the Elementary and Secondary Education Act in deference to the States' prerogatives in the educational field simply made money available for helping the disadvantaged without any very strong

stipulations as to how it would be used. The money dribbled into the educational system and mixed with other monies raises real questions about its effective use, and at the same time makes it almost impossible to measure that effectiveness in any kind of rigorous evaluative study. If we add the politics of the Congress, the agencies, and the Executive Office, problems both for the program designer and the researcher almost boggle the mind.

One hardly needs to dwell at length on this complexity in the social program areas as it is a well known phenomenon. Yet it is well to keep in mind that what social science must be organized to do is to address probably the most difficult of all areas of study, that of people interacting with each other in a large complex society.

2. Large-scale, multidisciplinary studies will be required to fulfill the data needs of social policy-makers; and the need for major outcome evaluations and field experiments that necessitate replication seem certain to usher in an era of "big social science."

The beginning signs of "big social science" -- research requiring relatively large amounts of money and a high level of organizational and administrative skills to operate the project -- are already apparent. The OEO/Department of Labor outcome evaluation of five manpower programs has over 10,000 people in the treatment and control groups and is estimated to cost 4-1/2 million dollars; OEO's performance contracting experiment is estimated to cost 6-1/2 million dollars;[2] and the cost of

[2]In a performance contract a local school system will enter into an agreement with a private company to provide classroom instruction. Payments to the contractor will be based on measured rates of classroom achievement (e.g., changes in reading levels over a contract year) with bonuses for high overall classroom achievement and penalties for poor performance.

the several negative income tax experiments now in progress will run into the tens of millions of dollars. While it is a central question whether or not social science in either a methodological/conceptual or organizational sense is capable at present of undertaking these large-scale studies, there is little doubt that such studies will ultimately be needed for intelligent social policymaking.

> 3. The social sciences have been characterized by a lack of orientation toward and organization for large-scale, multidisciplinary policy research.

The reward structure of the social sciences in the past militated against policy research in that policy-oriented researchers in the social sciences, except economics (and here the policy orientation has been in traditional areas such as monetary and fiscal policy, not social policy), were viewed as second class citizens by their disciplinary peers. This point is made strikingly in the recent National Academy of Sciences/Social Science Research Council report on the social and behavioral sciences which stated:

> Although there is a close relationship, in principle, between basic research and applied and developmental work, basic research tends to receive more attention from behavioral scientists in universities. Many academic scientists value the prestige that their contributions to basic research and theory give them in the eyes of their peers more than whatever rewards might be obtained from clients who would find their work useful. It is no wonder that university scientists prefer the kind of research that is satisfying in itself (because it is self-initiated and free of restraints) and leads not only to scientific knowledge, but also to respect and status tendered by those whose judgments they value most. It is no wonder, either, that their value systems are passed on to their students. Thus, much of the applied work in disciplinary departments is done by those who for one reason or another do not compete for

> the highest prizes of their disciplines. [9, p. 193.][3]

Moreover, there is little tradition in the social sciences of extended multidisciplinary work. Yet social problems in general cut across the established disciplines. For example, an effort to investigate means of increasing the capacity of the public schools to educate minority children may well require research not only by sociologists, psychologists, economists, and linguists but also biological scientists in areas such as nutrition and brain functioning. Collaboration, however, by members of different social science disciplines -- much less, joint research with biological scientists -- is the rare exception.

Furthermore, the social sciences in general lack organizations with the capacity to perform large-scale field research. Extensive evaluative research requires major organizations with large multidisciplinary staffs. These may in turn need high levels of administrative capability, elaborate divisions of labor and hierarchies of authority and status. Few such organizations exist either in universities or outside of them, and those that do such as the National Opinion Research Center and the Survey Research Center have traditionally specialized in sample survey activities. And these organizations certainly do not have proven records of high level evaluative research in the social areas.

This lack of orientation toward and organization for policy-oriented work have produced a severe shortage of policy-oriented researchers; a

[3] Currently I suspect that this statement is less true than at publication in 1969. But we simply do not have a systematic study of this question (see point 5., page 9).

minimum of graduate preparation aimed at developing competent researchers who view policy as their main area of inquiry; and an inadequate development of concepts (e.g., an explanatory model in the education area), methods, and field procedures for deriving evaluative results.

> 4. Present deficiences in staff size and skills within the government -- the Executive Office, the Congress, and the agencies -- severely limit the level and quality of evaluative activities in the social program areas.

Part of the explanation for the dearth of relevant policy research in the social areas derives from the fact that the government has done little to define clearly what types of studies are needed, or to encourage this socially important and methodologically challenging research in the social science community. Generally, government staffs do not have the technical and administrative capability to determine evaluative needs, design or work with contractors and grantees to design studies, and supervise the ongoing evaluative effort or monitor it in sufficient detail to determine the validity of the results. After studying the status of evaluation in 15 programs at HEW, HUD, OEO, and the Department of Labor, and in the General Accounting Office in the Bureau of the Budget (now, Office of Management and Budget), the Urban Institute group under Joseph Wholey observed: "The Social Security Administration's Office of Research and Statistics and the OEO Office of Planning, Research, and Evaluation, each with a substantial in-house evaluation capacity, are notable bright spots in an otherwise bleak staff picture" [12, p. 82].

It would be a distorted picture to stress only weaknesses. In the last few years, the social agencies, the General Accounting Office, and

the Executive Office have begun to develop evaluative capability. Such studies as the New Jersey negative income tax experiment, the performance contract experiment developed at OEO, and the OEO/Department of Labor longitudinal evaluation of five manpower programs certainly indicate competence. Furthermore, given the difficulties of assessing social programs and projects, the progress-to-date may be quite reasonable in terms of developing evaluative capability. Our purpose, however, is not to judge the past but to look toward the future. And in these terms, it would be a most serious mistake not to recognize that the present deficiencies in staff size and skills within the government severely limit the level of evaluative activities in the social program areas.

> 5. Little detailed information is available concerning the present capacity of the social sciences to engage in policy-oriented research and teaching.

While the present deficiencies noted above are quite apparent, it is also true that those concerned with future directions in the social sciences must operate with only the most limited information. Now available in published form are broad statements about the lack of social science activity in the policy areas and general descriptions of activities such as the number of Ph.Ds graduated each year in the social sciences, but little else. A brief example will illustrate this point.

The 1970 National Register of Scientific and Technical Personnel included in its Specialities List for economists such categories as "Economics of Poverty," "General Welfare Programs," and "Urban Economics and Public Policy." Five years ago, these categories did not exist, yet what do they imply? The American Economics Association which administered the questionnaire asked economists to specify their specialties from a long list that included no more than the name of the category and a number. To paraphrase Humpty Dumpty, "Economics of Poverty means just what I choose it to mean." In short, there is absolutely no qualitative information available, not even a simple description of what research the economist is engaged in, and the lack of information is hardly restricted to economics.

Little effort has been expended to study systematically the social science research community -- no one has researched the researchers. In response to perceived pressures to be more relevant, universities have established new schools, new institutes, and new areas of study. Established research organizations and new ones now claim competence in social areas. Yet one can barely find a listing of all of these entities, much less any detailed discussion of their scope of activities and quality. In addition, almost nothing has been written as to whether or not universities are modifying their curricula so as to make their graduates more oriented towards policy questions. There may be some hopeful signs of efforts to study the social science research process as witnessed by the recent report in Science of a study by Karl Deutsch and others on the conditions that make for research breakthroughs in the social sciences [3]. This type of work, however, is just beginning. And one cannot overemphasize the scant information upon which to develop some notion of the potential supply of policy research for social agencies.

> 6. The milieu for scientific activity has changed significantly in recent times. The general attacks of science both for errors of omission and commission, the turmoil in the universities, and the

> decrease in funds available for research and for university operations seem likely to be important factors, even if we cannot yet predict how they will affect science policy.

While it is far too early to try to assess the implications of the above changes in terms of their potential effects on policy research, it is not obvious that the effects will necessarily be negative. The looseness of the academic market place now may make it much easier to "guide" people toward policy research. As a government staff member in the mid-1960's, I found that in trying to generate policy research one had to beg and make all kinds of concessions to researchers to get them to consider policy issues. Those days may well be gone; the Office of Management and Budget, for example, is now able to push federal granting organizations like the National Science Foundation, and the National Institutes of Health much more toward funding "relevant" research. It is not clear that this is necessarily bad in the development of social science as a contributor to social policy.

> 7. Particularly at the present low stage of development of evaluative techniques in the social sciences, the use of evaluative studies in social policymaking poses risks both to society and to science.

The purpose of evaluation itself may be disruptive for program personnel and participants, and bring conflict in an agency between evaluation staffs and operating bureaus. Scientists (avoiding the mystic abstraction "science") can be completely wrong in terms of design of studies or interpretation of results, ignore or not perceive deleterious consequences of their discoveries, or manipulate data and theory in such a way as to support their political beliefs -- needless to say, with resultant harm. For example, an outcome evaluation indicating incorrectly that a program is not effective can bring reductions in program funding, unwarranted changes in staff and policies, and a shattering of the morale of staff and participants. Or, the argument that every new operating program ought to be tested and shown to be effective before operating on a large scale can be used as a facade for disparaging all new ideas and retrenching on social commitments. Legislators and administrators relatively ignorant of evaluative techniques may overvalue and hence over-react to quantitative data because of their aura of scientific accuracy.

Conversely, the undertaking of evaluative research may have harmful consequences for individual scientists and the institutions of science. The context in which information is used is very important -- the same data cited in a scholarly journal and on the floor of the legislature have different implications. Thus any information including evaluative data which can have a material effect on policy decisions, such as bringing significant cuts in program funding, is best viewed as "political" information. Political sensitivity can bring the researcher into the center of a raging controversy. This is well illustrated by the Westinghouse Learning Corporation evaluation of Head Start in which the debate over the validity of the results was carried on not only in the scholarly journals, but in major newspapers, the Executive Office, and the Congress. As I have observed elsewhere:

> [Westinghouse is] a stark illustration of what might be termed the implications of the iron law of absolute evaluation flaws. That is, as a general rule the absolute methodological and logistical deficiencies in any evaluation make political infighting a near certainty when evaluation results threaten a popular program. In short, "questionable evaluation practices" can always be attacked on methodological grounds for political and bureaucratic purposes. [14.]

Academic independence and objectivity can be threatened by the institutional relationship between the governmental sponsor and the researcher. Sponsors of evaluative studies may attempt to suppress unfavorable findings. They may tell an investigator what to find or to change results, or force the release of preliminary results in support of a particular policy position. Even without overt influence, a close and continuing relationship between an agency and a research organization may either raise doubts concerning the latter's objectivity or blunt its sensitivity to bad policies of the client. Mutual trust and the asking of embarrassing questions that might put basic programs of the agency in jeopardy are difficult to combine over long periods of time.

III. A COMMON FRAME OF REFERENCE

The above listing of factors -- many meriting a long paper which in most cases could be written only in terms of the grossest speculation because of the lack of firm empirical evidence -- is so formidable as to make sensible discussion difficult and recommendations concerning how to foster policy-oriented social research hazardous. If one does propose to engage in discussion and to proffer recommendations what seems to make sense is to first try to establish a common frame of reference. This involves specifying what proves to be a limited number of points about which most of us would agree, and from this common starting point examining my particular experiences, orientation, and biases.

It is pertinent to indicate the reason my biases are stressed in this discussion. I feel very strongly that almost no intelligent discussion has emerged concerning the critical issue of the potential contribution of social science to social policymaking, and that a major barrier to fruitful discussion has been a tendency to work from unstated assumptions concerning matters about which we have little or no solid evidence. Only if our biases and our fears, such as the possible contamination of science by its closeness to policy, are set forth and considered will we ever progress to reasoned debate.

A. Common Points of Agreement

A strong consensus would emerge in support of the following statements:

1. There are a number of grave social problems for which solutions urgently need to be sought.
2. Results from soundly conceived and executed studies that measure the effectiveness of existing programs and assess the merits of new policy ideas on a small scale before new large-scale programs are launched are urgently needed in support of social policymaking.[4]

[4]This is probably the one statement for which there may be disagreement. But I suspect it is more apparent than real in that those who oppose evaluative studies do so more in terms of perceived weaknesses in techniques (soundly conceived and executed studies cannot now be carried out) than of a rejection in the abstract of the principle of evaluation.

3. In all likelihood major decisions on social policy will not await research results, nor should decisions be held up for the lack of research evidence.
4. Evaluative evidence will be and, even if there were great improvements in its quality, should be only one of several kinds of available information in the policy process, with choices ultimately being political in the broad sense of that term.

These statements need only the briefest elaboration. Major decisions about social policy are now being made and are going to be made in the future with only limited empirical evidence. For example, the decision in any budget year to continue programs as they are presently operated is itself a significant decision. Moreover, even major decisions to change programs or introduce new ones will be made without much evidence. In view of the present deficiencies in techniques, personnel, and organizations in social research, it would be ridiculous to suggest that new social programs cannot be started unless there is strong empirical evidence showing their effectiveness. Moreover, great improvements in the development of research information should not make that evidence overriding. The propositions indicate support for a pluralistic process in which evidence of many types, including political and bureaucratic prerogatives, will be important -- often far more important than sound evaluative data. Still it would be nice to have some of the latter at a time when key social decisions are to be made.

B. The Author's Experiences and Biases

When I was at the Office of Economic Opportunity, I was on the staff of the Office of Research, Plans, Programs and Evaluation (RPP&E), the first central analytical office in a social agency. The office was responsible directly to the agency director for overall planning for programs aimed at reducing poverty and the barriers to equal opportunities, and also supported studies by outside researchers. RPP&E was a key office in the policy process in which decisions about social programs involving billions of dollars were made -- more often than not without benefit of any hard evidence or with very limited data. Decisions were made with these great gaps in knowledge not because RPP&E or OEO, did not try to find relevant, sound information; but because such information did not exist, and often there was not time (given the regimen of the policy process) to develop the needed data. Yet lack of time was not the only limiting factor, our attempts to develop policy research made clear the difficulties of policy studies both in terms of methodological and conceptual shortcomings and bureaucratic/political barriers.

This experience has made me both tolerant of imperfect, though better than present, information and sensitive to the difficulty and the challenge of social policy research. Let me try to crystalize my perspective by setting out and discussing three propositions that reflect how I approach the problem of developing more policy research:

> 1. In some substantive areas (e.g., education and manpower) present evaluation methods involving large-scale sample surveys of programs offer a real potential in the near term for increasing materially the useful outcome information directly pertinent to social policy decisions.

Notwithstanding, there are serious problems in carrying out such work, and the resultant findings will hardly be definitive or substitutes for good judgment. As the author has noted elsewhere:

> [T]here must be a far greater concern for the requirements of statistical design than has generally been exhibited in the past. These requirements would generally include a well-designed sample, early field interviewing to maximize the retrieval of information, repeated follow-up to reduce sample attrition, and a reduction in the importance of heroic assumptions in the model. In short, good evaluations are going to need well-qualified evaluators who are funded at "high" levels so that excessive shortcuts are not required, and who are given realistic planning time to develop a sound evaluation model.
>
> Even under such circumstances, it will be necessary to make arbitrary decisions and to recognize that many crucial questions are beyond our present capabilities . . . .
>
> Further, some caution is needed in interpreting evaluation data which generally will mean fitting the evaluation evidence into a mosaic with other reasonable evidence to "validate" a decision. In general, the present generation of outcome evaluations should be viewed as a piece of evidence, not <u>the</u> definitive piece of information that bowls over all other reasonable indications of a different policy decision. [13, p. 457.]

> 2. While useful evaluations can now be performed with present techniques, they will be time consuming; and in some areas in which concepts and techniques are deficient, only a very long time horizon is realistic.

The time exigencies of program policy needs and the time requirements of sound evaluative studies are certain to be in conflict. Haste is not compatible with the present deficiencies in evaluative methodology and field procedures. Even if one relies essentially on present techniques, it will take significant amounts of effort, time, and money to produce evaluative studies that are significantly better than those of the past.[5] Moreover, the present technology in some areas is sufficiently limited so that study results will not feed directly into the decision process, but will lead in succeeding stages to a decisionmaking input. Such exploratory activity would be expected over time to increase the capacity to produce significantly better future outcome data. But the payoff in terms of results directly relevant for decisionmaking may be many years away.

The idea of a long time horizon for some evaluative activities with an initial emphasis on exploratory work not leading directly to inputs into the decision process will be difficult for key officials to accept. Pressures on them to act quickly are tremendous. It would be a sterile exercise to push aside these political considerations. At the same time, if a realistic attitude toward the time required for evaluative research cannot be developed, it is difficult to see how real progress can be made.

> 3. Evaluative information may be used for policy purposes unless compelling *a priori* reasoning or strong empirical evidence support a claim of potential bias in the results.

[5] For example, this conservative time estimate would hold for an evaluation using an outside contractor to measure prospectively a manpower training program of six months duration: two or three months to get bids and award the contract; two to six months for the contractor to develop the evaluation methodology and sample; six months for the manpower training program itself; six to twelve months of on-the-job time by participants after the training (depending on acceptance of six months or one year wage experience); and two to six months to process the data and prepare a report. This estimate indicates a time range from one and a half to nearly three years from the start of an evaluation until results come in. At the minimum, data would be available not for the upcoming fiscal year but for the one after that.

To date, there has not been a field evaluation of a social action program that could not be faulted legitimately by good methodologists, and we may never see one. Only cursory inspection of an evaluative study will generally produce a number of potential biases, e.g., people will refuse to be interviewed, members of the treatment and control groups will disappear, or project directors will not allow random assignments. With some scholarly thought a fantastic number of subtle potential biases can be unearthed. However, policymakers should not reject evaluative results because some bias might possibly exist.

The recent Westinghouse Learning Corporation evaluation of Head Start referred to previously well illustrates this point. In this study, a sample of children who had gone on to the first, second and third grades after participating in the Head Start program were matched in terms of socioeconomic characteristics with a control group of children from the same local area, and both groups were administered a series of tests to ascertain levels of cognitive and affective development.

Surely possible biases in this case may be important. The standard argument in such *ex post facto* investigations is that those who enter the program are more likely to succeed either because they (or in this case, their parents) are more motivated than the controls or because the program personnel "cream" from the applicants picking the best in order to make the program look good.

In a recent paper, Campbell and Erlebacher [2] have argued the reverse of this proposition in suggesting that *ex post facto* evaluations in the social areas may be biased <u>against</u> the treatment showing a positive effect in that (a) program personnel purposely choose the most needy applicants, and (b) available statistical techniques such as those used in the Westinghouse study do not correct sufficiently for the fact that the treatment group is worse off than the controls. If both parts of the argument hold, an evaluative design requiring significantly higher achievement scores as evidence of Head Start effectiveness would be inappropriate -- catching up by the Head Start group might well be an acceptable performance. Given Campbell and Erlebacher's eminence as methodologists, let us simply accept as valid their argument concerning the present inadequacy of techniques to correct biases. However, Campbell and Erlebacher support the first part of the argument only with hypothetical data, not empirical evidence. So, if the alleged bias does not exist, the deficiency of the statistical tools does not matter.

I have my doubts that such biases are important. And these doubts are grounded in observation, if not systematic study. First, the situations in which individuals come to apply for programs are quite "chancy": they may be recruited (at times almost physically dragged in) or get program information almost by pure chance either from a friend or other sources. Second, program assignment procedures, in part, because of the great urgency of getting people into programs within time and fund restrictions, do not leave one with the impression that staff personnel can determine from among a group of individuals with similar objective characteristics (e.g., income, level of education,

job experience, etc.) who is more or less likely to succeed in the program. Thus, in <u>ex post facto</u> program evaluations with reasonably well-matched treatment and control groups drawn from a large number of projects, I would argue the following: The likelihood is small that the results will be sufficiently biased by the lack of random assignment procedures so as to render them misleading for policy analysis.[6]

In sum, a policymaker faces the inherent dilemma of running risks if he chooses to develop evaluative evidence (e.g., wasting funds on useless studies or, worse, using invalid results) and of incurring other risks if he does <u>not</u> undertake evaluative research (continuing ineffective programs or launching new flops). No available information permits a precise and objective weighing of these risks. So we are each reduced to a quite subjective assessment dominated by our own experiences. Mine lead me, on the one hand, toward developing in some program areas evaluations that are expected to be used directly in policymaking. "Fairly good" evaluative results available before decisions must be made will be both a distinct improvement over the past and superior to more rigorous results arriving after the decision. One may well trade quality for

[6]This formulation is hardly very elegant and depends heavily upon the matching of treatment-control groups. Some small but consistent differences may exist. There is no rule for saying when the difference is large enough to bias the results for policy. Moreover, in small geographic areas, a program such as Head Start could present problems. To the extent that the poorest people are relatively well known in an area and the program is able to "saturate" that group by getting almost all of the eligible children, the controls may be superior in socioeconomic terms. But, for the great bulk of programs in which the available slots will be far below the need level, saturation seems more unlikely to occur. In general, saturation requires far greater knowledge than present about the population, more sensitive selection techniques, more time available to those who determine participation than generally exists in the real world, and far more money.

speed. For example, in the just discussed Head Start evaluation, RPP&E chose an ex post facto study that would yield evaluation data in a year over a more sophisticated longitudinal design taking at least three years. On the other hand, in some areas I would restrict work to developmental studies, and in the longer run would require in all areas quite high standards for policy research. In short, it would be mixed strategy reflecting different needs and different time frames.

IV. WHERE DO WE GO FROM HERE?

The overriding issue in this paper concerns the steps over time that the government (the consumer of policy research) and the social science research community (the main potential supplier) should take to increase materially the development of soundly conceived and executed evaluative studies, and to reduce the dangers attendant with such development. The following sections will address possible future directions (some of which are clear but most of which are not) and some of the problems they raise. The discussion has three major topics: the federal government's demand function for policy research, the organization of the social sciences for policy research, and the means of minimizing the risks of developing and using evaluative results.

A. The Federal Government's Demand Function for Policy Research[7]

One thing is most clear -- a significant increase in soundly conceived

[7]The concern of this paper is restricted to the government's technical not its political capability to develop and use evaluative research. So this section will not address such questions as the location and status of the social agency evaluation office; the relationships among the social agencies, the Executive Office, and the Congress concerning evaluative studies; etc. For such a discussion, see [14].

and executed evaluative studies requires that social agencies as the primary developers and users of social program evaluative results within the government establish large, well-trained staffs with sufficeint technical and administrative skills to determine evaluative needs, to articulate these needs to outside researchers, to design or work with contractors and grantees to design studies and methodologies, and to supervise the ongoing evaluative effort. The skills and knowledge required for competent evaluation staff members are quite high: substantive knowledge about specialized areas (e.g., education) including the ability to specify evaluative needs; a sound background in designing evaluative studies and using statistical techniques including the ability to translate variables into measurable concepts usable in the field; and beyond these two sets of technical skills, the administrative ability to work with program personnel and researchers over time. Despite the fact that such skills in the social program areas are in short supply, the number of people needed for a viable evaluation staff will generally be substantial. For example, Wholey estimates that at least a GS-13 to GS-15 level staff member is required for every two to four (or $500,000 worth of) outside studies, with additional staff needed for special functions such as developing overall evaluative needs. [12, pp. 82-85].

It cannot be overemphasized that the government policy research staff must be made up of people with sufficient technical training and/or experience to interact with academic social scientists in a peer relationship. While the social scientist may have a national reputation in his discipline (an unlikely status for the government staff person), the latter should be at least a qualified journeyman in his academic specialty. One party may have a comparative advantage in terms of techniques and disciplinary knowledge, and the other in terms of knowledge about policy and policy needs, but they must speak the same disciplinary language.

The Congress and the Executive Office also badly need evaluative data, and at a minimum must have the staff capability to articulate their evaluative concerns to the agencies and to be intelligent interpreters and users of evaluative information. Beyond this, both may wish to develop sufficient staff to carry out a limited number of evaluations to keep the agencies honest. The limiting factor -- and it well may be an overriding one -- is the shortage of competent evaluators. The present situation is analogous to the one existing at the start of the Planning, Programming, Budgeting System government-wide in 1965 where a severe shortage of policy analysts thwarted the implementation of a basically sound concept for improving the governmental decision process. Thus, everyone attempting to draw on a limited supply of competent people may result in no staffs of sufficient size and skill to carry on a high level of evaluative activity.

If the federal government is to increase significantly the flow of sound evaluative results in the social areas, it must fund relatively more research directed specifically toward major policy problems and require that the research involve more interaction between government policy research staffs and outside researchers than it has in the past. It is important to stress that the statement does not imply only "applied" work but may also include "basic" research. However, the statement is meant to convey the notion of some structuring of the research both in terms of the

areas of concern and of the interaction with government staffs. For example, the adequate education of lower socioeconomic class, minority children is clearly a major social problem that badly needs investigation. In the search for causes of poor education and means of improving it, research might range from studies of the possible relationship between malnutrition and brain damage in a fetus to field experiments testing a new teaching process. Not only may policy-oriented research include "basic" research, it seems likely that in many social areas, major new applications must await the development of new knowledge from fundamental research.

At the same time, policy research should be structured at least to the extent that government policy research staffs will specify gaps in knowledge blocking more intelligent policymaking; and that the researcher will be committed to thinking about these needs. This commitment, including interaction with a government policy research staff, is extremely important. The need for interaction rests on the premise that in part the lack of useful social policy research in the past stemmed from ignorance about programs and policies, policy needs, and the form in which research results would prove useful in the policy process. What the government policy research staff should offer relative to the outside scientist is not superior intellect, but superior information about policy needs. This argument holds even for the most fundamental policy studies by social scientists in which the researcher has great freedom to determine the scope, character, and timing of the study; here also the value of the study for policy purposes is likely to be greater when the researcher has an appreciation of policy needs gained through interaction with government staff members.

The above formulation may raise in the mind of the social scientist a specter of government staff members with a new shibboleth, "policy relevance," dictating what scholars need to study and how it should be studied. Such a danger always exists. But if government policy research staffs are upgraded as envisioned, the interaction may for the first time provide researchers with knowledge about policy needs in sufficient detail to permit fruitful policy work.

In this regard, one other point needs to be made. The past has been one of significant interaction between the government and social scientists. But this interaction has been between researchers and government research -- not policy research -- staffs which themselves were not informed about or oriented toward program and policy concerns. As the author has observed elsewhere:

> The critical point is that the government men of power, themselves, have not been oriented toward policy process needs. That is, the government research bureaucracy, as isolated from policy as the academic, frequently shares the academic gestalt of what is proper (pure exceeds applied) but in their wisdom define popular pure areas. Thus the long tradition of government funding has supported the academic's distaste for policy-relevant studies while allowing him to believe that he is aware of what social research government does in fact need. [14.]

It is critical to recognize, however, given both the weaknesses of government evaluative staffs and the social science community to perform policy studies, that the shift toward policy-directed research should be relatively small and gradual. The great bulk of social research should

continue to be guided by the same concerns as in the past, with scientists performing research that in the long run may facilitate policy but which is not framed with policy concerns in mind. At the present stage of groping to formulate what research is needed for policy, it would be ridiculous to try to shift a large percentage of research funds toward explicit policy questions. The government simply does not have the technical capacity to use vast sums in search of policy relevance. Lots of money can be as destructive as small amounts -- agency research managers, just as program operators, must obligate all funds before the end of the fiscal year or lose them. The time that might be concentrated on developing a few sound projects with a high potential for producing policy results may instead be widely distributed over many projects of dubious quality and relevance in order to expend all funds.

For this reason, it is not appropriate to try to specify desirable absolute or relative levels of expenditure increases for policy studies or the time path of such increases. What is appropriate is a strong recommendation for a rapid build-up of policy research staffs. Evidence of staff capability must precede major funding increases. Also a firm commitment must be made to more policy-oriented research particularly evaluative studies, that is "validated" by immediate (but still relatively small) funding increases for policy studies at the expense of other research.

B. The Organization of the Social Sciences for Policy Research

The types and level of research on social problems now required to facilitate social policymaking strongly indicates the need for more special organizations (e.g., profit and non-profit research organizations, institutes, academic departments or schools) with explicit missions of large-scale, multidisciplinary research and/or teaching in the social policy areas. Recently the establishment of large-scale social policy organizations has been recommended by a number of groups and individuals, including major committees of the National Science Foundation, the National Academy of Sciences and the Social Science Research Council [5,9]. NSF which proposes the creation of "Social Policy Research Institutes" and the combined NAS/SSRC committees which in turn propose new "Graduate Schools of Applied Behavioral Science" develop their recommendations along quite similar lines. But there are differences. The NAS/SSRC proposal suggests that the new organizations be a part of the university and have a regular teaching function. The NSF report leaves the location issue open, neither requiring that the institutes be at universities nor rejecting that location but stressing a close relationship with the agencies (not a point of emphasis in the other report).

It is the question of location -- in or outside of universities and where in the university -- that seems to be the most controversial one. There does seem to be a convergence of views on one point about location. Both the NSF and NAS/SSRC reports state explicitly and strongly that the new organizations when on campuses should be financially and administratively <u>independent</u> of the established social science departments. And one can certainly make a strong normative argument that disciplinary departments ought not as departments undertake large-scale policy research which may draw researchers from teaching, and present even stronger positive

evidence that individual social science departments are not likely to undertake major policy studies, particularly those of a multidisciplinary nature.

It does seem a valid argument that a social policy institute, department or school should have the same financial and administrative power to choose its professional staff as the disciplinary departments. For example, in its appointments, a social policy institute should not be beholden to a disciplinary department that defines staff acceptability only in terms of theoretical or conceptual elegance validated primarily by publication in a handful of prestige journals. It is also important that social policy organizations not be isolated from the rest of the university, cut off from interaction with disciplinary peers or having lower quality standards for staff. Social policy organizations need not exist in a hostile relationship with the disciplinary departments. An institute in order to have a critical mass for research purposes may hire more people in a special area than a department would to fulfill its teaching functions, and the latter may be happy to make joint appointments (especially if the salary costs are low). A university at the inception or rapid build up of a social policy organization may want to provide special "quality checks" through a multidisciplinary university committee, and to place a high value on cooperation with disciplinary departments including joint appointments. What the university should not do, however, is recreate the disciplinary departments (especially ones that consider policy research an inferior good) by giving them a unilateral veto over the hiring of members of their discipline in policy organizations.

Difficulties of performing large-scale policy research in an academic setting and the potential conflict with other univerity functions have led some to recommend that policy research organizations be separate from the universities. This thesis has been particularly strong in physics where there is a wealth of experience with major applied and basic research centers some of which have been a regular part of the university, some with peripheral attachments to it, and others with no formal relationships. Alvin Weinberg has observed that "the ecology of the discipline-oriented university encourages the rise of purism and specialization and the denial of scholarship and application in science" [11, p. 176]; and Harvey Brooks has suggested: "Where a more programmed effort is desirable or the social need is so urgent that some technical effort is required, even though no very promising new approaches are evident, it should probably be centered at nonacademic institutions, with academic participation only when interest or ideas appears spontaneously from the academic community" [1, p.70]. Edward Teller has gone so far as to recommend that even without the collaboration of established universities, applied laboratories should be given the responsibility for the education of applied scientists [7, p.126].

There has been much less discussion of the location of social policy research organizations. However, Peter Rossi [6] in a provocative paper makes a strong case against locating at universities those types of social research organizations which have elaborate divisions of labor, distinct hierarchies of authority and status, and a professional staff drawn from several academic disciplines. Rossi labels such organizations

"research firms," and offers as prime examples the Institute for Social Research at the University of Michigan, the National Opinion Research Center at the University of Chicago (which he directed for several years), and the Bureau of Applied Social Research at Columbia University. He thinks that such organizations do not fit well in the university structure.

The basic argument is that having a regular academic appointment is more prestigeful than a nonacademic research status and offers more autonomy than mission-oriented work. Large-scale, social research organizations under these circumstances tend to attract marginal people. Some who are very good, but for a number of reasons cannot "qualify" for academic appointments, and more who are second rate. Further, staff members generally have all sorts of status conflict problems on the campus. Rossi suggests that the few successes of large-scale, social research organizations based at universities are explained primarily by a charismatic leader whose departure signed their demise. Institutional devices, independent of the strong leader, do not seem to protect the organizations in the hostile environment of the university.

Rossi then proposes that the appropriate research organizations for universities are ones "designed primarily to maximize foreign relations gains from organizing research . . . . [and which] are essentially collections of faculty members each pursuing his own research interests and using the center to provide letterheads, secretarial services, and political leverage within the university and with funding agencies" [6,p.14]. In short the organization that will function well in the university is one that is unlikely to do large-scale social research and because of its ties to one department is almost certain to be unable to do multidisciplinary work.

That all large-scale, multidisciplinary policy research or that the education of "applied" scientists should not be university functions are debatable points even in the physical sciences. In the social sciences with far more limited experience and debate, it would be premature to suggest a divorcing of all large-scale social research and the teaching of policy techniques from the universities.

A distinction, however, can be made as to types of research that may lead to a significant difference in functions between universities and non-university research organizations. For purposes of analysis, research can be distinguished as between studies in which the results are expected to have a direct bearing upon major agency policy decisions (e.g., an outcome evaluation of Head Start) and those in which the results are expected to have an impact on decisions only after additional research or testing (e.g., a tightly controlled laboratory experiment in early childhood learning).

As a general rule, "direct decision" and "earlier state" studies bring a differing set of demands upon the entity conducting the project. First, the former frequently necessitate a massing of staff including specialists both in substantive areas and administration, a number of whom may work full time or nearly full time on the project in order to meet agency time deadlines. Second, direct decision studies may require methodological short cuts that do not diminish greatly the results in terms of policy relevance but that render the study unfit for a prestige

journal thus lessening its academic worth. Third, direct decision studies have a high potential for bringing the institution conducting the study into a conflict situation within the agency decisionmaking process, a point well illustrated by the Westinghouse Head Start evaluation. The closer to a key decision, the closer to potential conflict is a good generalization -- a controlled laboratory study of learning will no doubt be done in relative quiet compared to an outcome evaluation.

As compared to universities, non-university research organizations such as the RAND Corporation generally seem better able in an institutional sense to perform large-scale research, the results of which are expected to have a direct effect on social agency decisions.[8] At least over the near term, say five years, non-university policy research organizations should be more capable of rewarding social policy work both in money and status terms, institutionalizing the "heat" from direct decision studies as a part of doing business, and massing the key substantive area, administrative and field procedure experts needed to mount a concerned effort. These organizations frequently face the problem of finding top flight scientists, but here the selective use of members of regular university departments will often supplement the operation.

These non-university organizations seem the likely candidates for expanding the supply of direct decision studies both significantly and relatively quickly. Nor should this comparative advantage rule out

[8]Institutes located on campuses but having professional staffs whose appointments are not in the established disciplinary departments may be an exception to this statement. In the following general discussion, however, the special and complex issue of such institutes will not be pursued.

more fundamental research since direct decision studies may give the organizations great insight into more basic problems. Further, they may be able to attract competent researchers for any kind of policy studies only if some basic work can be performed.

The universities, however, probably should have the major role in the more fundamental policy-oriented research. Here the articulation of needs and a "good selling job" by the federal government become paramount as many basic social problems fit well in the reward structure of social and behavioral scientists. The key point is that these more basic studies offer the traditional incentives of the past and, as an added attraction, may help overcome national problems to which science itself has oftentimes been a major contributor.

At the same time, the past experience with universities does suggest that funding agencies take a much firmer stance in requiring that university organizations demonstrate policy research <u>commitment</u> and <u>competence</u>. This "suspicion" should, of course, extend to non-university organizations, but historically it has been the universities that have performed most of the social science studies. And funding agencies should be far less willing than in the past to take for granted either university capability or desire to undertake policy research.

One final point needs to be made. The arguments that non-university organizations in the near term may have a comparative advantage in direct decision studies and that funding agencies be less willing to take university capacity for granted should not be expanded to allow the university to "cop out" on social policy research and teaching. <u>The</u>

<u>university presently has a major share of the qualified social science researchers and a virtual monopoly on the education of the social scientist at a critical career stage</u>. Even if one adopts the thesis advanced in a recent report by the American Academy of Arts and Sciences that research is appropriate at a university only when it facilitates the primary university mission of learning [8, pp.6-7], academics are going to have to get more knowledge than they presently have about policy research in order to train competent social policy researchers. And I would rank this training for policy research a highest priority item. It seems to me also university social scientists ought to engage in the search for means of solving critical social problems; hence social policy research qua research is an appropriate university function. But such considerations quickly carry one to a discussion of the university and society -- a topic that is difficult to avoid when focusing on the research function but one that is hardly to be treated in a couple of paragraphs at this time. Rather the basic point is that the arguments of this discussion are not meant to provide -- nor, I believe, <u>do</u> provide -- a basis for letting the university turn away from policy-oriented research and teaching.

C. <u>Minimizing the Risks of Developing and Using Evaluative Results</u>

Wide and careful scrutiny of evaluative activities both by various parties <u>at interest</u> (political) and by relatively disinterested researchers (technical) <u>before decisions are made</u> seems the most likely means of minimizing the risks that invalid evaluative results will be used in policy or that sound results will be misused through interpretations of them beyond their legitimate limits. Two comments need to be made concerning this formulation. First, analysis must occur before decisions are made for the obvious reason that after the decision is often too late. What is not obvious is how to get results on the table, given both the real time pressures of fiscal year decisionmaking and the desires of decisionmakers for flexibility. Second, wide discussion and debate will often leave the proper policy choice still debatable. In many ways the situation will resemble a courtroom setting in which each side has experts who score points with the final verdict resting on contradictory evidence. But better that the validity of the evaluative study design and the interpretation of results be subjected to wide political and technical scrutiny than looked at only in the comparative isolation of an agency or congressional committee.

Agencies supporting research and the researchers themselves are unlikely to want outsiders looking over their shoulders. Steps must be taken to facilitate and institutionalize access to evaluative information at a reasonably early stage. For example, a bill introduced by Senator Abraham Ribicoff would authorize the present General Accounting Office (reconstituted and renamed) to analyze ongoing evaluative studies going so far as to grant the (undesirable) power to subpoena the records of evaluation contractors, subcontractors and grantees. Other possible measures might include 1) a legal requirement for public disclosure at the letting of evaluative contracts and grants not scattered amid a deluge of other announcements but in a single, readily available source, 2) the requirement that contractors prepare for public distribution

interim progress reports, including methodological and procedural discussions, and 3) the establishment of independent bodies (perhaps funded by private foundations or quasi-public entities such as the National Academy of Sciences) to perform thorough methodological critiques. I hold no particular brief for these suggestions. In fact the critical point is that to date little thought has been given to the detailed procedures involved in establishing an institutional structure that will bring evaluative activity under careful scrutiny. At this point the most reasonable proposal is one calling for a search for such institutional means rather than one specifying particular approaches.

Scrutiny is not without its own dangers as a couple of examples will attest. First, in the New Jersey negative income tax experiment, the news media have made several attempts to interview experiment participants. The threat of "contamination" of the experiment must be apparent, and with a relatively small number of participants, the whole study could be destroyed. Another quite real danger is that almost any type of scrutiny (and consider the even more pervasive power to subpoena) can infringe on the rights of participants. Second, in a search for new alternatives in the education area, OEO has proposed experiments with performance contracting and educational vouchers. The former is now in progress, but widespread opposition by the educational community has kept the voucher experiments from starting. It is not necessary to judge the merits of this specific case to see the general dangers deriving from the fact that a concerted effort by an interest group can block an experiment and in so doing effectively block future policy. Yet, it is difficult to see how evaluative evidence itself can be evaluated unless it is made widely available just as other research. And with the potential effect on major decisions, it is imperative that results be subjected to a critique before decisions are made.

There is also an urgent need to investigate institutional means for protecting researchers in policy-oriented studies from government interference that lessens independence and objectivity or unduly restricts the scope of the investigations. In policy studies the federal government should draw on well-regarded (usually academically-oriented) researchers including scholars recognized as outstanding by their disciplinary peers. It should be apparent that with ample funds the government can always get plenty of second rate research. What is difficult is to bring to policy research top flight people who in general will and should adhere to the academic standards of independence and objectivity.

The above formulation is not meant to suggest a hands off policy by government staffs. Studies expected to have a direct impact on decisions will often constrain the researcher's effort -- and legitimately -- in terms of firm time deadlines, the relatively detailed specification of the objectives of the study and methods and procedures to be used, and detailed monitoring of the ongoing work by the agency policy research staff. However, unwarranted restrictions involving attempts to influence the findings, to suppress them, or to force early release of information must be minimized.

Pressures for early release are probably the most likely type of government interference, and have occurred recently in two politically sensitive evaluative studies: the Westinghouse Head Start evaluation

and the New Jersey negative income tax experiment. The White House first ordered OEO to sent it a copy of the earliest draft of the Westinghouse report and then made these very preliminary findings known in a Presidential message; and second, forced Westinghouse to make available to the Congress copies of a preliminary report (for a detailed discussion see [14, Chapter 7]). In the negative tax experiment, the researchers were asked by the Nixon Administration to analyze preliminary data so results could be provided to the House Ways and Means Committee then considering the Family Assistance Plan. The demand came sufficiently early in the experiment to raise legitimate questions about the wisdom and propriety of the analysis at that time. Further, the request had such severe time constraints that it did not permit a reasonable level of quality control in the analysis (see [10, pp. 8, 12]). Can anyone doubt that evaluative results are a form of political evidence?

There is little precedent in the social sciences for the problem of releasing preliminary results. The issue is first a methodological rather than a moral question. A researcher may believe for a number of technical reasons that his preliminary findings are not yet amenable to any kind of interpretation. For example, in the Westinghouse evaluation, the contractor when forced to release a preliminary report argued that additional statistical analyses were needed before interpretation of the data would be warranted. Individual experts may differ in their appraisal of whether or not it would be premature to release information; and, the question ultimately is judgmental in that a methodological purist might well claim that no preliminary release of results is justified. The issue then is what the researcher should do if he believes that his data should not be released in their present form.

To see the nature of this issue, let us assume that an agency is shown some preliminary findings in the normal course of agency-researcher relationships and sees the potential usefulness in an upcoming political situation. The agency might suggest that the researcher let it have the preliminary results including various caveats about their use, and that it will take responsibility for the release and interpretation of the information. In such a release the preliminary nature of the results can be played down so as to suggest that the findings strongly support the desired political action.

Should the researcher refuse to provide the preliminary information to the agency? If the agency gives an interpretation that is unwarranted in the view of the researcher, should he warn the public of the questionable interpretation? The situation is quite different from that of a final, readily available evaluative study that presents a sufficiently detailed account of methodology and concepts and enough data for a reader to assess the study's conclusions and draw his own. Under these circumstances the evaluative results are a part of the public domain, the researcher has no general responsibility for unwarranted interpretations of the published results. While it is a debatable point, there even seems to be no strong obligations to correct interpretations by the contracting agency, as long as their validity can be assessed from information in the final report. But with preliminary results, the researcher alone among nongovernment scientists knows the facts. Silence concerning

unwarranted interpretations means acceptance. Most certainly, if policy is being made on the basis of the preliminary results, corrections will not suffice in a final report. Yet if the researcher refuses to provide requested data to the agency or warns publically of the agency's interpretation, he may jeopardize both the single project and future funding from a (perhaps, the) major source of his organization's revenue.

Even more difficult to treat than overt pressure may be the subtle threat to research organizations dependent for the bulk of their funding on a single or a small number of government sources. Here the potential influence may be a fear of losing future contracts that leads the research organization to try to "please" the big client. Of course, it is tempting (and not completely unjustified) to observe that those who want chastity should avoid compromising situations. The problem, however, is not simply that of protecting the objectivity of research organizations (and hence letting the imprudent ones suffer the consequences), but of protecting the public against the unwarranted use of evaluative results. When a single evaluative activity can, at the extreme, influence decisions involving billions of dollars and millions of people -- to be specific, the New Jersey negative income tax experiment and the Family Assistance Plan -- it is no small problem.

The critical issue is whether or not policy research organizations funded by one or a small number of mission agencies can have extended contact with a funding agency and yet retain objectivity and independence. Enke has argued that in the case of the RAND Corporation and the United States Air Force extended interaction and independence have been blended so as to produce a high level of useful policy work, but that a number of unusual features were required:

> Project RAND's contract includes several little known peculiarities that together go far to explain the extraordinary success of RAND.
>
> First, the terms of reference were and are extremely broad, initially being "intercontinental warfare other than surface." Within this scope, RAND researchers worked on what they and their supervisors considered important, and not necessarily on what concerned Air Force officers. Many a general in the early days, learning of the existence of an Air Force-supported RAND, visited Santa Monica to announce what research he wanted done only to depart perplexed, having been told politely that RAND was deciding for itself what Air Force problems were important and tractable.
>
> Second, and a corollary, RAND decides what completed research to show the USAF. A Project RAND study does not have to be exposed to the customer until the management considers this desirable. Thus RAND imposes its own deadlines. But more important, it can "bury" those research ideas that lead nowhere. True research includes dead ends. With Project RAND the management does not have to "play it safe" by assigning staff only to prosaic undertakings certain to yield something that can be given the customer.
>
> Third, again a corollary, specific projects within the general terms of reference are defined by RAND. Recommendations are less likely to be narrow suboptimizations. Future problems are more likely to be anticipated in time for something to be done: for example, RAND was once criticized for conducting a large project on the defense of North America from air attack, until shortly afterwards when the Soviets detonated their first atomic device.
>
> Fourth, Project RAND has had relatively few financial worries. The initial funding was sufficient for several years. After some difficulties, Project RAND now enjoys something akin to institutional funding as it is given adequate notice of any reduction in level of support. Obviously this is important when recruiting. Also, because RAND does not have to sell research by the project, showing salary charges and rates for each, the management is more able to bid for real talent.

> Fifth, although many of RAND's departments are organized by academic discipline, the major projects are interdisciplinary and include members of different departments. Having discipline departments attracts and holds analysts with superior professional qualifications. It takes a mathematician to recruit a mathematician. But more important is the bringing together of economists, mathematicians, and engineers into a single project for 6 to 12 months, having a unique purpose, and forcing them to learn each others' languages and concepts. After initial misunderstandings, the final results are nearly always significant, often surprisingly so. Too many important and unresolved issues seem to remain unattacked in the No Man's Land that lies between established academic disciplines. At a university, it is only the maverick faculty member who strays into strange and suspect areas, and he alone can seldom achieve as much as can an interdisciplinary team.
>
> Sixth, RAND's staff is usually considered by Air Force officers to be part of the Air Force "family." This was especially true when Project RAND comprised almost all the activity of The RAND Corporation. Hence RAND staff members had access to sensitive information sometimes of the "skeleton-in-the-closet" variety. This close-yet-independent relationship contributes to productive research. It is another reason why certain federal agencies each need a "within-the-family" research organization of their own.
>
> Seventh, The RAND Corporation (established in 1948), has always had a Board of Trustees that effectively guided and protected it. This Board is truly familiar with much research at RAND—listening to three days of briefings twice a year. Even more important, the Board comprises men of such national stature that RAND has been able to preserve its independence through various attempts by the USAF to clip its wings. [4, pp. 4-5.]

The reader may disagree with Enke's claim of RAND's great success, but I suspect that far less controversial is his added argument of RAND's preeminence among the Department of Defense oriented "Think Tanks." None of the others had RAND's autonomy which Enke thinks is explained by the fact that "the USAF has often been more displeased than pleased with what it unexpectedly got for its money for Project RAND. Senior USAF officers have in the past advised their counterparts in the Navy and Army not to make the mistake of creating so independent, influential, and uncontrollable an organization" [4, p. 7].

It will not be easy to create a setting in which researchers can gain significant knowledge about policy problems, have a great deal of autonomy in performing policy research, and be free from undue pressures. This will be especially true for social policy organizations that owe their existence to one or a few mission agencies. The efforts to institutionalize wide public disclosure and scrutiny of evaluative activities may help in establishing such a setting, but clearly the question of the researcher's independence and objectivity requires immediate discussion and debate.

At the same time, I think it is critical that social scientists keep the matter of independence in perspective. Given the dearth of relevant social policy research, it would be wrong to make these potential threats to independence the principal problem. Far more important is the question of getting the social science community moving toward a material contribution to social policymaking.

It has been argued that even the most basic types of policy research will be enhanced if the researcher can gain detailed knowledge of policy problems and concomitant research needs. Such a formulation indicates a major role for the mission agencies in basic policy research even if they are not the funding source; and, under ideal circumstances, the social agencies would be the primary source of basic research funds in their areas of concern. This is true both because of the potential for interaction between agency staff and researcher and because the agencies

will be more likely to use research results supported by their own funds.

Funding of basic policy research by the mission agencies is not without its problems. First, the agencies more than organizations such as the National Science Foundation are likely to exert undue influence or put restrictions on research organizations. Second, in times of budget decreases basic research funds in mission agenices are usually cut back significantly. These problems suggest that NSF and the National Institutes of Mental Health (which is part of a mission agency but historically has been considered an independent funding source) support more basic policy research in the social areas. The big problems in making NSF and NIMH major funding sources of basic policy research concern 1) the means of developing greater knowledge about policy problems among potential researchers in light of the fact that these organizations in the past have had only a limited interest in social agency policy needs and 2) the means for getting the research results once they are available considered by the mission agencies in their policy process. These are complex institutional problems for which no ready solutions are apparent.

A shift in funds by NSF and NIMH to basic policy research at the expense of other basic research may have some major "indirect" benefits. The change could increase not only policy information but also <u>fundamental scientific knowledge in the social sciences</u>. Remember basic policy research differs from other basic research only in requiring that the researcher have an appreciation of policy issues and needs and thinks in terms of the research implications that flow from them. Thus I suspect that the support of basic policy research by NSF and NIMH, if coupled with satisfactory institutional means of imparting policy needs to researchers, will expand the frontiers of knowledge in the social science disciplines more than if only basic, but not policy-oriented, research is funded. This argument derives from my belief that the great social science questions of our times are bound up in the great social issues. Certainly, I cannot prove this thesis, but it is not one that on the face of it should be rejected out-of-hand.

D. A Concluding Observation

One final point needs to be made: If there is relatively more funding of policy research, if more interaction occurs between government policy research staffs and researchers, and if more special social policy research organizations are established, it will not guarantee a high level of social policy research. To label an organization, the Social Policy Research Institute, and to staff it with social scientists from several disciplines insures neither that the researchers will work together in an multidisciplinary effort nor that they will address relevant social issues more effectively than in the past.

This should come as no surprise. The decision to move toward more policy work must be implemented, and this requires a lot more than a name and a functional statement. However, the move toward more policy-oriented social research seems to be a sound one. Yet the fact remains that success depends on a complex dynamic process involving the government and researchers. Perhaps more than anything else, success depends on social science responding positively to the challenge and making a real commitment to confronting the great social problems of our times.

REFERENCES

1. Brooks, H., "The Future Growth of Academic Research," in H. Orlans (ed.), Science Policy and the University, Brookings Institution, Washington, D.C., 1968.

2. Campbell, D. T., and A. Erlebacher, "How Regression Artifacts in Quasi-Experimental Evaluations Can Mistakenly Make Compensatory Education Look Harmful," in J. Hellmuth (ed.), Compensatory Education: A National Debate, Brunner/Mazel, New York, 1970.

3. Deutsch, K. W. and others, "Conditions Favoring Major Advances in Social Science," Science, 5 February 1971, pp. 450-459.

4. Enke, S., "Think Tanks for Better Government," TEMPO, General Electric Company, Santa Barbara, California, 67TMP-126, December 1967.

5. Knowledge into Action: Improving the Nation's Use of the Social Sciences, National Science Foundation, Washington, D.C., 1969.

6. Rossi, P. H., "Observations on the Organization of Social Research," a paper prepared for the Symposium on the Organization, Management and Tactics of Social Research, Cleveland, Ohio, February 20-21, 1969.

7. Teller, E., "Education of the Modern Inventor," in B. R. Keenan (ed.), Science and the University, Columbia University Press, New York, 1966.

8. The Assembly on University Goals and Governance (A First Report), American Academy of Arts and Sciences, Cambridge, Massachusetts, January 1971.

9. The Behavioral and Social Sciences: Outlooks and Needs, National Academy of Sciences, Washington, D. C., 1969.

10. Watts, H.W., "Adjusted and Extended Preliminary Results from the Urban Graduated Work Incentive Experiment," Discussion Papers, Institute for Research on Poverty, University of Wisconsin at Madison, June 10, 1970.

11. Weinberg, A. M., "But Is the Teacher Also a Citizen," in B. R. Keenan (ed.), Science and the University, Columbia University Press, New York, 1966.

12. Wholey, J. S. and others, Federal Evaluation Policy, Urban Institute, Washington, D.C., 1970.

13. Williams, W., "Developing an Evaluation Strategy for a Social Action Agency," Journal of Human Resources, Fall 1969, pp. 451-465.

14. Williams, W., Social Policy Research and Analysis: The Experience in the Federal Agencies, American Elsevier, New York, forthcoming 1971.