December 1974


Follow this and additional works at: https://scholarworks.wmich.edu/jssw

Part of the Social Work Commons, and the Sociology Commons

Recommended Citation
Available at: https://scholarworks.wmich.edu/jssw/vol2/iss2/2

This Complete Issue is brought to you by the Western Michigan University School of Social Work. For more information, please contact wmu-scholarworks@wmich.edu.
# TABLE OF CONTENTS

**VOLUME II - NO. 2 SUPPLEMENT WINTER 1974**

Issue Editor: Alan M. Cohen, University of Western Ontario, London, Ontario, Canada

<table>
<thead>
<tr>
<th>Title</th>
<th>Authors</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Editorial</td>
<td></td>
<td>pp 208</td>
</tr>
<tr>
<td>Change and Program Evaluation in Social Organizations</td>
<td>ALAN M. COHEN</td>
<td>pp 210</td>
</tr>
<tr>
<td>Behavioral Science Influence on Legislation: The Case of Delinquency Prevention</td>
<td>ALBERT S. ALISSI</td>
<td>pp 227</td>
</tr>
<tr>
<td>The Demonstration Project as a Research and Change Strategy</td>
<td>LEONARD RUTMAN</td>
<td>pp 259</td>
</tr>
<tr>
<td>The Three Evaluations of Social Welfare Programs</td>
<td>DEAN HARPER</td>
<td>pp 271</td>
</tr>
<tr>
<td></td>
<td>HAROUTUM M. BABIGIAN</td>
<td></td>
</tr>
<tr>
<td>Evaluation Research and Evaluation: Scientific Social Reform Movement and Ideology</td>
<td>MICHAEL BAIZERMAN</td>
<td>pp 277</td>
</tr>
<tr>
<td>Anti-Poverty Policies and Evaluation: A Critique of the Pluralist Conception of Politics and Evaluation</td>
<td>ROBERT D. HERMAN</td>
<td>pp 289</td>
</tr>
<tr>
<td>Conflict and Compromise in Evaluation Research: A Case Study</td>
<td>JOE HUDSON</td>
<td>pp 298</td>
</tr>
<tr>
<td></td>
<td>PETER CHOMMIE</td>
<td></td>
</tr>
</tbody>
</table>

- 207 -
As Alan M. Cohen, Editor of this supplement, notes in his introductory editorial, this issue had its genesis at the annual meeting of the S.S.S.P. in Montreal. We have made several changes in the content. We have added papers by Baizerman and Herman. The content of these papers are directly related to the topic of evaluative research. Space limitations necessitated the withdrawal of one paper; and since Kronick's paper deals with the politicking necessary to establish a social welfare service, we are planning to publish it in the summer issue.

Norman N. Goroff
Publisher and Managing Editor

EDITORIAL

In August 1974, at the annual meeting of The Society For The Study of Social Problems, the Sociology and Social Welfare Division conducted a special session on "evaluative research and social sciences in welfare policy and programs." Instead of reading the various papers in their entirety the participants of this session agreed to briefly overview their papers. The balance of the session was devoted to a general discussion of the implications of their work within the context of evaluative research in the social sciences. This discussion provided the impetus for this special edition of "Sociology and Social Welfare."

The discussants returned again and again to what was referred to as the "folklore" that evaluative research, demonstration projects and other research techniques are undertaken to provide information that in turn leads to improved services. In other words, each participant commented on the widely espoused rationality of linking research findings to the subject being "researched." The obvious assumption being that if the evaluative research indicated inadequacies, this information would at least be seriously considered and would probably lead to changes in those areas shown by the research to be deficient in some way.

Despite almost universal agreement that actual rationality in linking research information to program change is exceedingly rare, new evaluation studies and research proposals continue to consume millions of dollars and man hours.

Why?

This special issue includes six papers that address at least a few of the possible responses to this question. Harper and Babigian discuss three distinct types of evaluation. They
emphasize these distinctions because methods of evaluation are partly dictated by the intent of the evaluation. The Alissi paper discusses "some of the issues and problems which arise in the field of delinquency prevention where there is overriding public interest and concern and calls for action to prevent delinquency but where the problem is elusive and where the variables are not easily isolated and controlled through experimental procedures." He provides many examples of the problems of linking research to legislative decision making. Rutnam uses the famous St. Paul "Family Centred Project" to elaborate on the incompatibility between the two primary goals of a demonstration project - evaluating the program and using it as an instrument of change. Hudson and Crommie outline many of the problems encountered in the process of conducting evaluation research in the field of corrections. They illustrate through their case material problems involved in planning and conducting evaluative research. They further illustrate the lack of rationality in utilizing the findings derived from the project. Kronick uses the case study of founding a group home to illustrate behind-the-scenes politicking necessary to get a project operative. Again, the lack of rationality in linking information to decision making is clearly demonstrated. Cohen briefly refers to three evaluation studies and demonstration projects that fail to link information to program change. He discusses possible reasons for this failure and suggests a pre-planning functional information base as an important sequence in creating a more receptive environment for organizational change.

The special issue of the Journal of Sociology and Social Welfare does not presume to grapple with all the complexities surrounding the politics of evaluative research. It is acknowledged that much has already been written about the problems of evaluation. It is further acknowledged that reasons for the very limited success of evaluative research to generate useful organizational change defy a simple clear explanation. However, it is the hope of the editor that the articles selected for this special edition will re-emphasize the IRRATIONALITY behind much of our heightened fervor for "evaluation." Hopefully this may further efforts towards closing the irrationality gaps between program information and organization change.

Alan M. Cohen
Supplement Editor
University of Western Ontario
London, Ontario, Canada
"Change and Program Evaluation in Social Organizations"

Alan M. Cohen
Sociology Department
University of Western Ontario
London Ontario Canada

There is an assumption of an inherent rationality in linking information on program effectiveness to program change. This article briefly discusses three typical evaluation studies and demonstration projects that fail to link information generated on the effectiveness of what people do, to program changes. Perceived inaccuracy of the information and the perceived threat of the information are emphasized as two reasons for this failure of program information to affect change in social organizations. A pre-planning functional information base is proposed as an important prerequisite in the sequence of creating a more receptive environment for organizational change.

It is often assumed that information on the effectiveness of what an organization does will be utilized to improve those programs. In social organizations these program changes are expected to provide more effective service to clients. This assumption of an inherent rationality linking information on program effectiveness to program change has led to a multitude of evaluation studies and demonstration projects to provide information on program effectiveness.

A review of actual cases of evaluative research and demonstration projects often does not substantiate the assumption of an inherent rationality in linking information to policy formation.

This paper will briefly review three typical case studies that fail to link newly generated program information to policy change. It will suggest two major reasons for this failure and then outline a pre-planning functional information base as an important prerequisite in creating a more receptive environment for program information to affect policy changes in social organizations.

Three typical cases where program information resulting from evaluation research and demonstration projects failed to affect program changes:

1. The original purpose of the St. Paul Minnesota
"Family Centred Project" was to develop a master plan for reorganizing health, welfare and recreation services in American Communities. (Rutman, 1974:2) National and international attention to the 1948 and 1954 phases of this project was so extensive that a 1962 survey found 143 communities in Canada and the United States had multi-problem families projects using at least some aspects of the St. Paul model, at some stage of thought, planning or operation. Despite this apparent success in elaborating and disseminating new conceptualizations and case work techniques for multi-problem families, the Family Centred Project eventually abandoned its treatment focus. For many reasons, some of which are elaborated upon in Leonard Rutman's (1974) paper, the widely publicized information on potential improvement in service had very limited long range impact on relevant program change.

2. A second typical case where relevant program information failed to affect program changes is a four school study in New York referred to by Kruse (1968). This project was undertaken in the 1960's to ascertain the effect of extensive agency involvement in the schools. All four schools were in roughly similar circumstances in disadvantaged areas of New York City. Two of the four schools were subjected to extensive group and case work involving both students and their families. The other two schools were used as control groups. After two years, the findings indicated there were NO significant differences in the social problem characteristics of these two groups. Despite these findings, the usual agency services continued to operate uninterrupted.

3. A final example where pertinent information failed to affect change is a 1970-71 evaluation study on the relative effects of parole supervision as provided to a group of juvenile boys and girls (Hudson, 1974). The major aim of the study was to determine whether juveniles released from two major state level correctional institutions in Minnesota would adjust as well on parole without formal supervision from parole officers, as a corresponding group receiving conventional parole supervision. The research techniques are elaborated upon in considerable detail in Hudson's
This study found that exposure to parole supervision does NOT have a significant positive effect upon the parole adjustment of juvenile-aged boys and girls. Hudson (1974:15-16) states "the major conclusions were included in a summary of the project which was distributed to all parole officers, supervisors, and administrators as well as to the members of the Parole Board. (All these groups had extensive involvement and commitment to the research). In conjunction with this summary of the project findings, the investigator communicated his willingness to meet with and discuss the programmatic implications of the research. No response to this offer was ever made by either the practitioners, supervisors or administrators of the parole division or by the members of the Parole Board. Almost totally ignored, the project became, for the purpose of the state, a "non-study -- left to gather dust on a shelf".

There is a considerable danger in generalizing on a subject as extensive as social change. However, two of the major reasons why program information, such as generated in the three cases referred to above, is not more extensively utilized in policy making can be outlined.

TWO REASONS WHY LINKING PROGRAM INFORMATION TO DECISION-MAKING FAILS TO YIELD PROGRAM CHANGE

The first reason why program information often fails to lead to change is the perceived INACCURACY of the information.

Evaluation of social programs is often beyond the capabilities of existing methodology. The number of studies that measure up to minimum standards of scientific adequacy is generally accepted to be exceedingly small. The following nine factors account for much of the perceived inaccuracy of the information generated by evaluation research and demonstration projects in social organizations.

Nine factors that affect the accuracy of the information generated by evaluation research and demonstration projects are:

1. The difficulty of measuring the effectiveness of
the programs. For example:

(a) what constitutes the appropriate criteria of outcome.

(b) output definition and measurement problems, especially where programs are multi-purpose.

(c) some outputs are easier to measure than others.

(d) problems of attributing the effect of the process on the outcome. In other words, the research design, selection and administration affect the outcome. An important question is how extensive is this effect?

(e) difficulties in directly observing states of "well being". For example: Is a person who is mobile with considerable pain healthier than someone who is bedridden? etc.

2. The need to include client opinion, as well as professional judgement, in considering treatment outcome. This creates many methodological problems including:

(a) how should the client's judgement be weighed in the evaluation of services? For example, the client may express gratitude or frustration for service rendered, rather than an actual change in condition.

(b) some researchers have noted that treatment outcome is closely related to client motivation. Given our understanding that equality of opportunity cannot be equated with the equality of condition, we can predict that different clients may perceive identical treatment in entirely different terms.

3. The inability to check the validity of the research. Is the project evaluating what it claims to evaluate? Problems here include:
(a) outside criteria as an external source to compare the measuring instrument are usually not available.

(b) There is often a lack of uniformity in recording case records between agencies.

4. The difficulty of obtaining a scientifically valid sample of clients. For example, even the "mix" of the population to be served, e.g., self-referral, private agency referral or government agency referral, may affect the results.

5. The lack of a clearly conceptualized treatment method in many research programs. For instance, what treatment variables should be included. e.g., Length of involvement, sex, insights, etc.

6. The difficulty of fitting research into the normal routines of the agency and the resultant bias on outcomes.

7. The complexity of both the internal organizational structure and the many levels of interaction with the external environment.

8. The inability to replicate most evaluative studies because of continued changes in both the internal and external environments.

9. The difficulty in assessing how the values and attitudes of the professionals affect the service outcome. Evaluation rests upon a set of values but technical procedures of evaluation rarely challenge these values. (Fried, 1968:43)

A second reason why linking program information to policy fails to yield program change is that the information is often perceived as a THREAT.

This perception of program information as a threat is shared
by both the professional workers and administrators. For example, if a counselling program is found to be relatively ineffective it may appear to imply the workers are not performing properly. It may also appear to question the validity of the professional approach, and, it may even appear to threaten the very raison d'être of the organization. Dedication to organization survival and maintenance often leads to a "success cult". (Carter, 1973:17) When the success of programs are thought to rely on results of evaluation to "pass", the threat of not passing often leads to obscuring evaluative results.

Another threat can develop from the potential for drawing unwarranted conclusions from those who "don't understand" the limitations as outlined above, on the accuracy of program information.

The lack of credibility of the information generated by evaluation research and demonstration projects, plus the perceived threat of program information, have both contributed to the failure of this information to affect program changes. The author has elsewhere (Cohen, 1974) discussed other characteristics and imperatives that affect change in social organizations. While acknowledging the broad complexity of the subject of change in social organizations, it is proposed that a pre-planning functional information base can assist in overcoming the failure to link information on program effectiveness to program change.

THE FUNCTIONAL INFORMATION BASE

The creation of a more receptive environment for program information to affect policy changes is a primary goal of a functional information base. A key question to ask in most organizations at the present time is "How can you plan when you do not know what you are planning for"? Detailed budget records, statistics on staff employed and number of clients serviced, provide the basis for most of an organization's present knowledge of its operations. A serious shortcoming with this traditional approach to classifying organizational data by "objects" of income and expenditures is that it does not directly identify how all input resources including rent, administration, travel costs, etc. are distributed to the various service programs which in total constitute what the organization does. A number of recent studies have found large discrepancies between what the staff and the board of an organization think or say they are
doing, and what they are actually doing in terms of a total deployment of resources to all actual service programs and functions.

The information base envisaged here considers that all data is only data unless it can be used for decision making, and only then does it become information. The types of decisions to be made must be anticipated. In social welfare organizations for instance, this usually implies the need for common caseload reporting with data related to categories of people. (e.g., emotionally disturbed, physically handicapped, one parent families, etc.) These categories of people have needs. Measuring the extent of organizational success in meeting the needs of these categories of people provides output information for making decisions on any changes to be made in existing programs. The decisions relate to resource allocation to services to meet the needs of the people.

To provide an appropriate information base for decision making a functional information base is considered a pre-planning necessity. Following the implementation of this functional accounting system an organization should move towards the processes of P.P.B.S. (Planning, programming, budgeting system.) (See Goodman (1969) for a discussion on the dangers in making functional budgeting an end unto itself.) The spirit of P.P.B.S. is a marriage between program planning and budgeting. However, the program information utilized in P.P.B.S. must be mature if the marriage is to last. This maturing comes through the processes of acceptance that the functional information base is reasonably accurate. A major deficiency in many information systems such as P.P.B.S. is they ignore the need for a sequential maturing stage before attempting to link program information to the decision-making processes of organizational change.

ADVANTAGES OF A FUNCTIONAL INFORMATION BASE

There are a number of reasons for assuming that the information base described above will create a more receptive environment for organizational change than existing evaluation research and demonstration project techniques.

A functional information data base reflects with reasonable accuracy what resources are actually providing the various programs the organization is undertaking. This data collection on existing programs is not a one shot project but an ongoing process that allows ample opportunity for correctional feed-
back from all levels of the organization.

This paper has discussed implications of perceived inaccuracy of information as an important variable inhibiting organizational change. The involvement of all levels of the organization in determining how resources are allocated to the various organizational programs should result in an acceptance that this functional information base is at least a reasonable representation of what resources are allocated to what programs. All three cases referred to earlier in this paper lacked this commitment to basic program information.

Beside an increased consensus on the accuracy of organizational program information, another advantage of the functional information base is that the provision of information by programs and services encourages all levels of the organization to focus on ends rather than means. For example, travel expenditures are not categorized separately, but are grouped to programs by considering travel for what . . . to do . . . for what services? This need to collect unit cost information by at least crude benefits to the people served logically leads to questions regarding effectiveness in meeting the needs of the people being served. For example a counselling service with costs over $200 per unit of service will probably lead to questions about possible less costly alternatives . . . without a forced review through a formal evaluation with its many limitations already referred to.

A further advantage of a functional information base is a reduction in the perceived threat to the workers and administrators. Criticisms can be countered with information that reasonably reflects what services are being provided and what it costs to provide them. It is also possible to relate these costs to at least crude output indicators that reflect the impact these services appear to have on the client population.

Obviously the information generated by this functional information base will not always reflect what everyone in the organization would expect or readily accept. This data is called a pre-planning information base because it is specifically emphasized that policy making should follow the data collection. It is true that the information engineering is always possible. However, by strictly limiting this information base to what programs are actually being undertaken and how all organizational resources are utilized to provide these programs, the political reality of change is placed in sequential order.
In other words difficult political decisions can be reviewed from the perspective of a more accurate organizational information base that is accepted by those responsible for the provision of the services.

EXAMPLES OF FUNCTIONAL INFORMATION

The scope of this paper does not permit detailed examination of the processes of the program information generation and its subsequent linkage to organizational change. However, a few specific examples should illustrate the potential effect on program change.

Table 1 is the actual functional budget of a small family agency in a fairly typical North American community of approximately 300,000. The left hand column lists the proportion of total resources including administration required to provide these services, while the final three columns show actual and proposed dollars. In this case the funders of the organization were under the impression their dollars were providing family and individual counselling. When the newly generated functional accounting results of Table 1 indicated only little over 50% of the agencies' resources were being devoted to this service, the major funders threatened to withdraw part of the next year's funding. This led to significant changes in the organization.

Tables 2 and 3 provide information on the major functions of a Y.W.C.A. in a North American community along with functional accounting information on dollar income, expenses and deficits. The fact that 42.6% of the projected deficit resulted from the Health and Physical Education program and in particular from the operation of the swimming pool, led to considerable discussion of possible changes in this program area. The functional information on the proportion of resources devoted to the various program areas, and the projected deficit attributable to these various programs have already led to organizational change.

Functional information on two adoption agencies operating in the same city revealed that the rate per adoption between the two agencies was as follows:

Agency A - Rate per adoption - $338.00
Agency B - Rate per adoption $1,368.00
<table>
<thead>
<tr>
<th>Service</th>
<th>Total</th>
<th>100.00%</th>
<th>Sub Total</th>
<th>1973 Budget</th>
<th>1974 Budget</th>
<th>1975 Budget</th>
</tr>
</thead>
<tbody>
<tr>
<td>Family Life Education</td>
<td></td>
<td></td>
<td></td>
<td>8.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family &amp; Individual Counseling</td>
<td></td>
<td></td>
<td></td>
<td>53.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job Placement Services</td>
<td></td>
<td></td>
<td></td>
<td>20.85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Information &amp; Referral</td>
<td></td>
<td></td>
<td></td>
<td>16.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Development</td>
<td></td>
<td></td>
<td></td>
<td>7.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community Planning, Coordinating</td>
<td></td>
<td></td>
<td></td>
<td>1.19</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Functional Budget (With Projected Costs) of A Family Agency

Table 1
<table>
<thead>
<tr>
<th>Function</th>
<th>Proportion of Total Resources used for function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Hostel and Transient Program (The provision of shelter)</td>
<td>39.9%</td>
</tr>
<tr>
<td>2. Health and Physical Fitness (Pool and Gym)</td>
<td>30.2%</td>
</tr>
<tr>
<td>3. Personal and Social Development (Courses other than vocational)</td>
<td>9.4%</td>
</tr>
<tr>
<td>4. Vocational Training (typing courses)</td>
<td>6.0%</td>
</tr>
<tr>
<td>5. Big Sisters (Volunteer parental substitutes)</td>
<td>2.6%</td>
</tr>
<tr>
<td>6. Head Start (Including a nursery service for paying customers)</td>
<td>7.6%</td>
</tr>
<tr>
<td>7. Other (including the restaurant service and a number of others)</td>
<td>4.3%</td>
</tr>
<tr>
<td>Function</td>
<td>Annual Income</td>
</tr>
<tr>
<td>-----------------------------------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Hostel</td>
<td>78,558.</td>
</tr>
<tr>
<td>Health &amp; Physical Education: Recreation</td>
<td>38,718.</td>
</tr>
<tr>
<td>Personal &amp; Social Development</td>
<td>10,137.</td>
</tr>
<tr>
<td>Vocational Training</td>
<td>8,764.</td>
</tr>
<tr>
<td>Big Sisters</td>
<td>-</td>
</tr>
<tr>
<td>Head Start, etc.</td>
<td>13,150.</td>
</tr>
<tr>
<td>Other</td>
<td>7,542.</td>
</tr>
<tr>
<td><strong>Totals</strong></td>
<td><strong>$156,869</strong></td>
</tr>
</tbody>
</table>

1 Including central services and administration.
This huge discrepancy in costs had existed for many years but major funders were not aware of it. In fact Agency A was considered by many, to be an "inefficient operation". Substantial differences in the philosophical approach to adoption between the two agencies exist, with the more expensive service utilizing a foster care program prior to adoption. However, it was only after this functional information was available that serious consideration was given to determining the possible advantages of one approach over the other. This type of questioning, arising out of actual differences in programs purporting to accomplish the same goals, has considerable potential for productive evaluation research.

Vancouver Canada has a uniform functional information system operating across many non governmental social service agencies. (Vancouver, 1973) Twenty-four agencies reported they were providing a service defined as "Family and Individual Counselling" at a unit cost ranging from $3.11 to $29.11. The Girl Guides of Vancouver had a unit cost of $1.41 for a service defined as "Group Work", while the Y.W.C.A. unit costs for the same "Group Work" service category were approximately $17.00. Obviously these two agencies were not "doing" the same thing in their provision of "group work".

The relevant point in both these Vancouver examples is that simply knowing the wide cost variances within supposedly similar services, permits the asking of questions about effectiveness of these services that would not be possible without this unit cost functional information.

SUMMARY

This paper is not proposing that the provision of a functional information base will eliminate the problems related to change in social organizations. What is suggested is that existing information in most social organizations is not adequate to cope with the consequences of evaluation research and demonstration projects. The result is that many important implications for social change are ignored. The creation of a more receptive environment for organization change will not just happen. A functional information base for social organizations is one way to increase the potential for more effective change in meeting the needs of people.
1. Social organizations for the purposes of this paper refer to the broad range of organizations providing services rather than producing products as their primary function. Health, education and welfare organizations are obvious examples of social organizations.

2. See appendix for a brief discussion on the definitions of evaluation studies and demonstration projects.

3. For instance, London, Ontario spent over $50,000 on a multi-problem family project in the 1960's. This demonstration project attempted to replicate many of the St. Paul concepts and operations. After three years and the publishing of an extensive report the entire London project was terminated with no long term effects on program change.

4. Geismar (1973:14) has noted that "research relating service characteristics to treatment results are extremely rare".

5. This information base should not be equated with the total needs for information in an organization. For instance it initially does not emphasize objectives, program evaluation and many areas of internal and external communication. This information base is concerned with detailing where all input into the organization is deployed in terms of what the organization is doing (its programs) and accounted for in units that allow for eventual evaluation of the effectiveness of these programs.

6. See Cohen (1974) and Frank (1973) for relevant discussions on P.P.B.S.
APPENDIX

Evaluation studies and demonstration projects have been defined and redefined, organized and reorganized in so many ways that it is impractical to overview the subject in this paper. For example, Carter and Wharf (1973) list fifteen pages of bibliography related to evaluating social development programs. This list of eighty books and one hundred fourteen articles, periodicals and papers is far from a comprehensive list related to this subject alone.

Within the context of this paper "a demonstration project is defined as a field experiment designed to test the value and feasibility as well as the practical implications of

(a) a new program or new social structure
(b) an inventive approach to providing services or dealing with new and unresolved social problems
(c) adopting an operational program to a new setting
(d) modifying existing programs to make them more responsive to current social problems and needs with the objective of improving the character and/or quality of the social environment and of services rendered by the organizations involved". (Canada, Demonstration Projects, 1970.)

For the purposes of this paper evaluation studies, evaluation research and similar designations assume the Webster's dictionary definition of evaulate: "determine the worth of; appraise".

Although the implications of the extent of scientific sophistication in evaluations is referred to in this paper, a more detailed analysis is not practical here. For instance, some attempts have been made to differentiate evaluation studies from evaluative research. Evaluation studies often refer to the broad spectrum of attempts to determine how effective programs are in meeting their objectives. Even informal interviews and observations are included in this classification. On the other hand, evaluative research is often limited to a more rigid scientifically disciplined approach. This more rigid evaluation would usually include criteria such as Coplans" (1968) three fold classification of process evaluation, achievement evaluation and administrative
evaluation, using scientifically accepted procedures to measure effectiveness. There is, however, little consensus on these definitional differentiations.

Within the context of this paper it is considered appropriate to include a wide spectrum of definitions by assuming the broad generic dictionary definition of evaluation.

REFERENCES

Canada

Caplan, G.

Caro, F. G.

Carter, N. and B. Wharf

Cohen, A. M.
1973  "Change and Social Organization". Journal of Sociology and Social Welfare (Fall): 86-92

Cohen, A. M.

Frank, J. E.
Fried, M.  

Geismar, L. and I. Wolock  
1973  "Research Data as Aids in Formulating Agency Policy". Rutgers University Graduate School of Social Work.

Goodman, N.  

Hudson, J.  

Kruse, A. H.  

Rutman, L.  

Vancouver  
1973  "Vancouver P.P.B.S. Project-Progress Report". United Community Services of Greater Vancouver. Vancouver, B. C.

Weiss, C. H.  
I. THE PROBLEM

Although progress has been uneven and the parameters for establishing interdisciplin ary approaches remain to be settled, new ways of using the behavioral sciences in the field of law are beginning to be explored more systematically. At this point, the relevance of behavioral science materials in legislation, litigation, administrative and evaluative activities of government does suggest some of the most promising areas for further exploration and development.

Social science knowledge can be used, of course, to influence legislation in various ways. It can serve, for example, to shed light on some aspects of behavior which are of concern and which may be potentially subject to legislation. Hence, knowledge regarding the nature of deviance such as delinquency, criminality, mental illness or of the nature of the normal childhood, adolescent, maturing and aging processes can form the basis upon which sound social policies are advanced and desirable legislation enacted. Behavioral science may also be useful in the more direct examination of some of the legal institutions and social control systems to determine how they work and to assess outcomes. For example, studies of the criminal justice system, that is, the relative functioning of the police, courts, juries and correctional personnel can offer new insights into how agencies of social control operate and point up the need for legislative reforms.

A variety of research tools and techniques may be applied to provide the "hard" data to influence legislation. Ideally, where the problem may be clearly defined and the significant variables controlled through careful experimentation so that findings can be shown to be "Conclusive", the arguments for influencing legislation would appear to be compelling. But what happens in the more usual case where this ideal is not met?

This paper will discuss some of the issues and problems which arise in the field of delinquency prevention where there is over-riding public interest and concern and calls for action to prevent delinquency, but where the problem is elusive and where the variables are not easily isolated and controlled through experimental procedures. What, in fact, constitutes the body of knowledge in the field? What uses can be made of behavioral science materials where there has been little or no experimental successes, or data upon which to build action programs? What credence should be given to the largely unverified "practice" knowledge and experience in delinquency prevention programs? These questions will be explored here in an effort to provide a more responsible frame of reference for suggesting and assessing legislation in the field of delinquency prevention.

Basically, it will be shown that although there have been many theoretical breakthroughs, numerous studies and even more action programs, our knowledge as to
how to go about preventing delinquency is surprisingly limited. Although it hardly
increases confidence in one's ability to influence legislation, it does not follow
that there is no role of the behavioral sciences in the matter. To be sure, it may
be argued, for example, that lack of knowledge itself can provide a basis for influ-
legislative action which would substantially increase our knowledge of program ef-
ficiveness. Or, at the very least, what is already known may provide a rationale
for rejection poorly formulated schemes which claim to be able to "eliminate"
delinquency.

Our discussion will focus briefly on the nature of prevention and the ways in
which different perspectives come to influence definitions of the problem and limit
the scope of the proposed solutions. Notwithstanding the importance and need for
adequate theory, major attention will be given over to a review of delinquency pre-
vention programs in an effort to highlight the variety of program goals, limited use
of research methods and paucity of the findings. And finally, some of the problems
and issues regarding the ways in which programs selectively define, label and treat
individuals in the name of prevention will be explored to provide some guidance in
further assessing how social science knowledge may be useful in legislation.

II. VARYING PERSPECTIVES OF DEVIANCE AND THE CONCEPT OF PREVENTION

The development of effective conceptual tools has been a slow and tedious
process in the behavioral sciences. Some doubt whether it will ever be possible to
invent the adequate means to portray social reality in its totality. This would not
surprise us, however, for it is true that we never see things in their total con-
creteness. We see only certain aspects - those that we have been taught to abstract
using the currency of our own cultural symbols. As Walter Lippman expressed in his
famous aphorism, "First we look, then we name and only then do we see." A group of
young people hanging on the corner, obstructing pedestrians, cluttering entrances
to buildings, brings about different reactions depending upon whether they are de-
defined as demonstrators, strikers, Christmans carolers or members of the loach
"Black-Hawk" gang. It is therefore, in the naming and defining of behavior that we
come to "see" its significance.

What passes for knowledge and understanding, then, must center on how we
arrive at these definitions. In exploring the sociology of knowledge, Mannheim
introduced the notion of relationism which states that truth is not necessarily a
fixed commodity but is predicated on the historical and situational context in which
it is found. As cross-cultural studies have repeatedly demonstrated, our own
involvement and narrowed cultural frame-of-reference in a sense, institutionalizes
our own versions of the truth. As Hall put it, "Culture hides much more than it
reveals, and strangely enough what it hides, it hides most effectively from its own
participants." In short, what we see or overlook depends on the concepts available
to us.

So it is with the helping professions where even the most carefully description
of delinquency are not automatically objective for all behavioral definitions have
pre-established connotations. There is a ready set of concepts when it comes to
delinquency. Thus, we have an image of the delinquent which serves as a prophecy
we are altogether too eager to fulfill as selected aspects are highlighted, blotting
out others not conveniently covered by the label. The pre-fabricated conceptual molds come in all sizes and shapes. One set forces us to recognize anti-establishmentism, the generation gap, contradicting sub-cultural values, etc. A psychoanalytic set calls attention to dependency, ambivalence, psychosexual conflicts resurfing oedipal conflicts. And still another perspective sets us searching for anomia in the social system, status deprivations and role confusions in social and interpersonal relationships.

Similarly, concepts of prevention vary considerably depending on the set of concepts one employs. Hence, there are major differences of opinion regarding what is to be prevented, who is to be deterred, and what constitutes successful intervention. There are, first of all, those who see solutions in terms of broad attempts to change and upgrade life styles. Those who are particularly impressed with social and cultural determinants of behavior tend to emphasize programs aimed at treating the entire social milieu in which all children grow up. The theory is that the best way to prevent delinquency is by enhancing healthy growth and development. Thus, the focus of attention is on improving community and family life. As housing, employment, recreation, social and cultural conditions improve, delinquency will accordingly be prevented.

The second approach is less global in that the focus of attention is limited to those children who are deemed to be "at risk" or have given some reason to believe that they will become delinquent unless something is done to prevent it from happening. Advocates of this approach emphasize early detection and direct "treatment" of those children who are identified as vulnerable, disturbed, troubled, or potential delinquents.

The third approach is the most constructive insofar as the notion of prevention is more precise and limited insofar as it applies solely to known or adjudicated delinquents. In this sense, programs which are designed to correct or rehabilitate delinquent youngsters prevent delinquency by reducing recidivism and lessening the likelihood that further delinquent acts would be committed.

These three perspectives form a convenient framework for review of delinquency prevention programs.

III REVIEW OF DELINQUENCY PREVENTION PROGRAMS, RESEARCH METHODS AND FINDINGS.

A. Improving Community Life Styles

Perhaps one of the better known approaches to delinquency prevention which sought to reach the delinquent in his natural social environment was the Chicago Area Project initiated by Clifford Shaw and his associates in 1934. The program was based on the premise that delinquency, in most cases, was a process of social learning, experienced by youth growing up in delinquency areas. The social milieu in which delinquent boys were nurtured - the family, play group and neighborhood - all had to be considered in any preventative measures to reduce delinquent behavior.

Accordingly, the key to prevention was seen to depend on broad level change in the character of the neighborhood life whereby the community would become more
"wholesome" for children. Moreover, there was the conviction that change could best be implemented by the self-help efforts of neighborhood residents, themselves. By working together on commonly defined neighborhood problems it was reasoned the increased sense of neighborhood feeling and responsibility would develop. Natural leaders from the neighborhood were recruited to prove recreational and educational activities and to provide healthy role models for growing boys. The idea was "to give the pre-delinquent or delinquent boy an opportunity to form an attachment to, or come under the influence of, a person or persons from whom he would receive recognition for conforming to conventional standards of conduct." 5

Unfortunately, the project was not subjected to strict experimentation and whatever results were measured were inconclusive. It was demonstrated that residents in low income areas could, indeed, organize themselves, sustain community organizations and attract local "natural" leadership which would have some impact on community life. But there were no precise measurements or "hard" evidence to show that there was any actual reduction in delinquency which could be attributed to the program. 6

A series of other community-based programs were conducted which were similar in conception to the Chicago Area Project. The Midcity Project was initiated in a lower social economic area in Boston in 1954. 7 This was a multi-faceted program focused on improving community conditions to reduce delinquency. One of the major goals in the program included efforts to coordinate the work of existing social agencies to help "problem" families and to reach out to delinquent youths. Over a three year period, seven project field workers maintained contact with approximately four hundred youths between the ages of 12 and 21. It was determined, on evaluation, that there was "no significant measureable inhibition to either law violation or unethical behavior as a consequence of the project efforts." 8

Similarly, the South Central Youth Project, 9 established in 1955 by the Minneapolis Community Welfare Council sought to coordinate social agencies to detect and reverse the processes leading to delinquency by improving community and neighborhood life. Again, no adequate measures of program outcome were available although one finding was that the lack of cooperation and communication between agencies and community residents resulted in "ineffective services", generally.

More recently, the Office of Juvenile Delinquency and Youth Development of the Department of Health, Education and Welfare funded a number of community-based programs 10 aimed at delinquency prevention all of which may be viewed as modern-day counterparts of the Chicago Area Project. The Mobilization for Youth, 11 for example, sought to organize neighborhood groups to engage in self-help social action to overcome apathy and defeatism evidenced in a sixty-seven block area in a lower socio-economic area of New York City. It was conceived that as poor people organized and developed as a power base, though would be able to deal more effectively with poor housing, poor schools, inadequate police protection, etc. In the process, deprived youth would be offered new opportunities to develop motivation and skills to advance themselves. It was held that with increased competition would come increased social controls and less delinquency. The focus on developing a political power resulted, however, in increased conflicts and resistance from other governmental agencies.
Similarly, the Syracuse Crusade for Opportunity\textsuperscript{12} sought to mobilize community groups to engage in social action to counteract the conditions of poverty, transience and chronic dependency on the theory that delinquency would, in turn, be prevented. It was found that contrary to the experience of the Chicago Area Project, friction between professional staffs and neighborhood residents over decision-making powers was most apparent. No evidence was found to show how delinquency was reduced although some activities provided the youth appeared to have a positive impact. The United Planning Organization\textsuperscript{13} in Washington, D.C. represented another effort to mobilize citizen action and involvement in creating new opportunities for youth. Unique it relied on decentralized decision-making activities where a variety of relatively independent block clubs and social betterment programs were organized. But, gaps, in communication and coordination of effort was held to be a major shortcoming of the project.

Although the Houston Action for Youth\textsuperscript{14} was also designed to mobilize community groups, its program which included stable, working population utilized low-pressure tactics in approaching area problems on the theory that confrontation with other governmental agencies would be counterproductive. Varying degrees of success were recorded but no clear evidence regarding delinquency prevention could be produced. Another low pressured approach was used in the Action for Appalachian Youth Program\textsuperscript{15} in Charleston, West Virginia where much of the effort was on the improvement of a rural target area through the use of neighborhood organization associations. In a sense, the program sought to reduce the stresses and strains between rural and urban life styles which were seen to be instrumental in creating delinquent behavior. Although some problem solving occurred, the net impact on the isolated communities seemed to be relatively small.

And finally, the Harlem Youth Opportunities Unlimited Associated (HARYOU ACT)\textsuperscript{16} was initiated in Harlem, New York to reach a target population of a quarter of a million people. Based on the now familiar goal of creating opportunities for youth, the program sought to introduce broad based community change through citizen participation and action. But again, the massive urban problems were not solvable and the results of the effort were judged to be negligible.

In most of these recent programs many of the potential benefits were somewhat obscured by the fact that controversy and conflicts emerged where competing interest groups and political powers often redirected efforts into diffuse channels. In addition, a number of the programs were found to be administered in a weak and ineffective manner. And, the broad scope and multi-faceted nature of such programs made it unlikely that any conclusive evidence could be gathered to show direct cause and effect relationships between the intervention and delinquency reduction. The lack of any controlled comparisons makes it difficult to verify claims of success.

A number of studies have been aimed at evaluating community recreation and informal education programs. In the field of recreation, there are many different kinds of organizations and programs embodying varying philosophies, objectives and programs. Given the broad scope of its activities, it would appear that virtually every effort to reach out and channel or direct the energies of youth...
could in one way or another be considered to be a program of prevention. To be sure, the public has come to accept and recognize recreation as a major tool in delinquency prevention. Yet although many surveys have been conducted to determine the need for recreation and to assess the quantity and quality of the programs being offered, few evaluative studies have been conducted to demonstrate how positive recreational experiences lead to a reduction of delinquency.

Moreover, the few studies that were conducted seemed to indicate that there was no direct relationship between the two. In fact, one study even suggested that perhaps recreational activities may increase opportunities for delinquent escapades.

Ethel Shanas and Catherine Dunning's early study, conducted in the 1930's for the Chicago Recreation Commission, concluded that not much could be expected from the traditional recreation programs in prevention of delinquency. The study compared delinquents and non-delinquents' use of recreational facilities in four slum areas and one middle class area in Chicago. Among the approximately 15,000 boys and 8,000 girls, 10 to 17 years old, involved in the study, it was found that although a larger number of non-delinquents used the recreational facilities, delinquent boys tended to spend more time in the recreation activities but less time in 'supervised' programs. Generally, however, recreation tended to attract relatively few children especially among the older groups where delinquency rates were found to be the greatest.

Another study was conducted during the same period by Andrew Truxall for the National Recreation Commission. He tabulated juvenile court cases by health department census tracts and compared them with the existence of recreation centers. In spite of certain methodological problems, it could not be shown that there was lower delinquency rates in areas where there were recreation facilities compared to those without such facilities.

Frederick Thrasher conducted an early study of one voluntary "leisure-time" agency - a Boys Club in the City of New York - in an effort to discover its effectiveness in the prevention of juvenile delinquency. In contrast to the Shanas and Dunning findings, the Boys Club was able to attract larger proportions of delinquents and "apt-to-be-delinquents" into its recreation programs. It was found, however, that of the total official offenses of Boys Club members, 18 percent occurred before joining the Club, 28 percent took place after some participation in the Club program, and 61 percent occurred while the boys were actively affiliated with the program. Some of this could be explained in terms of the larger proportion of older boys (where delinquency is more prevalent) who were in the program. Based on case studies of 60 "problem" boys, Thrasher concluded that "acute behavior problems...peripated the various combinations of family disorganization, dire poverty, school maladjustments, gang activities, association with older hoodlums and underworld characters - were beyond the power of the Boys Club to neutralize."

In another study of the effectiveness of the Boys Club on the prevention of delinquency, Brown and Dodson compared delinquency rates of white boys in three areas of Louisville, Kentucky. They reported reduced rates of delinquency which there was a Boys Club while delinquency rates increased in the areas
without Club programs. In the absence of experimental controls, however, there
was no way to assess the influences of other social variables on the changing
rates. The authors had to conclude that although it might be comforting to
know that the Boys Club "contributes to the alleviation of the delinquency
problem" the findings add "little to basic knowledge on the prevention of
delinquency."22

B. Redirecting Potential Delinquents

Various innovative programs have emerged in an effort to reach out and
work with potential delinquents before their difficulties become too serious.

The Cambridge-Somerville Youth Study23 was established in 1935 with the
dual purpose of preventing delinquency and providing scientific research data
on delinquency prevention. The study ranks as one of the better studies which
utilized a built-in experimental design. The major objective was to determine
whether on-going relationships with friendly adult counselors would prevent
vulnerable children from becoming delinquent.

Teachers, social workers and juvenile justice personnel referred hundreds
of boys under the age of twelve who appeared to be heading for delinquent
careers. The project staff arrived at a population of 325 matched pairs based
on data supplied from questionnaires, psychological and medical tests and interviews
with parents, teachers and the boys, themselves. One half of the boys were randomly
assigned to receive ongoing counseling while the other half constituted a control
group for comparison purposes. Treatment lasted over a seven year period.
While the program was in progress, psychological tests were given, checks were
made on school adjustment and court records were examined. These failed to
disclose any significant differences between the treatment and control groups.

Edwin Powers and Welen Witmer24 conducted an in-depth study of the program
almost three years after the program ended. They found that according to juvenile
court and criminal records, the control group had committed as many and as serious
offenses as had the experimental group. Again, in 1956, Joan and William McCord25
conducted another follow-up study, this time tracing 253 of the matched boys into
adulthood. Their many comparisons led them to conclude that treatment variations
and subsequent convictions provided no evidence that the treatment had successfully
deterred criminality.26 The data did, however, suggest that "intensive" treatment
might have been more beneficial than the more general "Friendly guidance" approach.
Although judged to be a failure, the program has been acclaimed as a model in
experimental design which when applied consistently would provide a sound basis
for assessing program effectiveness.

Another careful experimental research design was employed by Meyer, Borgatta
and Jones27 in their study of four entering cohorts of girls in a vocational high
school in New York City. The major goal was to determine the effectiveness of
social work intervention in preventing "problem" behavior. Records containing
data on school behavior, personal characteristics and family situations were used
to screen potential problems cases. A random procedure was used to refer 189 of
the cases to a social work agency for on-going casework services to be compared
with 192 non-treated control cases. A wide range of data regarding in-school and
out-of-school behavior was collected over a four year period (1955-59). Comparisons of the treatment and control group revealed that the impact of the service on the girls had been negligible. Few statistically significant differences were found leading the researchers to conclude that the "limited demonstration of effectiveness raises important questions of appropriate goals of service programs as well as issues about social work practices and its evaluation."28

Another effort to study the effectiveness of a program designed to prevent potentially problematic behavior was conducted by Ahlstrom and Havighurst.29 All seventh grade boys in inner city schools in Kansas City were screened to arrive at a target group of 13 and 14 year old "socially and educationally maladjusted" boys who look like "loosers." One half were randomly assigned to receive special attention through participation in a half-day program. The control boys participated in the regular school program. Data were collected on school, work, community and family adjustment on all the boys through the age of eighteen and nineteen. The independent variable, "work experience" included data on three different developmental stages in required work skills, supervision, etc. It was found that only about one fourth of the boys "seemed to profit" from the work/study experience. The authors concluded that their findings failed to demonstrate that supervised work experiences could materially reduce delinquency "among youth so disposed."30

Reckless and Dinitz31 report on another research demonstration project aimed at preventing anticipated delinquent behavior. Teachers and principals of 44 elementary schools in the inner city of Columbus, Ohio provided ratings of boys who were randomly assigned to receive specialized instruction with a strong emphasis on "role model" building; a control group of vulnerables who were randomly assigned to receive regular classroom instruction; and a group of non-vulnerable (good-boys) who served as another control group. The boys were compared on the basis of psychological tests, value orientations, attitudes, conduct records, school performance and encounters with the police.

It was found that there was no statistically significant differences between the experimental and the control group of vulnerable boys. No differences were found in the number of police contacts, in the seriousness of offenses or in any of the school attendance, drop-out, and achievement indicators. Police contacts tended to increase with age where it was found that by the end of the tenth grade almost one-half of the nominated "bad" boys in both experimental and control groups were known to police. School performance also deteriorated with age. The comparison "good-boys" on the other hand remained somewhat insulated over this time period. In spite of these results, it was suggested that perhaps the role model lesson provided the experimental group was not sufficiently intense to insure the desired "internalization."32

A number of other programs have been conducted without using such experimental designs where measures of success have not been as carefully measured and control yet, experiences gained in many of the programs are widely publicized after acting as a stimulus for initiating other similar programs.

The Passaic Children's Bureau33 was established in 1937 to consolidate the facilities of the public school and police department to reach out to all cases
involving "bothersome behavior" who had been subject to complaints by police, teachers, social workers, and others concerned with youths. The bureau staff was composed of a director from the school's guidance department, counselors, attendance officers, social workers, psychologists and police officers. Based on statutory powers giving broad attendance officers in the school broad investigatory and disposition powers, a variety of incorrigibles and truants were provided psychological, psychiatric and other forms of social treatment. Kvaraceus in a limited study of police arrests of Bureau clients, found only fifteen percent of the boys and who had reached the age of sixteen to have been arrested at least once. Although the percentage appears small, it reveals very little concerning the effectiveness of the program. Similarly, other data showing declines in juvenile court cases and reductions in juveniles committed to institutions cannot be held to be conclusive insofar as there were no adequate research controls on intervening variables.

The St. Paul Experiment in Child Welfare was set up in 1937 by the United States Childrens Bureau to study the problem of identifying and treating children with personality and behavior problems, and to study ways of integrating children into the community at large. A team consisting of a psychiatrist, psychologist and social worker was organized to locate children with incipient behavior and personality difficulties in a high delinquency area. Based on contacts with schools, police and social agencies, children were recruited into the program and provided services which included intensive psychiatric help, social casework and group work and community organization and a variety of referrals to foster care, health care, etc.

After six years of operation, it was determined that 739 children had received group services; 727 individual services not including an unstated number of referrals. Effectiveness was judged by staff members on termination of the project. It was found that "major improvements" were seen in 18 percent of the children. Sixty five percent made what was judged to be "partial improvement" and seventeen percent made "no improvement."36

The New York City Youth Board, established in 1947, was hailed as one of the first comprehensive state-wide delinquency prevention programs undertaken. Basically, the program sought first to identify children who had behavior and personality problems to make appropriate referrals for needed services and second, to expand treatment facilities to more effectively meet the needs of these youths. The program was confined to the areas in the City of New York with the highest delinquency rates. Through an intricate system of coordinating referrals, through the City Department of Welfare and Education, and "contracting" with as many as twenty different voluntary family, youth and special treatment agencies, new services were generated to reach out to potential delinquents. Although this program represented a bold venture into new practice areas, research was limited. In assessing results, however, it was acknowledged that there were great variations in outcomes depending upon the "family's capacity for change" and the "skill of the worker" involved.38

A number of the prevention projects have centered around the public school as the more suitable focal point for working with potential delinquents.

The "600" and "700" special schools in New York City and the Montefiore and Mosely Schools in Chicago serve as examples of specialized schools aimed
at "normalizing" problem students through individualized instruction, scaling down expectations and lessening competitive situations. Results have not been conclusive although reports indicate that there has been some lessening of truancy and teacher-student conflicts. Still to be tested are a variety of different school related programs such as the all-day neighborhood school, community school and other special work/study programs.

A school program known at Higher Horizons was begun in a junior high school in Manhattan in 1956 which has since been expanded to over 76 public schools. The focus of the program was on "disadvantaged children" who were provided specialized counseling and guidance, special instruction and remedial services to broaden cultural experiences. There has been evidence of improved scholastic achievement although there was little to show how delinquent behavior was deterred by the services.

Glen Wallace described a research program which used a more rigorous design in the experimentation of intensive planning, coordination of guidance and counseling services in the Tulsa School System to "alleviate behavior problems" of children. The study found that there were no statistically significant differences in the efficacy of interagency coordination regarding attendance, nature of school offenses and court referrals of children although there were increased grade point averages achieved by participants in the program.

The Maximum Benefits Project in Washington consisted of a special program of social casework which included a variety of supportive psychological and home treatment services for children with severe school problems in two elementary schools. During the years 1954-57 a group of 179 children were referred, evaluated and treated. These were matched with an untreated control group by age, sex, race and Glueck scores. Follow-up data, based on police and juvenile court records, were described as being "far from encouraging" as no significant differences between the treated and untreated groups were found. Sixty-nine percent of the 104 treated children were determined to be delinquent compared with 63 percent of the untreated group. The authors concluded that social casework would be of little value to potentially delinquent youths.

The Seattle Atlantic Street Center, supported by a U.S. Public Health Services Research Grant, did a study conducted to test the effectiveness of social work with "acting out youth" in the school and/or community. Based on the predictions of sixth grade teachers, children were provided with social work services over a two year period. Fifty-four boys who were served in the program were compared to a control group on school records, and on a "severity scale" to measure the degree of offenses committed. No significant differences were found in the frequency or severity of offenses between the two groups.

And, a final approach to preventing delinquency through work within the school was evidenced in the relatively new concept of the police-school liaison programs which sought to create better understanding between police and youth. The program originated in Michigan where police officers were assigned to work along with specific schools to act both as a law enforcement officer and resource counselor. An evaluation by Wierman compared a group of students in a school
with a liaison officer with a group without an officer. Whereas negative attitudes towards police were found to increase in the group where there was no officer, it could not be shown, on the other hand, that there was an improvement in attitudes towards police in the experimental school.

A number of programs have also been introduced over the years to reach out into the community and work directly with youth in their natural groups or gangs, to prevent anti-social delinquent behavior.

Perhaps the best known gang out-reach worker projects were identified with the Chicago Area Project and the New York City Youth Board work discussed above. Based on the theory that gang members were not likely to be attracted to conventional types of building-centered programs, these projects set out to use neighborhood street corner contacts to establish relationships, develop trust and provide direct help to troubled youth in their natural groups. For the most part, the program relied on the concept that group structures could be modified, group values changed and behavior redirected. Unfortunately, there was little experimental evidence to support the otherwise positive findings of work and program planners who were immersed in the day-to-day provision of services.

Walter Miller, an anthropologist, reports on what he calls "the corner group method" of preventing gang delinquency which was used in the Boston Delinquency Project. The focus of the program was on restructuring gangs into more or less formally organized clubs to help shift the members from law-violating behavior to law-abiding behavior. The project was conducted in three neighborhoods and involved 193 boys ranging in ages from 12 to 18 years who were members of corner groups. Seven trained group workers worked with seven such groups intensively and seven other groups less intensively. Data were collected from worker recordings and official records including information regarding court appearances, group law-violating behavior, and changing relationships between group members. The finding showed that there was a definite impact on patterns of group behavior which resulted in a decrease of law-violating behavior especially during the early phases of the project. The lack of control groups, however, make it difficult to assess the program adequately.

Another gang out-reach project which relied heavily on social group work as the Hyde Park Youth Project conducted in Chicago. Groups of teenagers who were deemed to be potential delinquents by the Hyde Park Center were identified. Three street club workers served eleven street clubs during the three year period in which evaluation took place. Data from police and court records, community institutions report and worker impressions were combined into an index to show degrees of anti-social behavior before and after service. Of a total of 326 youths in the program, reduction in anti-social behavior was observed in 151 or 40 percent. Ten percent were found to have increased their delinquent activities. Of 156 youths found not to be involved in anti-social activities, almost half were so classified from the earliest contacts. It was concluded that there was "little increase" and "some decrease" in the frequency of anti-social behavior by the youths. Again, the lack of control group data constitutes a drawback in interpreting the findings.
Henry Street Settlement House in New York City sought to prevent delinquency through early identification and treatment of pre-delinquency gangs. The program engaged parents as well as youth in joint problem-solving activities within the settlement. Experience with five delinquent gangs and parents seemed to show "positive effects" in behavior although no hard data, based on experimentally controlled observations, were offered to support the conclusions. Nevertheless, it was reported that the key to checking pre-delinquency in pre-adolescent groups lies in helping to reinstate parental influences where it rightfully belongs and in building stronger ties between children and their parents.

In an exploratory study of delinquent subcultures by Irving Spergel, it was shown how "opportunity structures" within a neighborhood influenced the kind of delinquent subcultures which prevailed in a neighborhood. However, although a later detailed presentation concerning the theory and practice of street gang work was presented, there is little solid evidence presented to show how the intervention of a gang worker, in fact, reduced delinquency as it is nurtured and furthered through gang affiliations.

C. Reducing Recidivism and Rehabilitation Delinquents

Typically, the bulk of the delinquency prevention programs aimed at treating delinquents to prevent further delinquency are found in the juvenile justice systems as they are found to operate throughout the private mental health and child guidance clinics.

The earliest psychiatric clinic for treating delinquents was set up in the Chicago Juvenile Court by Dr. William Healy in 1909. Most of the programs soon to follow sought to deal with delinquency as a problem within the child's personality. Therapy was to consist of clinical services aimed at treating behavioral disorders. The efforts of psychiatrists, psychologists, social workers, teachers and other professionals were coordinated through staff conferences designed to assess, diagnose and formulate treatment plans based on mental hygiene principles.

Whereas the early efforts were aimed primarily at coordinating other treatment resources and environmental manipulation, the later trend was towards providing more direct psychiatric treatment for delinquents. A common feature of the program was the emphasis on modifying parent-child relations.

A number of research evaluation studies were conducted at the Judge Baker Clinic in Boston which was one of the best known, well-staffed treatment programs. In early years, the program relied heavily on study, diagnosis and served as
resource to make recommendations to the court. In a study conducted by Sheldon and Eleanor Glueck, based on the records of 1,000 boys who had been referred by the Juvenile Court to the Judge Baker Clinic, it was found that five years after being examined and diagnosed, 88 percent of the boys had continued in their non-conforming delinquent behavior. Seventy percent had been convicted of serious offenses while about one-third of the boys had been arrested four or more times. The research staff of the clinic compared court records of another group of 1,000 boys who were before the juvenile court at the same time but who were not examined in the clinic. The comparison showed that both groups had similar rates of recidivism.

Shifting away from the study-recommendations approach, the focus of prevention was placed on direct intervention with the delinquents and their parents. Healy and Broner conducted a study of treatment effectiveness in Boston, Detroit, and New Haven. One hundred and five delinquent boys were compared to 105 non-delinquent siblings along with 40 other unmatched pairs. It was found that about one-half of the boys treated had not engaged in misconduct during the ten year period following treatment. While this was significant, there were great differences in the nature of the success. For example, almost none of the more seriously disturbed responded favorably to treatment. Those with pathological social circumstances seemed to fare better than expected while the greatest successes were with the more average maladjusted boys. The findings were influenced somewhat by the fact that the programs studies were above average compared to the more typical child guidance clinics.

Based on these findings, the Judge Baker Clinic modified their intake policy in efforts to refine and make treatment more effective. But, subsequent studies continued to show ineffective results with the more seriously disturbed and with social problems which were more extreme. However, it was recognized that services were extended primarily to families were treatment if desired and motivation for help influences participation. Hence, the programs have been criticized in that they tend to serve clients who are not representative of the total delinquent population.

The Highfields project in New Jersey represented a pioneer project in the use of guided group interaction as a major method for the rehabilitation of delinquents. The program consists of residential treatment aimed at a population of approximately 20 boys, adjudicated delinquents, aged 16 and 17, who were referred to the juvenile court. The boys remain in residence for from three to four months. Days are spent in working at a mental institution and evenings are devoted to the group counselling sessions.

At least two evaluation studies have been conducted on the Highfields project. McCorkle et al. found that in comparison to general parole violation, statistics show there was a better adjustment of the Highfield boys compared to other parolees, over a five year period. Only eighteen percent of the Highfield boys had violated parole compared to 33 percent of a control groups. Ashley Weeks compared the Highfields boys with a group of boys who had been sent to a traditional reformatory delinquents. The experimental control groups were not, however, randomly determined but were more accurately follow-up studies on placements determined by the juvenile court judges. It was found that of the total 229 boys
to Highfields, 60 percent, as compared to 47 percent of the reformatory boys, completed their treatment and did not get into serious trouble so as to be re-institutionalized. The higher percentage of success found among Negro Highfield boys in interpreted to counteract criticism that the Highfields projects tends to be sent "better prospects for success." Because of the question regarding adequate controls of variables, these evaluative studies have been subject to dispute.65

The Provo (pinehills) Experiment stimulating by the apparent success of the Highfields Project in the rehabilitation of delinquents was begun in Provo, Utah in 1956 by a group of professional and lay volunteers. Funded by the Ford Foundation, the program sought to treat habitual offenders, 15 to 17 years old, assigned by the juvenile court. In contrast to the residential program at Highfields, in the Provo experiment the boys continued to live in their homes and were free to participate in community activities throughout their treatment which lasted from four to seven months. The major form of treatment during the early stages and community adjustment at later stages.

The Pinehill experiment consisted of a comparison of program participants in two control groups - one of boys and probation and another of boys committed to a regular training school. Whereas initially the design called for random selection from a common population of repeated offenders, the methods were not rigidly followed. Findings showed lesser percentages of rearrested within six months compared to 77 percent of those on probation and 42 percent of those in the State School. It should be noted, however, that there was relatively little difference between the Pinehill group and those who were placed on probation.

Other short term residential treatment programs aimed at rehabilitation have been conducted by the California Youth Authority. In the Freemont Experiment for example, small group treatment for example, small group treatment and work experiences were used in conjunction with home visits to try to prevent recidivism. A group of Freemont Boys were compared with a control group of regular institutionalized delinquents. In a two year period follow-up study it was determined that no significant differences were found between those receiving specialized treatment and those in the regular program. In another similar study, the Fricot Rand training school's effectiveness was examined by comparing resident delinquents who were divided into experimental and control groups, the experimental group being "more often exposed" to peer group and staff member contacts. Although the experimental group tended to be less troublesome for longer periods of time, it was found that 80 percent of both treatment and control groups were in trouble some time during the three year follow-up period studies.

The California Youth Authority's controlled experiments in correction offer other illustrations of the use of research in measuring "community treatment" effectiveness. Based on careful screening procedures, boys and girls were either randomly assigned to community treatment or referred for correctional institutional treatment. Treatment consisted of group counselling, therapy, family therapy, and school tutoring, etc. A repetition of psychological tests were used as one indicator of successful treatment along with a measure of "failure rates" determined by re-institutionalism after release. It was reported, on a fifteen month follow-up
that only 28 percent of those paroled from the experimental group had been revoked compared to 52 percent of the control group.

Later extensions of the program into Watts and Oakland—high delinquent areas—were similarly evaluated but with less rigorous research methodology inasmuch as random assignments were not made to control groups for comparison purposes. After fifteen months exposure, 39 percent of the project participants were subject to parole revocations compared to 48 percent for what the equivalent state-wide rates for youth of similar ages.

IV SOME POTENTIAL USES OF BEHAVIORAL SCIENCE KNOWLEDGE TO INFLUENCE LEGISLATION

A look at some of the common themes which seem to underlie the delinquency prevention programs suggests some problems and issues arise in contemplating any legislation. In the first place, it is clear that programs vary greatly in philosophy, programming goals as well as in operational styles. Even where some share common preventive orientation, no standardized prescription exist and no single treatment method, however defined, seems to be applied throughout the programs. The complex social realities of differing environments allow for too many social variables to enter the picture and there is no way to control these given the state of our knowledge. In addition, the failure to integrate and coordinate desperate approaches is significant in that it reflects the tendency for agencies to function in relative isolation from each other, offering services based on differing criteria which may or may not contribute to the overall solutions.

To be sure, there is no reason to recommend that a standard approach be adopted. On the contrary, the diffuse nature of the delinquency problem would imply the need for multi-facted approaches. But, the separatism does suggest that even if answers were found, it is not likely that new knowledge would be evenly applied nor would programs change significantly given the nature of their differing definitions of the problem, target population, and intervention techniques. What guidance can be given regarding the matter of problem definition? Can social sciences play a role in developing strategies for concerted action? Insofar as definitions are crucial to all subsequent activities, the definition of delinquency constitutes one of the major issues to be addressed by the social sciences.

That the definitional dilemma is real, is perhaps best illustrated when one tries to think about delinquent acts in light definitions which apply not only to criminal offenses but to the so-called status or juvenile offenses as well, it has been estimated that as much as 90 percent of all young people have committed at least one act which could have brought them before juvenile court. A careful analysis of 2,490 "self-reported" delinquent acts by 522 teenagers in Flint, Michigan by Martin Gold revealed that only seventeen percent of the youths reported that they had committed no chargeable offense within the past three years. Hence, by definition, about 83 percent of the sample were confessed delinquents. Generalizing to all of Flint's 13-16 year olds, Gold estimated that somewhere between 10,560 and 11,220 "delinquents" were walking the streets at the time of the study. Given our added concern for those who might just become delinquent, one begins to wonder whether there would be any youngsters left over who would not be subject to conscious prevention intervention programs.
A second major question emerges in relation to the stated program goals. As was suggested earlier, social science data may not only be useful in helping shed knowledge on the nature of delinquency (i.e. characteristics of the actor in his social context) but may also be employed to examine the social control agencies as they function and interact with delinquents. For the most part this potential seems to be largely ignored in the projects reviewed.

The target for change in most cases is either the delinquent, his family, the gang, neighborhood or community. While these certainly constitute appropriate points of intervention, there is no reason why social control institutions such as the police, courts, schools cannot also be included among the targets for change. But, evaluation is aimed at determining the efficacy of one or another program, the subject of study is such that it takes for granted that social control agencies cannot but result in anything except "positive" outcomes. Few study the unanticipated consequences which might show how, for example, police judgments "create" delinquency, how juvenile court decisions add to some juvenile problems, how the school actually encourages truancy, how correctional programs help maintain deviant identities and so on. It would appear that social policies may be enlightened considerably should there be more data showing the role these institutions play in the dynamics leading to deviant behavior.

Much of what has been said gains support from the current attention in the social sciences which is given to the "labeling" perspective on deviance and social control.

Increasingly, it is recognized that deviant behavior is not exclusively the result of disruptive and disorganized forces in society but reflect the normal social processes of control which account for conventional behavior, as well. Whereas sociological attention has been given to understanding the social order to discover the structural conditions which encourage deviance, the labeling theorists have been more impressed by the social processes which define, label and articulate negative societal reactions which, in turn, contribute to the deviant response. Deviancy in this perspective is not an attribute or quality in any individual but rather "a consequence of the application by others of rules and sanctions". According to the labeling theorists, deviant behavior is behavior people so label.

John Lofland speaks of "pivotal" categories which emerge as short hand ways for people to impute significance to behavior. Categories such as "delinquent" are being constantly revised and re-created through the efforts of "moral entrepreneurs" and "imputational specialists" - otherwise known as behavioral scientists. The interesting fact is that the public's readiness to accept or reject definitions may be materially influenced by these specialists. Yet, as the number of imputational specialists increases there is a corresponding increase in those who are imputed to be deviant. Social workers, psychologists, psychiatrists, police, teachers, etc. become proficient in the imputation of delinquency and tend to "ensure the flow of persons" defined as such.

Unanticipated consequences of increased sensitivity may be illustrated by the enlightened judge who in recognizing the importance of rehabilitation refers more rather than less children to training schools. Similarly, as one study has shown, a juvenile police department composed of officers with more...
advanced training and academic degrees more often considered delinquents as "problem-children" amenable to treatment and thus, had higher rates of juvenile arrests than a department where there was less "sensitivity" and accordingly less delinquency "crimes".

Enough has been said to make the point that the success of defining, identifying, labeling and classifying cannot be taken for granted any longer. Yet, studies have not produced any substantial evidence that our programs can, in fact, be more successful than what might occur by chance. Some have suggested in this regard that the stigma and labeling influences may be so important they should, under certain conditions, take precedence in decision making. For example, Wheeler et al. state

Social policy formation...may not be able to wait for the results of such research and it is necessary to formulate a position on these issues without the carefully gathered and assessed data that would support a more clearcut choice for one or another alternative. The choice seems clear; in the absence of evidence on the beneficial effects of official contacts, every effort should be made to avoid the use of a formal sanctioning system and particularly the official pronouncement of delinquency. Such a position is justified on grounds of the potentially damaging effects of the labeling process. The primary reason for use of the official sanctions should be the seriousness of the conduct and its potential damage to the community.

A third problem area which flows from the issues regarding problem definition, and labeling, centers on the somewhat constructed concept of intervention as evidenced in the preventative programs reviewed. Admittedly, programs vary considerably in seeking to remedy problems. Yet, in one sense at least, the alternative solutions seem to be fairly parochial, unimaginative and reflective of typical solutions based on what is familiar rather than what may be theoretically sound but practically difficult inasmuch as few "models" exist to emulate.

To illustrate, Cohen's theory of delinquent subcultures states that much of the delinquent behavior serves as a kind of solution to problems of adjustment which stem from the press of middle class values on youths who are suffering from status deprivation. The theory appears to go untested inasmuch as there is no delinquency prevention models that bring representatives from the middle-class together with delinquents in conscious ways so as to modify the arrangement. Another illustration stems from the notion of the generation gap and the degree to which delinquent responses may represent efforts to break with the prolonged childhood status. Few programs have been developed to bring adults other than parents together with adolescents to deal with intergenerational struggles more constructively. Then again, absent in most all of the approaches is any attention to the subjective experiences of the delinquents, themselves. A reading of the programs seems to treat the delinquent as a passive object subject to manipulation and change. This, in itself, tells a lot about the ways our program continues to overlook fundamental realities and truths regarding the importance of interpersonal relationships in the solution of social problems.
The problem regarding definitions, labeling and treatment have not been dealt with effectively in the research reported. It is to be noted that generally where rigorous research methods have been employed, as in the experimental research designs, "verifiable" data produced discouraging results. Studier repeatedly demonstrated time and again that there "were no statistical differences" between experimental or control groups, or that no differences were found in the amount of severity of the offenses between those who received the treatment and those who did not.

On the other hand, in many of the programs where such research controls were not imposed, the findings were generally more encouraging. When left to the judgements of the staffs, the program, planners, estimates of success such as exemplified in assessments such as "significant improvements", marked progress in behavior was offered as evidence that the method did, in fact, have an impact on the participants.

Obviously, the determination of success had to rest on some agreed upon indicators either of reduced delinquency, recidivism, better functionings, etc. A range of instruments adopted in the studies discussed undoubtedly were designed to get at these data. And, many of the findings rest on the ratings of success and failures in these terms. But, the role of the various labeling instrumentalities such as police discretion, citizen complaints, predispositions towards visible and non-visible deviant acts, are examples of some of the factors neglected. Differing and changing reactions, varying efforts to enforce or neglect to enforce or otherwise deal with behavior cannot but influence these indicators. To rely on a statistical record in light of these intervening variables case serious doubts on the utility of such indicators. Again, as the more or less mechanical manifestations of delinquency to not only the changes in individuals in questions but also to the identifying agents - schools, police, community - who play a role in determining whether a person becomes a delinquency statistic. Of the studies reviewed, a few, if any sought to control for such intervening variables.

In any event, the issue here is what constitutes successful intervention? Is success to be measured in terms of "hard objective" data based on less than perfect indicators or will success in the final analysis have to be determined by how people feel about the problem? What rationale is there for showing that the researchers who rely so heavily on admittedly difficult to substantiate data can be found to be anymore accurate than the workers who are so immersed with the clients? One, of course, may point to the self-serving interest of the workers who claim success. But there is little to suggest, at least from this review, that the researchers are any less identified with demonstrating success or failure, although the methods of research obviously serve to check bias. Surely, massive intervention programs cannot be based simply on clinical judgements anymore than they can be based on the findings of researchers who have limited access to controlling intervening variables. The measures of success to be employed must remain an important item on the agenda for legislative consideration.

In summary, it has been stated that at least four major problem areas or issues seem to play a part in determining direction for future development.
These concerns were evidenced in 1.) the need for more accurate problem identification, 2.) the need for more careful attention to the role played by social control agencies in labeling delinquent behavior, 3.) the need for re-examination of concepts of intervention and 4.) the need to balance and weigh the relative contribution of practitioners and researchers in the determination of appropriate directions for future police development.

These will be considered further in relation to legislative concerns and directions.

B. Re-formulating the Delinquency Problem

What are the implications of changing definitions in legislation. In the first place, there may be a definite role in legislation regarding social policies as to what should constitute delinquency. One direction may be to marshal evidence to show that many of the behaviors now labeled as "delinquent" may be "normalized" through redefinition. The most obvious example would be seen in efforts to eliminate juvenile status crimes from the jurisdiction of the jurisdiction of the juvenile courts.

There is considerable evidence which shows how status offenses occupy much attention of the juvenile justice system. It was reported, for example, that one out of every four children - 20 percent of the boys and 15 percent of the girls - appearing before the juvenile court were there for the so-called status crimes of ungovernability, incorrigibility, truancy, running-away offenses which are "crimes" not for adults but only for children. A national census of the almost 60,000 youths incarcerated in the 722 institutions throughout the country in 1973 revealed that one out of every four boys and three out of every four girls were detained because of status offenses.

Social data may be used to influence legislation in this issue as many kinds of inquiries may be pursued. One can examine the stigmatizing effects of being labeled deviant relative to the "social" harm of the acts. One might note how ungovernable and incorrigible behavior is by definition an interactional phenomenon where the target for intervention would more properly require attention to the social unit such as family or parent-child, teacher-pupil diads. Another area which could be examined would note the significance of adolescent "acting out" behavior which may, in fact, be a sign of health rather than disturbance in that the developmental task of adolescence can be interpreted as a period when normal outbursts occur which functionally prepare adolescents for adulthood. One can point to the nature of the social system which reluctantly opens positions, statuses and roles for adolescents. Strategies may be developed to modify the social system to more effectively accommodate the needs of potential adults rather than focus on revising the needs of potential delinquents. And finally, such an inquiry could point out the need for new interpretations of the role of social control agencies. Should the juvenile court be a social agency treating psychosocial needs or can its role be narrowed to treating youths whose behavior was clearly criminal? If police do not intervene into the family, what other facilities such as school, public social agencies, private agencies can play a role in re-defining and serving the truant, runaways?
The role of the social scientist in directing legislative attention may prove to be controversial. Consider, for example, the testimony and congressional discussions surrounding two legislation proposals: Senator Mondale's attempts to create a stronger alliance between social scientists and policy makers through the creation of a Council of Social Advisors which would act as a counterpart of the Council of Economic Advisors; and Senator Harris' proposal for the creation of the National Foundation for the Social Sciences to utilize social science knowledge in dealing with highly visible and social problems, to offer solutions, and to mobilize funds, manpower and facilities for their solution.

As Senator Mondale pointed out, there is a certain historical mistrust on the part of many lawmakers regarding social sciences. The attitude is partly because of unfamiliarity, poor communication or perhaps, because lawmakers like to think of themselves as successful practitioners of applied social science as witnessed in their ability to get elected to office. Many see behavioral science study as a trend toward the society of Orwell's 1984 and are wary of invasions of privacy in social research, and fearful that data banks will be abused and render the individual vulnerable.

But, he also indicated that if the social sciences had not been developed the necessary sophistication to fully participate in determining social policy, then they "must - and very soon" for government at all levels is going to seek advice and value judgements which is a responsibility which may be forced on the social sciences. The council of Social Advisors, argued Mondale, would facilitate the transformation of social scientists into activists, from observers to participants.

Similarly, in urging the creation of a National Foundation for the Social Sciences, Senator Harris pointed out the need to enhance the status of social science by legislative mandate and to acquire Congressional support to "Conduct innovative and sometimes controversial research." As he states:

Social scientists should help to design as well as evaluate programs and to sell them to the people. Overall, there is a continuum from knowledge to power with active roles for social scientists to play along this entire spectrum. But if we are to get the political consensus necessary for real change, the people also must gain that knowledge and exercise some power.

As a matter of practical politics, the passage of legislation requires a constituency, inasmuch as most laws grow out of a need that has immediacy and relevance for a sizeable portion of the population. Most proposed legislation has some kind of constituency urging passage. To return to our example, any social science support of legislation to eliminate status offenses from the statutes is likely to be interpreted as a threat to the traditional authority fo the parents who are traditionally backed by the law in their efforts to control their offspring. Revising the arrangements may raise outcries although it is not altogether clear that some modification is parent-child relationships power relationships would not be acceptable.

For example, in a study of community reactions to parental authority done by Cohen, Robson, and Bates, it was found that the majority in the community
would favor greater legal restrictions on parental authority over the child than the law presently required. The community would have the law grant more legally enforceable claims to pre-adolescent children than it now permitted. Whereas the law by and large, does not distinguish between the preadolescent and adolescent in granting autonomy, the community did recognize a need to have the law accord increasing degrees of autonomy recognized in terms of increased responsibilities as well as rights at the age of adolescence.

C. Assessing Delinquency Prevention Programs

In addition to the use of behavioral science knowledge in assessing human behavior and social institutions and in re-defining the problems of delinquency, it seems likely that social science data would still be sought to assess delinquency prevention programs and give guidance to legislation. Given the limited nature of existing knowledge in this area, what would be an appropriate role?

Certainly, the need for more and better research would be high on the agenda. Few proposals would be supported unless attention would also be given to including some "research evaluation" component. Moreover, it would be urged that every effort be made to clarify objectives, procedures, explicate methods of intervention, provide adequate and accurate reporting, etc. But in addition to these efforts, there will have to be assessment and interpretation of program proposals where social science data are simply non-existent or unavailable.

One way to approach the problem would be to make certain that knowledge from a variety of sources be used. A paradigm might blend theory, research and experiential knowledge to be applied on a case by case basis. In other words, it would seem that not proposal however conceived can be established on a sound basis without paying attention to 1) theory - the theoretical and conceptual framework underlying the proposal; 2) research - the way it takes into account and builds on previous research; and 3) practice - how it draws from or reacts to previous practice knowledge and experience. Similarly, it would appear that legislators would be more receptive to arguments which take all three factors into account. Decisions would be influenced by program "concept" (theory) and the possible reactions of the public degree of "Scientific proof" (research) and demonstrated usefulness (practice).

The paradigm, of course, is not new although it has perhaps not been consciously applied in all cases. A few widely divergent proposals will be briefly discussed to illustrate how it might be useful in providing a framework for assessing delinquency prevention potential.

Consider a recent proposal for a preventative experiment by Cortez and Gatti. Based on assessment of delinquency studies and especially the findings on the role of the family in creating delinquency, they propose that pilot projects be set up in different cities where there is the greatest delinquency problem. All families with children under 7 years would be included in the
screening process where Glueck's Special Prediction Scale would be used to identify and predict the potentially delinquent families, not necessarily potentially delinquent children. The thrust of the service would be provided by social workers and psychologists in conjunction with government programs to "develop techniques to tactfully and helpfully inform and train the parents in the necessary modifications of their child-rearing practices and in their relationship with each other. The proposal seeks to strengthen family life relationships, attitudes and behavior not only in the home but in educational centers and clinics attached to school systems and communities, as well. The authors state, "The reader will agree that from every point it is more humane and sound to predict potential delinquency and concentrate our efforts on preventing it rather than wait for a period of years without using preventive measures in order to determine whether or not our predictions were valid."

The proposal may be compared to another by Hutschnecker a number of years ago which received national attention when it was suggested that mass psychological tests be given to detect children who are apt to become antisocial and that special camps be set up to retrain them. Whereas Hutschnecker relies on psychological tests, Cortez and Gatti prefer testing physical somatotypes to "reveal delinquency potential."

Certainly, such proposals if backed by the force of law would raise some serious constitutional issues especially if such tests are used as a basis for removing a child from his home or as a basis for differential treatment. Although Cortez and Gatti's proposal is limited to working with parents who are receptive and cooperative, one can never be certain that hidden social influences would not subject some families to unwelcomed intrusion on basic liberties, family life styles, etc.

But, perhaps the more difficult problem aside from the legal question is evidenced in the impression given that we do have the knowledge and techniques to effectively help these families. What evidence is there to support such a contention? Cortez and Gatti rely on what they see to be the success of Healy and Bronner, the St. Paul Family Centered Unit, Tefferteller's material on revitalizing parent-child relationships, among other studies. They argue that the preferred treatment should seek to strengthen the role of the father, improve discipline and supervision, provide love and support in the home and create cohesiveness and togetherness in the "good family life." The proposal seems to overlook the anticipated negative effects of labeling, potential abuses in the intrusion of the rights of individuals, "guilt by association" when a child's test score becomes a family's problem and the potential neglect of families not desiring help who perhaps maybe more a cause for concern than those families who "are willing."

In terms of our paradigm, such proposals may conceptually fit some of the public impressions regarding the role of the family in the creation of juvenile delinquency, and may have some theoretical support. But the research evidence, other than the claimed successes of the prediction instruments such as the Glueck scale, lacks experimental data to support the intervention method. Although there is some practice knowledge in the studies cited, one would hardly feel confident that there is an adequate treatment that could be applied without considerable refinement.
In another less ambitious suggestion, Wolfgang and his associates provide data based on a birth cohort analysis of 9,945 boys living in Philadelphia from the 10th through 18th years of age. Information was gathered from schools, selective service lists, police and other records. It was found that some thirty-five percent of the boys in the study had been involved with the police at least once. Race and socio-economic status (based on income) were the most significant factors discriminating delinquent from non-delinquent groups. Although there was a greater likelihood that other offenses would follow after a first offense, knowing the nature of the first offense offered little predictive power concerning the character of future offenses. The probability of delinquency increased with age up to sixteen. Non-whites tend to be less "remided" by police intervention but in general the effect of intervention on the cohort was unclear. The authors conclude that it would perhaps be wasteful to intervene prior to the third offense inasmuch as 46 percent of the delinquent acts are not repeated. The target for change in this view would be the individual youth who comprises the largest proportion of chronic offenders. It was found that the chronic offender which comprised 18 percent of the total group of offenders was responsible for over half of the offenses. As the author states, "Because forty-six percent of delinquency stops after the first offense, a majority of expensive treatment programs at this early point would be wasteful." They suggest that intervention be held in abeyance until after the commission of the third offense for an additional 30 percent of the second offenders desist from then on.

At least two problems may be raised. In the first place, little is said about how to go about treating the third time offenders. Presumably we are forced to rely on our out-moded unsuccessful approaches. Secondly, there is in such a recommendation the implicit belief that the public would accept as rational a plan which would allow about half of the offenses to go unattended on the theory that such offenses would stop of their own accord. Even if the treatment in rehabilitative prevention were effective, the delinquency problem would still continue unless we learn to look at the one or two-time offender quite differently than we now do.

Applying the paradigm, the concept although not likely to raise any constitutional equal protection issues, would conceptually call for a major re-orientation to the delinquency problem. While the research data are convincing, there still are no experimental findings to support any plan of intervention. And finally, practice knowledge seems to be taken for granted insofar as the proposal implies that we know how to work with third-time offenders in a less wasteful manner.

A third illustration of a still different kind of study proposal which could conceivably provide some guidance in delinquency was conducted in a small close-knit Italian-American neighborhood in Cleveland, Ohio. Based on a study of all boys between the ages of 12 and 18 living in the neighborhood, the study sought to determine the degree of social class heterogeneity within the neighborhood which influenced deviant behavior. Three major variables were identified and found to be highly correlated: individual values and behavior styles, family types and peer group associations. What was usually

-249-
taken to be a homogenous social unit was seen to consist of interacting social systems, locating the delinquent within the neighborhood maze of sub-system influences could offer some potential for developing specific preventative approaches. In this regard, one could argue for intervening either in the family, individual or peer group.

With respect to the paradigm, the theoretic or conceptual arguments would, although speculative, not appear to be controversial. Certainly, the average person is receptive to the idea that these are the "important" factors. Although the research showed that the variables were associated there was no demonstrated cause and effect relationship nor was any specific intervention technique tested. Social data would not be available to support the merits of one or another treatment. And finally, it would appear that there is much practice wisdom to support a proposal which seeks to deal with such variables although again, much remains to be verified.

V. POSTSCRIPT

When applying behavioral science materials to influence legislation one often thinks in terms of a simplistic model. I would include a statement of the problem, hopefully a fairly sophisticated study process culminating in a final statement of findings recommending legislative change. Then, as logic would have it, such scientific evidence would hardly be ignored by the law makers who are at their best when passing laws which would have the sanction of scientific authority.

But, the process of enacting legislation is anything but logical for it is more likely to reflect the ebb and flow of a series of conflicts and positions lacking in internal consistency and continually sensitive to the varying pressures and competing interest groups. It may start with litigation in a courtroom where a judge scrutinizes the constitutionality of a statute in regard to a single individual. It may constitute a phase where collective negotiations for compromise is sought without any attempt to charge the law. Or there may be campaigns of direct confrontation between proponents and opponents press to get a bill passed or killed. Victory of defeat does not necessarily mean that the war is over. Even when a law is passed, the cycle may only begin again with new litigation over other issues.

Moreover, the desired change is anything but insured even when all recommendations may be finally enacted and apparently accepted into law. In this regard one may make a distinction between the instrumental and symbolic functions of the law.

Governmental acts commit all of us to the public norms of morality. When these laws are enforced and, in fact, control our behavior, then they may be said to be truly instrumental. But some laws have only symbolic value in that they command public affirmation but relatively little compliance and are seldom enforced. There are, for example, "patterned evasions of norms" which are more or less sanctioned ways of evading laws without being punished such as is found in gambling, prostitution, public drunkenness. Such activities in a functional sense act as safety valves minimizing conflicts between cultures.
The law symbolically proclaims public morality on the one hand although the behavior is permitted to persist on the other. This may not incidentally be altogether bad for in this process other institutional forces are often mobilized to contribute to this public affirmation of morality.

One must be prepared to acknowledge these differential functions of the law. For example, is the passage of a fair housing law condemning racial discrimination in housing and employment which might provide significant symbolic functions likely to satisfy or frustrate its backers if enforcement is left to an agency which has no power to enforce the law against offending landlords and employers? Or should the Supreme Court decision in Brown vs. Board of Education be accepted as a victory when it can only be interpreted as symbolic in nature? Indeed, the significance of how the law itself functions should become part of the problem to be examined and assessed.

When the subject matter for legislation is deviance the problem is even more complicated, for the public definition of deviant behavior is itself changeable. It is open to reversals of political power, to twists of public opinion and to the influence of social movements and moral crusades. For what it attached as criminal today may be seen as sick next year and fought over as possibly a legitimate rate by the next generation. And indeed, the lifting of deviant activity to the level of a political public issue can very well be a sign that its moral status is at stake and that legitimacy is a strong possibility. But when community values and attitudes differ strongly social evidence for change is likely to be viewed with suspicion and labeled as anything but scientific.

We would like to think also that somehow problems will be solved where social science recommendations are enacted into law. But, just as it is true in practice that one senses that activities may bring about desirable results they may also bring about unanticipated and sometimes not so desirable consequences for "clients" or for other parts of the system. This implies an obligation to think through with the law-makers all alternatives carefully noting possible consequences. This process is particularly frustrating when one recognizes that our contributions are not very likely to be accepted unaltered into the decision-making process in any event.

These, of course, are just a few of the considerations. Perhaps as we get more experiences in doing more influencing, the day will come when we can do best in relating our knowledge and skills to the legislative process.


3. This framework has been used by a number of authors. See especially Helen L. Witmer and Edith Tufts, The Effectiveness of Delinquency Prevention Programs (Washington, D.C.: U.S. Government Printing Office).


6. Kobrin, op. cit.


13. Trojanowicz, op. cit., p. 204.


16. Ibid.


20. Ibid., p. 78.


22. Ibid., p. 52. During the same period other Boys Club delinquency studies were conducted which revealed similar results. See Witmer and Tufts, op. cit., p. 22.


24. Ibid.


26. Ibid., p. 92.


28. Ibid., p. 207.


30. Ibid., p. 15.


32. Ibid., p. 162. The authors have done considerable work on examining the role of self-concept in the identification, prevention and control of juvenile delinquency. See for example, Walter Reckless, Simon Dinitz and Ellen Murray, "Self-Concept as an Insulator Against Delinquency" 21 American Sociological Review 744 (1956).

33. Arthur Hill, Leonard Millor and Hazel Gabbard, "Schools Face the


38. Alice Overton, "Aggressive Casework" in Furman, ibid., p. 60.


40. The 600 schools deal with severe behavior problems while the 700 schools are designed for children with "records", Lemert, Instead of Court, p. 27.

41. Adele Franklin, "The All-Day Neighborhood Schools," Annals AAPSS, p. 42. There seems to be some evidence of reduced truancy and minimized vandalism and "less delinquency generally" where all-day schools have been tried.


45. Ibid., p. 242.


47. See Trojanowicz, op. cit., p. 207.


51. Walter Miller, "Preventive Work With Street Corner Groups: Boston Delinquency Project," Annals AAPSS.

52. Ibid., p. 103. 'See also Walter Miller, "The Impact of a 'Total Community' Delinquency Project: 10 Social Problems 168 (Fall, 1962).


62. Witmer and Tufts, The Effectiveness ... op. cit. p. 38.


67. Joachim P. Secker, The Fremont Experiment, Assessment of Residential Treatment at a Youth Authority Reception Center (Sacramento, Cal.: Department of the Youth Authority, 1967).

68. See Carl F. Jesness, The Fricot Ranch Study (Sacramento, California: Department of Youth Authority, 1965).


70. Wheeler, ed. Controlling Delinquents, op. cit., p. 317 ff. outlines similar conclusions based on the "relevant data and perspectives on issues of delinquency control" reported in that volume.


81. All jurisdiction allow the courts to intervene to deal with childhood behavior which would not be illegal for adults. For a comparative of delinquency statutes state by state see Irene Solot, Report on State Laws Concerning Detention of Children, Yale Legislative Services (September 1973). Twenty six states have special categories for juveniles variously labeled as a person in need of supervision (PINS) minors in need of supervision (MINS) where there is a special attempt to provide differential treatment to status offenders although in practice there is little difference in the treatment afforded them. See Paul Lerman "Delinquents without Crimes: The American Approach to Juvenile Status Offenses" in Norman K. Denzin, ed. Children and Their Caretakers (New Brunswick, N.J.: Transaction Books, E.P. Dutton and Co., 1973).


84. Mondale, op. cit., p. 216.

85. Ibid., p. 215. It is interesting to see the arguments against the legislative provision of a greater role for the social sciences. For example, the Journal of the American Medical Association voiced the complaint, "The social sciences are themselves so young and their techniques at present so experimental and poorly controlled as to indicate some doubt as to whether or not their development has proceeded sufficiently to warrant incorporation in the proposed agency." That argument was made against the inclusion of the social sciences in the National Science Foundation.
86. Harris, op. cit., p. 15.

87. Ibid., p. 17.

88. Julius Cohen, Reginald Robson and Alan Bates, "Parental Authority, The Community and the Law" in Rita James Simon, ed., The Sociology of Law: Interdisciplinary Readings (Scranton, Pa.: Chandler Publishing Co., 1966). The researchers offer three major reasons for the differences between the community and the law's conception: the differences in the impact of tradition upon law makers and upon the community; the relative lack of pressures exerted upon law makers to signal the need for change; and the inadequacy of prevailing techniques by law makers for ascertaining the moral sense of the community.\textsuperscript{,} pp. 633-4.


91. Ibid., p. 298 ff.

92. The authors note that somatotyping might reveal not only delinquency potential but also "entrepreneurialship" potential. They suggest that more attention and greater amount of training be given to those children who are mesomorphic and potentially aggressive, p. 299.

93. Ibid., p. 291 ff.


95. Marvin E. Wolfgang, Robert M. Figlio and Thorston Sellin, Delinquency in a Birth Cohort (Chicago, Ill.: The University of Chicago Press, 1972). It should be noted that this work was not meant to provide a delinquency prevention proposal as such.

96. Ibid., p. 254.

The demonstration project is becoming a major instrument for social planning. In sponsoring demonstration projects the overall goal is for small scale "pilot" programs which include some form of research to contribute to program change and policy-making (14, 16, 19, 21). It is generally expected that the lessons learned from demonstrations, through the rigours of scientific research, will somehow result in large scale adoption and major shifts in aims, styles and resources, and effectiveness of social service programs. Models or prototypes for future operational programs are tested to determine their effectiveness in meeting states objectives. For this reason, they are undertaken in natural settings which presumably resemble the non-experimental conditions in which such programs might be later introduced.1

This paper will attempt to elaborate the incompatibility between the two primary goals of a demonstration project - evaluating the program and using it as an instrument of change. It will show the manner in which the aim of using the demonstration project as a rational planning instrument is undermined when research and change strategies are pursued simultaneously rather than sequentially. By examining a nationally prominent demonstration - the St. Paul Family Centered Project - we can isolate relevant issues and illustrate how this incompatibility is manifested. Although there is limited generalizability from case studies writings by Marris and Rein (10) as well as Moynihan (11) support the findings of this analysis. The advantage of a detailed examination of a case study is that it provides an empirical and vivid illustration of a phenomenon which appears widespread. Moreover, it yields questions for large-scale research. Therefore, this paper is not meant to describe the history of the Family Centered Project. Rather, this Project provides the empirical data for grappling with the major issues - i.e., the demonstration project as a research and change strategy.

---

1 Although the expressed purpose is to use demonstrations for rational social planning, there are other covert purposes for undertaking such projects. For example, the funding organizations' purposes could include such concerns as: postponing needed action, placating particular constituencies or challenging existing programs without a major concern for supporting data. And the groups who sponsor the demonstration projects may be more committed to the implementation of service programmes to which there is great commitment rather than testing an experimental project.
The Family Centered Project

In November, 1948, Bradley Buell and his colleagues from Community Research Associates in New York conducted a Family Unit Report Systems Study in St. Paul, Minnesota. The original purpose of this study was to use it as a basis for developing a master plan for reorganizing health, welfare, and recreation services in American communities. St. Paul was selected as the sample city for the purpose of data collection and 108 agencies provided information on persons and families receiving service during the particular month. A major finding of this study was that about six percent of the families in St. Paul were using about 50 percent of the community's services. Among these families who monopolized services there seemed to be a high concentration of problems such as dependency, ill health and maladjustment.

In 1952 the Hill Foundation in St. Paul provided a grant of $90,000 for a three-year period to develop a plan for improving the conditions highlighted by the 1948 study. A pilot project was then undertaken in 1954 to implement family centered casework techniques with "multi-problem" families. Treatment was provided by social workers who were loaned to the Family Centered Project by several agencies. This program lasted until 1959 when funds could not be obtained to continue an expanded treatment service.

Unable to continue with the treatment program, the Family Centered Project shifted its focus to a) training staff from various agencies in the use of family centered treatment and b) the development of a reporting system. The aim was for the reporting system to provide accountability by measuring the changing prevalence and incidence of social problems and the success of family treatment provided by workers in different social agencies. The Family Unit Register was ultimately completed in 1967 to provide such accountability. And immediately afterward the Family Centered Project was terminated.

Evaluation of Family Centered Treatment

Many experts in the field of evaluation research seem to agree that program evaluation is often beyond the capabilities of existing methodology (20, 23). Since evaluations of entire programs are so complex and because they have limited generalizability, a more common strategy is variable testing - singling out specific components of the program and testing their effectiveness in meeting

---

2 The data referred to in this paper is in the Family Centered Project Collection, Social Welfare History Archives, University of Minnesota, Minneapolis, Minnesota.
more limited goals. The research of the Family Centered Project included both types of evaluation. This paper will not examine the efforts to evaluate the St. Paul Family Centered Project. Rather, in order to concentrate on the focus of this paper (the incompatibility of simultaneously pursuing research and change) the discussion will be restricted to an examination of the issues involved in evaluating a particular program component - family centered treatment.

The evaluative research design should be determined according to the purpose of the evaluation, the characteristics of the program, and the host of constraints (legal, administrative and ethical). In examining the demonstration project as a research strategy, it is necessary to examine the appropriateness of the research design. In addition, considering the concern for using the demonstration as a change strategy, it is important to identify the manner in which the pursuit of spread and spillover affect the research component.

The Family Centered Project devoted considerable effort to developing movement scales which could measure changes in the functioning of families, not merely changes of individual members in the family. With the family as the focus, the concept of social functioning was chosen as a basis for analyzing change in the following sub-categories: 1) family relationships and family units; 2) individual behavior and adjustment; 3) care and training of children; 4) health conditions and practices; 5) social activities; 6) economic practices; 7) household practices; 8) relationships to the family centered worker; and 9) use of community resources (6, 7, 8). Information on the social functioning of families was extracted from the case records kept by the social workers on the basis of this source of information, independent raters completed a Family Profile which included a seven-point scale for each of the sub-categories of social functioning. On this scale, adequate and inadequate functioning represented the two extremes and marginal functioning was the mid-point.

Through special research on the research instrument - the Family Profile - there was statistical evidence showing that these nine categories did in fact constitute a scale of social functioning. And the procedures involved in completing the Family Profile were found to be reliable (i.e., there was high consensus among people who independently rated the same families on the nine categories of social functioning). This study could not determine whether the recording in the case records was reliable because it only examined the extent of agreement among raters who read the case records. Also, questions pertaining to the validity of their evaluation procedure were not examined. In other words, it could not be concluded that certain changes took place merely because raters were generally in agreement when determining a family's level of functioning. To ensure validity it would have been necessary to compare changes measured by this instrument with some outside criterion.
1. Demonstration projects and exploratory studies. An assumed feature of demonstration projects is a clearly conceptualized treatment method or program variable which can be tested. In reality, this is often not the case among sponsored demonstrations. And in such situations the project is not particularly amenable to experimental study. Instead major attention can be devoted to exploring innovative ideas and approaches which could perhaps ultimately be developed into testable programs. The type of research strategies required by exploratory projects (i.e., programs which lack a clearly stated program and objectives) and demonstration projects differ. Formative research would be appropriate for explorations and summative research for demonstrations (18). Failure to recognize these differences often result in inappropriate research designs which are consequently limited in their use for planning purposes (22).

2. Monitoring the treatment process. Evaluation research is generally most concerned with outcomes which reflect the effectiveness of intervention. Nevertheless, there are several reasons for also monitoring the treatment process or program inputs: 1) to see whether the treatment was actually carried out in the expected manner; 2) to identify those aspects of the treatment which seem to account for the results; and 3) to determine the variations of outcome resulting from the manipulation of certain program variables (e.g., intensity of treatment, and practitioners' level of professional training).

Since the Family Centered Project was mainly concerned with exploring ideas of family centered treatment, it would be expected that the treatment process would have been the major focus of the research efforts. However, this was not the case. The research publications clearly show that the major emphasis was devoted to determining whether change occurred for the families who were receiving family centered treatment (6, 7). The task of developing concepts of family centered treatment was not totally ignored. Although this was not the focus of the research activities, the Project Director carried out this responsibility. On the basis of her meetings with workers and supervisors, she articulated the concepts of the family centered treatment in a widely circulated publication entitled the Casework Notebook (13). This publication did not, however, result in research which attempted to associate outcome with the various components of family centered treatment which had been identified. Instead, the research continued to be primarily concerned with rating the families on the nine categories of social functioning for the purpose of measuring movement which supposedly was the result of treatment.

Demonstration projects are often viewed as missionary enterprises in which there is a high level of enthusiasm among the participants. According to the Project Director, the sense of complete investment made by the practitioner was a stronger factor than casework skill in regard to performance. She claimed that in some instances improvement took place even where the casework was on the slumy side. This illustrates the importance of monitoring the treatment process - partly to determine whether outcomes are attributable to the Hawthorne effect rather than the treatment approach.
It is possible to infer the reasons for the major concern with outcome results even when the experimental variable has not been clearly defined. Demonstration projects are not commonly funded for the purpose of exploring ideas and formulating testable concepts. Rather, they are funded mainly to solve particular problems requiring intervention. The Family Centered Project was funded for the purpose of reducing and preventing dependency. This necessitated data which would show whether the results of treatment actually reduced dependency. In other words, when demonstration projects are funded for the purpose of determining the effectiveness of particular approaches to the solution of major social problems, the research focus is on outcomes, regardless whether the treatment or program variable has been clearly identified and monitored.

3. Outcome measures. The major thrust of the research efforts in the Family Centered Project was the measurement of movement among families receiving family centered treatment. And as indicated earlier, the procedure for determining movement involved "before" and "after" ratings of the families' social functioning.

Success or failure of projects are largely dependent on the criteria selected for making such assessments. Although the rating procedure was shown to be reliable, there was no assurance that the ratings of social functioning were in fact valid. These rating procedures showed general positive change (about 65 percent) among the Project families. However, studies which relied on more tangible measures showed little change. The overall trend for the families seemed to be an initial rise in public assistance costs and then a slight decrease in subsequent years (2). In a study of police contacts, there was a small increase among children whose families were in the Project. However, the increase was somewhat less than the increase for children in the whole community (1).

Thus, the more tangible criteria reflected little change while the rating procedure showed considerable improvement. Yet the studies of movement were widely published and used to convey the success of the treatment approach. This illustrates the importance of critically examining the success criteria to determine whether results are accurate and objective indicators of success.

4. Use of control groups. Since the experimental variable - family centered treatment - was still in the process of growth and could not be clearly distinguished from other interventions which were in common usage, formative research was needed. Such an emphasis did not require a control group because the treatment variable that would be offered only to the experimental group could not be differentiated from the treatment approaches being used with families who might constitute the control group. Nevertheless, the research focused on testing family centered treatment. And there was considerable debate over the inclusion of a control group. What is particularly interesting, however, is the extent to which concerns about spreading family centered treatment influenced the decisions.
The presumed innovativeness and high visibility of demonstration projects attract widespread attention, particularly since project participants appeared at public meetings, presented papers to professional conferences, published papers and monographs, and conducted training sessions. Even before the concepts of family centered treatment were clearly formulated, the impression that it constituted an effective approach was conveyed throughout the United States and in other countries. The goal of spreading the Project's concepts and methods was logically inconsistent with the primary purpose of the demonstration project—i.e., to test the effectiveness of the family centered approach. How do you spread something which is still in the process of being tested and has not yet been proven effective?

The Family Centered Project deliberately attempted to spread its family centered techniques among the local social agencies. This compounded the problem of obtaining a sufficiently rigorous differentiation between the experimental variable and the "traditional" casework practices. In other words, if a control group had been established and if the results showed that their outcomes did not differ greatly from families in the experimental group, this would not necessarily imply that family centered treatment was ineffective. Such results could also suggest that families in both groups received similar treatment because the techniques of family centered treatment were widely transmitted.

As an alternative to establishing a control group, it was decided that families who had considerable contact with several social agencies in the past provided a built-in control group; i.e., there was a type of contrast control by comparing agencies' past experiences with these families (which could be considered as unsuccessful because the family still manifested many problems) with the current experiences under the demonstration project. The weaknesses of such comparisons are obvious. There was no means of determining the nature of the services provided previously and the situations which brought the families to the social agencies earlier may have differed substantially from the current circumstances in the type and severity of the presenting problem.

5. Service orientation. There have been many writings on the conflicts between research and service in evaluating social programs (12). In addition to the administrative constraints which often necessitate compromises in the research design, considerable attention has also been devoted to the struggles between practitioners and researchers.

Collaboration between researchers and practitioners in evaluative research can be mutually beneficial. A logical hypothesis would be that practitioners are more likely to cooperate in research if the instruments have value for practice and if the research activities do not interfere with their normal work patterns. The Family Profile was originally designed as a research instrument. It was developed by extracting categories of social functioning from the case records kept by the practitioners. In addition to satisfying the research requirements, the practitioners found it useful as a diagnostic tool and the supervisors found it helpful in the process of supervision. Data
collection procedures involved a compromise. Caseworkers continued open-ended recording in accordance with specific guidelines. And researchers used this information to quantify levels of social functioning for the purpose of measuring movement of the Project families. The Research Associate of the Family Centered Project said that the reaction of the practitioners to the evaluation efforts ranged from indifference to mild hostility. Although the evaluation was primarily aimed at examining the effects of a particular method it nevertheless constituted a threat to the practitioners (5).

The Demonstration as a Change Agent

Martin Rein suggests that the concern of demonstration projects for continuity (i.e., survival) usually results in abandoning the strategies for promoting spread (duplication of the project elsewhere) and spillover (using the demonstration as a catalyst for other changes (16). This was not the case with the Family Centered Project. While this Project received widespread national and international attention, it was unsuccessful in its attempts to survive. Since this paper is restricted to the pilot project phase when the major focus was on family centered treatment, this analysis will be restricted to the termination on this particular phase of the Family Centered Project.

1. Spread and spillover. The spread of the Family Centered Project's concepts and methods were not fortuitous. Rather, there were deliberate efforts to popularize this demonstration project. Even prior to the 1948 survey, St. Paul was the focus of national attention as the plans and purposes of that survey were described in Bradley Buell's article "Know What the What Is" which appeared in the Survey Mid-monthly, then a popular social welfare journal (4). Soon after the survey, a National Conference on Appraising Family Needs was held in St. Paul to discuss the study's findings. This resulted in further national exposure as 150 leaders in the social welfare field from across the country were invited to attend this conference. One particular finding of the 1948 survey received widespread attention - that six percent of the families in St. Paul were using about 50 percent of the community's service (3). Yet this figure was misleading as it was computed by using the total number of families in St. Paul as the base, not the number of families who were actually receiving service from social agencies. This point was made by Frank Rarig, Jr., then Executive Secretary of the Wilder Foundation in St. Paul:

... the project reveals that out of a total of 20,264 problem families and non-problem families, 6,466 or 32 percent absorbed up to 50 percent of the volume of services in the study month - admittedly a slightly less dramatic statement.(15).
This particular result was used to justify a coordinated program for families considered "multi-problem". In addition to the establishment of the Family Centered Project in St. Paul, the survey's findings were largely instrumental in the emergence of several other related demonstration projects: Winona, Minnesota, focused on economic dependency; Washington County, Maryland, concentrated on physical disability; and San Mateo, California, was mainly concerned with behavioral disorders.

It was during the pilot project phase (1954-59) that concepts and methods of the Project were most widely publicized. It is important to note that this widespread publicity occurred during the demonstration period, not after the methods had been tested and proven effective. The presumed "innovativeness" of the demonstration project created considerable interest in it, involving the dissemination of information about the experimental variable - family centered treatment - even prior to its final evaluation.

In 1962 a survey was conducted to determine the extent of multi-problem family projects in all communities of 100,000 or over in the United States and Canada. Despite the lack of specificity in defining such projects, it was found that in 1943 communities (or 60.1 percent of the communities reporting in the survey) programming for multi-problem families was at some stage of thought, planning or operation (9). Further evidence of the widespread influence of the Family Centered Project is found in Schlesinger's publication which includes an annotated bibliography of 300 items on the multi-problem family (17).

Efforts were made to assist local agencies to incorporate the concepts of family centered treatment into their regular practice. To facilitate the spread of the Project's methods locally, there were: neighborhood meetings, a speaker's bureau, and training institutes for staff from social agencies. The structure of the pilot project facilitated the "osmosis" of the Project's ideas as workers, supervisor and executive directors of the participating agencies shared their experiences with the non-participating staff. However, it was recognized that this osmosis process was difficult to implement because there was a general lack of interest in the Family Centered Project, as staff felt that the Project did not belong to them and it was not pertinent to their regular work. There was also the feeling among some practitioners that their agency was not family centered and so why should they be interested in the Project. It is interesting to note that there was some reluctance among local practitioners to accept the Project's concepts and methods. In fact, it was thought that there was more familiarity with the Family Centered Project and greater use of the Casework Notebook outside St. Paul than locally.

2. Continuity. Demonstrations are generally funded for time-limited periods. Since their survival is generally dependent on tentative funding arrangements, the extent of control over its own program and its ability to influence long-term changes is somewhat limited. Instead, the demonstration project's program and decisions regarding the transformation of its approach into permanent programs
belong to the organizations which have the necessary financial resources to implement such changes.

The funding organization can influence the demonstration project's activities through the disbursement of funds. When the Hill Foundation renewed the funding of the Family Centered Project it included the following tasks which would have to be pursued as a condition of funding: 1) formulating the concepts being used into transmittable form; 2) identifying, describing and defining the special methods used in the Project; 3) specifying criteria for measuring change resulting from treatment and evaluation of treatment outcomes; and 4) developing specific steps for bringing about changes in the community's pattern of organizing its health and social services. These requirements reflected the focus which the Hill Foundation as the funding body expected the demonstration project to pursue. The first three tasks did not necessitate any major change in direction; the treatment and research activities could continue but the lessons learned from these activities would have to be clearly formulated. However, the latter task - developing a community plan for organizing services - represented a concern which according to the Hill Foundation was neglected in the initial phase of the pilot project. Although the initial grant was made for the purpose of developing a community plan for the prevention of maladjustment among multi-problem families, the pilot project was mainly concerned with treating them.

The community plan ultimately proposed included a centralized unit with social workers carrying the major responsibility for casework services which would be monitored for research purposes. Throughout the community another group of social workers in their respective agencies would only provide family centered treatment and they would not be expected to make any major time investment in research. This plan essentially called for an expanded treatment program.

In justifying a centralized unit, it was claimed that carefully documented studies and evaluated experiences supported this type of structure. However, there was actually no real basis for proposing a central unit. There was nothing recorded to justify the necessity of centralization, particularly since the alliance of social agencies in the pilot project was considered a success. Nevertheless, a central core structure was advocated because it would ensure that the experiment would survive. The Ford Foundation refused to fund this proposed plan and the Hill Foundation could not continue to support such a venture. Not being able to implement this plan, the Family Centered Project abandoned its treatment focus. In its revised request to the Hill Foundation, the change in emphasis was expressed as follows:

our program calls for shifting our emphasis from experimentation, research, and demonstration, to the application of our findings of the past three years into the work-a-day practices of our major governmental and voluntary agencies (8).
The Family Centered Project shifted its focus from providing family centered treatment to training staff in the local social agencies to use this approach. This shift occurred largely because funds were not available to continue the treatment program but the Hill Foundation was willing to finance an endeavour involving the training of social workers in the methods of family centered project.

Conclusion

The espoused purpose of demonstration projects is to ultimately influence long-term change in programs and policies. This paper has shown that attempts to simultaneously use the demonstration project as a change instrument can result in inappropriate and less than adequate research designs: focusing on outcome measures when the treatment variable was not yet clearly conceptualized; using "soft" rather than tangible criteria for evaluating treatment outcomes; and not establishing control groups because of deliberate contagion. On the other hand, there are formidable obstacles in using research findings as a basis for program development. In the Family Centered Project, premature results were widely spread to convey the impression that the treatment program was successful. Due to the tentative funding arrangements it was shown that the participants in the demonstration were limited in their efforts to use the project as a change instrument. It was the sponsoring organization that was able to influence the program focus of the demonstration that was able to influence the program focus of the demonstration project. And organizations with sufficient resources to implement programs could determine whether they wish to implement the experimental approach of the demonstration project.

Despite the growing popularity of demonstration projects, there has not been systematic study to determine the extent to which they actually facilitate program or policy changes. In this regard, it would be interesting to determine under what conditions and through what process do demonstrations have maximum leverage in influencing change.

REFERENCES


15) Rarig, Frank Jr. Executive Secretary, Amherst H. Wilder Charity, St. Paul, July 14, 1952.


As new and innovative social welfare programs are being attempted, there has been an increased concern with evaluating the effectiveness of such programs. To what degree is a new program effective? For which kinds of clients is each type of program effective? What elements are crucial in a program which has been judged to be effective? These are just a few of the questions that evaluators would like to answer.

There is a large literature on evaluation research--some of it reporting or reviewing the results of specific evaluations (6, 8, 9, 10, 11) and some of it presenting general discussions, essays or models (1, 2, 3, 4, 5, 6, 7, 8). This literature outlines many important problems and suggests imaginative solutions to a number of those problems. However, in our judgment, some of the literature has viewed all evaluations as alike; insufficient attention has been given to the basic differences between different types of evaluation. The objective of this paper is to clarify this issue.

We argue that there are three distinct types of evaluation that can be articulated on the basis of the intent of the evaluation: (1) The evaluation is a test of some basic but general scientific hypothesis that underlies or motivates a proposed intervention. (2) The general scientific hypothesis is found or assumed to be valid, and the evaluation is an assessment of the effectiveness of a particular program that supposedly embodies the hypothesis. (3) After a particular program has been shown to be effective at some point in time, the evaluation is a continuing assessment of the effectiveness of the program. An awareness of these three types is important because the methods of evaluation are partly dictated by the intent of the evaluation.

These three types will be referred to, for the nonce, as evaluation-as-experiment (testing a scientific hypothesis), evaluation-as-assessment (judging the effectiveness of a particular program), and evaluation-as-monitoring (continuously examining the effectiveness of the program). Each of these will be discussed in turn.
Evaluation-as-Experiment

As indicated above, each type of evaluation attempts to answer a distinct question. The first question is a question of "pure science."

Suppose that X represents some intervention which is thought to have some desired result, Y. Then, the evaluator asks: Is it true that X will, in fact, result in Y? Under what conditions Z will X result in Y? That is, he is interested in whether or not there is a direct and strong causal link between X and Y; and he is further interested in learning what other variables, Z, moderate or strengthen the causal relationship.

A clear example of this is found in medical research on new drugs. Suppose that a physician thinks that a particular new drug might reduce some of the symptoms of, say, schizophrenia. He needs to determine the effectiveness (Y) of giving the drug (X) to different types of patients (Z_1), in different dosages (Z_2), and in different settings (Z_3). Obviously an experiment is the appropriate way to evaluate the drug.

Less clear examples of this are found in the evaluation of social programs. Consider a program directed to dealing with behavior problems (aggressive and rebellious behavior, withdrawal and the like) perceived in some school children. Suppose that a social worker concluded that the problems of these children were manifestations of family problems. Since the social worker could not work with individual families he decided to conduct small group therapy sessions, with, say five or six families (i.e., pairs of parents) in each group; his decision was based on the conception that through discussion with others, each set of parents might gain some insight into their own conflicts and problems and from that insight alter their interaction with and treatment of their children. Thus, the general behavioral proposition that would motivate this approach is that discussions with others who have similar problems leads to insight into one's own problems which, in turn, leads to change in one's interactions with others which, in turn, leads to an alleviation of the behavioral problems in those others.

It may be that group therapy with parents has, in general, little impact on the behavioral problems of offspring. It has some effect but much less than, say, group therapy with the adolescents. Or, alternatively, the behavioral problems of adolescents may co-occur with failures in school--academic and social failures--and further, the behavioral problems are a consequence of the failures. That is, educational activities which permit these adolescents some experiences of success have the consequence of altering their other behavior which was defined as problematic or psychopathologic.

Here, then, are three distinct principles (group therapy with parents, group therapy with adolescents, experiences of success) on which remedial programs could be based; there are many others. These principles or propositions have the status of scientific hypotheses; as such they need to be tested in a rigorous experimental fashion.
An experimental investigation could, of course, be devised whereby families are randomly assigned to one of several treatment groups (no therapy or treatment vs. group therapy with parents vs. group therapy with adolescents vs. experiences of success vs. group therapy with parents and with adolescents, and so on), with the outcome measure being some set of systematic observations on the behavior of the adolescents. And these should be tested in a variety of contexts: (1) varying economic and educational background of parents, (2) type of community, e.g., racially mixed or not, working class or middle class and so on, (3) type of behavioral problems of children.

The point is not to show whether these treatments are effective or not but to obtain some measure of relative effects. Thus, it might be the case that, in general, group therapy has some effect but relative to experiences of success it is of little importance. That is, if children have some rewarding academic or social achievements in school, then this has a greater impact than gaining insight into their own problems.

There are few instances of planned social intervention that have been preceded by careful experiments conducted in a variety of situations. The problems of the day are so pressing that welfare personnel clearly cannot wait until definitive experimental results are obtained; they must create plausible programs and proceed in a trial-and-error fashion. However, at the same time, some experimental investigation should be conducted. The experimental investigation constitutes one type of program evaluation.

**Evaluation-as-Assessment**

The question asked in the first kind of evaluation is a scientific question; this does not imply, however, that the question underlying the second kind of evaluation is unscientific. Rather, the second question is a question in applied science.

The second question assumes that the first, the question in basic science, has been answered affirmatively. That is, it has been reasonably demonstrated that under conditions Z a specified X has, in general, a desired result Y.

Then the questions are the following. For a particular instance of the intervention, i.e., a program of planned change that is operating in some particular community: (1) do the conditions Z exist in the community? (2) does the program embody the specified X, the agents designed to bring the change? (3) are the desired consequences, Y, achieved?

An example from medicine may make this clear. A surgeon may have discovered and reported that for patients of certain ages and constitutions, a particular surgical technique is effective in correcting certain heart defects. Over a period of time this surgical technique gains wide use. Another surgeon might then be interested in evaluating the effectiveness of this surgical procedure in treating patients with these heart defects. The conditions, Z, then, are the characteristics of patients for whom the technique is useful. The change agent, X, is the surgical technique. The desired objective, Y, is the correction of heart defects.
In conducting the evaluation the surgeon might learn that $Y$ is not achieved; a significant number of heart defects are not being corrected. Several possibilities may account for this. The conditions $Z$ do not exist; the surgical technique is being used on inappropriate types of patients. Alternatively, where the operation is attempted, it is not being performed in the specified way; that is, physicians throughout the various hospitals are not performing it in the way it was originally contrived. In our formalism, the specified $X$ is not embodied in the medical care programs. Still a third possibility is that the original research leading to this surgical technique was in error.

There is considerable difference between medical care programs and social welfare programs; one difference, of course, is that social science has not supplied general research findings that clearly dictate one program rather than another. As we have indicated above, social scientists should direct some of their efforts to obtaining this knowledge.

If the necessary basic research had been done, which had demonstrated that some principle is effective in ameliorating some social problem, then programs using that principle could be assessed. For example, suppose that the definitive experiment had been conducted which demonstrated that group therapy with parents and school success of adolescents was significantly more effective than either alone and that group therapy with adolescents did not have much additional impact. And, suppose that this definitive experiment had articulated rather precisely the nature of the group therapy and the experiences of success so that the program could be copied easily in appropriate settings. Finally, suppose that the program was begun in Community A. Now an investigator in Community A wants to evaluate the effectiveness of this program as it is conducted in his particular community. Since the principles underlying the program have been experimentally validated, he is not going to replicate the scientific experiment. Rather he wants to know (i) are the principles and concepts of this program satisfactorily embodied in the particular activities of Community A? and (ii) is there a reduction in the behavioral problems of adolescents?

This investigator may learn that the desired results are not achieved; there is no reduction in the frequency or degree of behavioral problems of adolescents. But because he asked and attempted to answer the first question above, he can suggest a cause for its ineffectiveness; he learns that the program did not involve experiences of success. Academic tasks that the adolescents could excel in were not devised; there were only a few abortive social affairs, none of which resulted in social successes for these young people. Hence, the investigator is able to indicate why the program failed.

Although the first kind of evaluation requires an experiment, this is not necessary in an assessment; this does not mean, however, that it is easier. In the experiment, some outcome (e.g., reduced behavior problems) in experimental groups is compared with that outcome in the
control groups. In the assessment, however, some previously established standard corresponds to the control group; that is, the investigator compares the results obtained or observations made in the evaluation with the specified standard. The conclusion about the effectiveness of a particular program in a particular setting depends on whether or not its results equal or exceed the standard.

Perhaps we have overdrawn the distinction between evaluation-as-experiment and evaluation-as-assessment, at least as far as evaluation research is practiced, but we do believe that this distinction reflects an important problem that has not been sufficiently appreciated by some of those conducting evaluations.

**Evaluation-as-Monitoring**

The second kind of evaluation can merge into or become a third kind: monitoring. Less needs to be said about this third kind; however, it is necessary to keep the second and third distinct--both conceptually and in practice.

Like an assessment and unlike an experiment, monitoring is directed to a particular program rather than to the class of all such programs. The assessment can be viewed as a discrete investigation with a beginning and an ending. On the other hand, the monitoring of a program is a continuous evaluation. It can be likened to the quality control techniques of the industrial engineer.

The assessment of a program rests on the experimental evidence that dictates the principles underlying the program. Thus, logically, assessment follows experiment. Likewise monitoring follows assessment; the monitoring of a program rests on an investigation that revealed that the program was effective during a previous period of time.

Suppose for example, the administrator of a school mental health program in Community A has had his program assessed with the result that it is functioning fairly effectively. The next step he may wish to take is to devise some means of continuously or periodically evaluating its functioning; that is, he begins some procedure of periodically making observations of the program in operation. These observations may lead to immediate changes in its operation. Monitoring--the systematic and periodic observation--is a "servo-mechanism" in the program; it feeds back data which leads to self correction.

**Conclusion**

The evaluation of the effectiveness of social welfare programs is an important and necessary activity. Among the many considerations--political, social and scientific--we have focused on one: the methodological issues of types of evaluations. We have argued that there are three types of evaluation--all of which are necessary. In making an
evaluation, an investigator should know what he is and is not doing; he should see the formal relationship between his own research activities and that of others, which though seemingly similar to his own, may be different in some significant way. If an evaluator can locate his own attempt at evaluation in terms of the structure of evaluation research that has been outlined in this paper, then his attempt will be both more efficient and effective.

References


4. Donabedian, A. Evaluating the Quality of Medical Care, Milbank Memorial Fund Quarterly, 44 (July 1966): 166-203.


INTRODUCTION

The idea that human services programs should be evaluated is accepted by most practitioners and researchers. Discussion is about the technical aspects of this research and about how practitioners can be encouraged to utilize evaluative findings in their everyday practice. Emphasis is placed also on the organizational barriers to this utilization. These ideas and issues are found in a growing literature in the social sciences and, increasingly, in the even faster growing literatures in the professions and in the human services, including the social services, medical and health services, criminal justice and the like.

Here, evaluation research is discussed from a different perspective: as a scientific social reform movement and as an ideology. Attempt is made to offer a beginning analysis of some elements in the evaluation ideology and to focus attention on some of the consequences of this ideology and social movement for staff in human service agencies.

SCIENTIFIC SOCIAL REFORM MOVEMENT: THE CASE OF EVALUATION RESEARCH

Evaluation (or evaluative research) is now accepted as necessary by many in the human services. Until recently this acceptance was often rhetorical; now services are being studied, often at the request of service administrators who are responding to laws, rules and regulations which demand program evaluation. These requests and these rules have resulted from a successful social reform movement which sought to make evaluation central to practice in the human services.

Evaluation can be thought of as a "scientific social reform movement" (Eaton, 1962). A scientific social reform movement is a type of social movement distinguished from other types by the participants and by the social legitimation sought and given to the reformist ideas. Sources of this movement are found in early sociology, social philosophy, and social reform (Caro, 1971). At base and simply put was the notion that "science" could contribute to "making society better." In the recent past, particularly since the War on Poverty programs, evaluation has become a basic idea, one found in several forms.

For the difference between these notions, see Suchman, 1967.
Evaluation research was a goal of those who sought to learn if programs "worked"—legislators and social scientists among others. These people sought to make research, in general, and this approach, specifically, a central process in the funding and administering of human services. To practitioners in the human services, evaluation was a means of learning about their services so as to make these effective. This idea included two notions: the true desire to learn about one's program, and, it seems, the preventive notion that unwillingness to accept this outside monitoring could lead to more severe forms of outside control and accountability.

The word evaluation came to be used by different groups of people to mean different things. The word diffused into many professional vocabularies—but in each vocabulary the meaning of evaluation was different. In common was the idea that evaluation was a "good thing," something which "should be done." Evaluation became part of a social movement which sought to counter critics of the human services at a time when the Great Society was no longer great, when funds for such services became scarce, when the social philosophy and political style of the government became more corporate and efficiency oriented:

No matter which political party dominates the legislative process, or which academic viewpoint is held at any given time, or whether the mood of the country is for increased spending or cutbacks, constant modifications and innovation in human service programs is bound to continue. There is every reason for dissatisfaction with the current state of intervention on problems of health, economic security, education, housing—and indeed on the entire range of social disorders that confront our urban communities. . . .

Numerous limitations surround current efforts at social action: inadequate techniques, scientific knowledge, and manpower are commonly cited examples. . . . But knowing what to do and when to do it requires another tactic as well. Neither the rhetoric of politicians nor the pleas of do-gooders of various persuasions are sufficient to guide program development. Similarly, neither the theories of academic in nor the exaggerated statements of efficacy by practitioners are an adequate basis for the support and expansion of various human service activities.

Evaluation research, not a new but nevertheless an increasingly robust enterprise, can have a major impact on social problems. While it would be foolish to argue that all the deficiencies of current programs or all the political and conceptual problems can be swept away by evaluation studies, the adequate assessment of existing and innovative programs can be a vital force in directing social change and improving the lives and the environments of community members. (Caro, 1971)

EVALUATION AS AN IDEOLOGY

The ideology of evaluation in the scientific reform movement included several notions: That "the evaluation findings can and should be utilized in practice;" that "the findings are appropriate and necessary for management and program development;" that social scientists have the techniques to do evaluative research.
And that policy and administrative decision-making about social policy, program funding and the like can be rationalized, (outside of or within a notion of politics) and become more effective: i.e., social policy and planning using evaluative research as the technology for getting appropriate data would be an a priori "better" way to decide.

Evaluation is a particularly powerful ideology of change because (by its nature) it blurs the distinction between the culture systems of science and ideology (Geertz, 1964). Ideology has the social function of making "an autonomous politics . . . possible by providing authoritative concepts that render it meaningful, the evasive images by means of which it can be grasped." (Ibid.):

And it is, in turn, the attempt of ideologies to render otherwise incomprehensible social situations meaningful, to so construe them as to make it possible to act purposefully within them, that accounts both for the ideologies' highly figurative nature and for the intensity with which, once accepted, they are held . . .

The differentiae of science and ideology as cultural systems are to be sought in the sorts of symbolic strategy for encompassing situations that they respectively represent. Science names the structure of situation in such a way that the attitude contained toward them is disinterestedness. Its style is restrained, spare, resolutely analytic: By shunning the semantic devices that most effectively formulate moral sentiment, it seeks to maximize intellectual clarity. But ideology names that structure of situations in such a way that the attitude contained toward them is one of commitment. Its style is ornate, vivid, deliberately suggestive: By objectivizing moral sentiment through the same devices that science shuns, it seeks to motivate action. Both are concerned with the definition of a problematic situation . . . Where science is the diagnostic, the critical, dimension of culture, ideology is the justifying, the apologetic one - it refers to that part of culture which is actively concerned with the establishment and defense of patterns of belief and value.

Evaluation is an ideology which blurs the distinction between science and ideology:

Interest in evaluation research has been greatly stimulated in the past decade by widespread concern for domestic social reform. Searching questions have been raised about the adequacy of organized programs in such institutional sectors as health, justice, education, employment, housing, transportation, and welfare. In an atmosphere charged with demands for rapid and significant change, a great many innovative action programs have been introduced. Some reformers have urged that the quest for more effective institutions be orderly and cumulative. They have argued that careful program evaluation is needed as a basis for continued planning and have recommended that the methods of social research be utilized in the evaluation of reform programs. (Caro, 1971)
Evaluation would be used also for policy and administrative decision making in the domain of human services.

Evaluation as an ideology is part of the scientism ideology. Both of these ideologies fit into the model of a scientific social reform movement.

EVALUATION AS A SCIENTIFIC SOCIAL REFORM IDEOLOGY

Evaluation was the ideology of the evaluation research scientific social reform movement. This ideology sought to give legitimacy to the social change efforts of those who wanted to implement evaluation research in human services programs by arguing that evaluation research on a grand scale was a new and rational way to decide about these services. In this sense, evaluation was a particularly potent example of the scientific social reform ideology of "newism" (Eaton, 1962).

The newism ideology is "the presumption that new developments or practices are superior to those 'not quite so new' or old." (Ibid.) In this ideology, . . . the attribute of novelty is presumed to be indicative of validity . . . . Plausible rather than well-documented facts are used to support the theory that stylistic changes also represent a gain in operating efficiency and durability. (Ibid.)

Evaluation research on a grand scale was the novelty in that it is presumed to be a more effective tool for rational decision-making about human services. An examination of the symbolism of the evaluation ideology suggests, however, that the scientific social reformers sought more than the widespread implementation of evaluation research of the human services. Using the ideology of evaluation, they sought to rationalize the human services system. This is seen in the notion of "accountability" which is integral in the symbolism of the evaluation ideology.

THE SYMBOLISM OF THE EVALUATION IDEOLOGY

Evaluation as an ideology is understood in part by an analysis of the "figurative nature" of its words, notions, ideas, metaphors and symbols; and by the consequences of these symbolic forms on and for particular groups of people who comprise the socio-political human services institution.

An essential character of religious symbolism is its multivalence, its capacity to express simultaneously several meanings the unity between which is not evident on the plane of immediate experience . . . This capacity of religious symbolism to reveal a multitude of structurally united meanings has an important consequence: the symbol is capable of revealing a perspective in which diverse realities can be fitted together or even integrated into a "system." One cannot sufficiently insist on this point: that the examination of symbolic structures is a work not of reduction but of integration. One compares and contrasts two expressions of a symbol not in order to reduce them to a single, pre-existent expression but in order to discover the
process by which a structure is capable of enriching its meanings.
(Eliade in Edelman, 1971).

"Evaluation" can be thought of as a symbol in which are combined several ideas, notions, words, meanings and social processes. Among these are "knowing," "monitoring," "research," "scientific," "applied research," "field study", and, most important, "accountability." It is our notion that there is in the symbol "evaluation" a meaning of "accountability," and that this meaning is, in turn, a language of rationalizing, organizing and controlling. That is to say, the notion of "evaluation" can be viewed as an "administrative language" (Edelman, 1967) by which political control is exercised or attempted in the social domains of human services as a means of bringing "order and reason" and socio-political legitimacy to public and private expenditures for human services. Evaluation is a "managerial ideology" (Krause, 1973).

The word evaluation has come to include several notions, ideas, meanings, and social processes. Among these are a type of empirical socio-behavioral research and a technology of learning about something (Robinson, 1971). To human service practitioners, evaluation means a requirement that their program be examined by "outsiders." It is a research process in which data will be collected, analyzed, etc., "to help us learn about what we are doing so that we can change and correct it," or "really show how good we're doing." It is something "hard," "scientific," "objective" and "scary," it is something "scary" because evaluation is understood by practitioners to mean "accountability" in a new and different sense.

The Notion of Accountability

The notion of accountability includes ideas from sources such as moral, political and legal philosophy, and it has the formal meanings of "answerable to, capable of being explained, and responsible to (for)." In this last meaning, accountability is found in the political idea that an organization has many constituencies which hold expectation of it, that some of these expectations and requirements are about funds, "treatment" of people, "quality service" and the like. Since the expectations and requirements cannot be learned about a priori, it is necessary that facts be collected so as to learn about whether or not and to what degree these expectations are "met." That is, "research" is a method of learning about the agency. Evaluation (and evaluative research) is one type of research (technology) which can be so used. Evaluation and accountability have come to be joined in the mind of the practitioner.

Accountability also refers to the social process of administrative/political/management decision-making and control. In particular, to practitioners it refers to fiscal decision-making - to funding and refunding. Hence, accountability means agency survival.

The symbol evaluation, then, includes in its penumbra the notion of agency existence. This is one psychological and social reality for human service workers and is likely part of the reason why evaluation is "scary."
EVALUATION AS A LANGUAGE OF ACCOUNTABILITY

"Accountability" like evaluation are themes in the larger "human service ideology" (Baker, 1974). "Control" is one notion found in the ideas of accountability and decision-making in the evaluation ideology. Control over agency survival is one example of this. Another example is in the idea that evaluation means to some a language of rationalizing and organizing, i.e., of controlling the human services "system." These notions and ideas are found in the language used to discuss accountability and evaluation.

Four sets of language are heard in discussions about evaluation and accountability: The language of the human services and politics long used by participants in the traditional bargaining systems; that of the new bargainers - new professionals, new client groups and the like; the language of corporate business - efficiency, markets, PR, products, cost-benefit ratios - as this has become the language of the federal executive branch; and the language of socio-behavioral research - as evaluation has become a means of achieving efficiency and effectiveness.

The language of corporate business is a managerial language of "hardness," coldness, clear-headedness, and data.* It is a language of ends, of goals and objectives, and it is a language of numerality. Like legal language, it is a language of specificity, of things which can be measured. It is a language of objectivity. It contains words, symbols and metaphors which present images of rationality, organization and order;** of facts and "science." (Boulding, 1969).

This language brings semantic and cognitive order to the human service system: inputs and outputs are clear(er). It is a language of control as order and regulation facilitate management. It is a language which makes things "researchable" - discrete, clean, time-limited. It is a language of accountability. It is a language of evaluation; of science; of scienticism - the ideology.

This language is not traditional in the human services, particularly in small, private agencies; it is certainly not the language of the alternative agencies. There are well-documented differences between the languages of research and human services (with clear exceptions in some branches of clinical medicine and psychology, among others). This was of major import to researchers who hoped early on to do evaluative research which would have utilizable findings incorporated by service workers. Most practitioners gave symbolic, not substantive support to their efforts (Eaton, 1969). Now the import is more severe, for, as suggested, there is increasing congruence between the languages of research and management. There is actual power to use the symbols. Language has become one medium of administrative control over human services accountability.

*See Adams, (1974) for the relations between and among the notions and words such as "hard - soft," "male - female," "right-handed - left-handed," etc.

**Contrast this to the notion of "muddling through."
Participants in the evaluation research movement sought the widespread use of this research approach in the human services. They were successful in that the idea that human services should be evaluated is now socially legitimate. It is found in laws, regulations and rules for the funding of human services in both the public and private domains. The stated purpose of this movement was to implement this type of research in order to make decision-making about human services a more rational and effective process. The argument for this position was found in the evaluation ideology - a culture system with a symbolism and a language. In these latter was embedded the notion of accountability. The social movement sought control of decision-making by offering a rational and seemingly objective way to insure agency and worker accountability. To achieve these ends, those advocating evaluation research had to be accepted by, or to become, an elite group in the human services system.

In the recent past, researchers and evaluators were marginal to the centers of decision-making in human service agencies and in the human service system. The opportunity to change this status occurred after the Nixon Administration reduced the War on Poverty and spending for other social programs. Funds for social programs became scarce at a time when many programs were new so that there was a high relative scarcity of funds. At the same time, the ideologies of business, of management, of efficiency and of accountability were introduced by the Administration.

The War on Poverty brought more than new human service programs and large public funding for these. The ideologies of "citizen participation" and "community (control)" were legitimated. Attempts were made to include clients of the service on decision-making boards. "New careers" were conceived and implemented. The poor, the minority and the "new professionals" were active participants on the local levels of publicly funded programs. These groups were an emerging new elite in the human service system.

Regardless of intent, one consequence of the evaluation research social movement was a challenge by researchers and their allies to the citizen participation movement: The rationality of research against the experience of being poor. Evaluation was the ideology which gave legitimation to this new elite of researchers (Krause, 1973).

The evaluation research movement challenged the older human service elite, too. Here the issue was the "sloppiness" of human service systems. The parts of the systems were disjointed when examined from the perspectives of client service, interagency relations or the like. The old elite could be held responsible for these inefficiencies and for the lack of service effectiveness.

Evaluation in its symbolism and language offers cognitive and social means to rationalize the human services systems (Rich, 1973). It is a way to think about and to achieve "order." The problems of actually doing and utilizing evaluation research of the human services are very real. Researchers, now successful, may soon face a severe problem: "How to deliver." It may be at
that point that the demands for evaluation research become modified. For the moment, evaluation research "is in." Examined next are some consequences of this successful scientific social reform movement.

SOME CONSEQUENCES OF EVALUATIVE RESEARCH FOR THE HUMAN SERVICES

On the Level of the Worker

The language of evaluation is spoken increasingly by human service workers. However, very often only the words are learned and the meanings are changed or lost. When the meanings are kept, social control is gained for "language can be thought of as a system of social control with vocabularies socially canal-izing thought" (Pitkin, 1972).

Practitioners want and need evaluation research, hold unusual expectations of it and are disappointed by its outcome. In response, they have created a vocabulary of "reasons why not" i.e., why statistically significant findings were not found:

--"The therapeutic relationships examined or the impact of the program is 'too subtle to measure with statistics'."

--"The presence of outsiders disturbs the normal conduct of the program or the group or the session."

--"Even though they may come back to prison, they are better or happier or more emotionally stable people for having participated in the program."

--"The effects of the program can only be measured in the long run, not just during the first six months or year after release ."

--"The program or the technique is OK but it is not designed for this particular individual."

--"The reason that the program failed is that it wasn't extensive enough or long enough or applied by the right people."

--"The program is worth it if it saved one man." (Ward and Kassen- baun in Weiss, 1972).

Silly comments, defensive statements, these? To some ears, these can be read also as a solid critique of evaluation research as now practiced. The conflict in orientation between the researcher and the practitioner is clear. This conflict is emerging into public discourse as another consequence of the social movement for evaluation. Related directly is the issue of measurement: What do the evaluation researchers measure in their studies? Often "objectives."
Practitioners have begun to "manage by objectives," to quantify expected outcomes to/in their clients and to measure these (most often with the help of researchers). Too often, only the easily measurable is examined regardless of its centrality to the program, its management, or to its service ethos (Rich, 1973). "Management by objectives" (MBO) is both a technology and an ideology. So too is "research utilization." Measurement as a technique used in MBO to collect data to utilize in programs is also used in political and ideological ways. It is these uses which contribute to the emerging widespread clash between practitioners and researchers (Rich, 1973).

For the worker, evaluation research means records, forms and instruments. To him, his success and failures become numbers, not people. Often, he tries to get around these papers. He fudges. He gets cynical, often angry. He begins to avoid certain tasks. There emerges a deviant social role for the human service worker.

Another consequence of evaluation often is that the instruments of the process - the forms and records and the like - become incorporated and institutionalized. Information sought for one end is used for no end. More red tape and bureaucracy follow, with the attendant difficulties in changing these ways.

Yet another consequence is that the data are moved from the practitioners and the managers to the researcher who designs, collects, analyzes, interprets, etc. - all in his own corner. In turn this seems to make the goal of management utilization of findings more difficult to achieve. Further, this contributes to a sense by workers that the researcher is an "outsider" - one who is hard to reach, who doesn't understand practice, etc. And, maybe most important, the practice of organizationally separating research from practice and researchers from practitioners is to disengage the worker's act of giving treatment from (a) feedback about the act and the particular worker - client dyad; and (b) the worker's professional responsibility for that act. Responsibility and moral accountability becomes effectiveness - "did it work? The researcher will tell me." The idea of and the social values basic to professional practice and professionalism are changing.

On the Systemic Level

Along with emerging notions of professional behavior and responsibility are other consequences of the institutionalization of the evaluation ideology in practice. Among these are the following.

A new form of social legitimacy and social control of the human services has emerged. With this has come a new elite: social scientists, and systems and operations-management people. These people, styles and ideologies are

*See Rhinelander, (1973) for suggestive comments about the "optical model" of the mind and the relation of this to notions of "objectivity" and, by extension "objectives."
important participants in the socio-political bargaining systems of the human services. This is a gain for the technocrat and the methodologist.

A third consequence is panic on the part of practitioners and managers as they begin to see that the question of "does it work?" could lead to their program being closed. A finding of "no proof of effectiveness" can be a death certificate.

Another consequence can be thought of as the "achievement crisis" (Eaton, 1974); researchers will have to produce too. They will have to devise technologies of evaluation which will meet the needs and wants of the socio-political decision-making process or else they will lose the legitimacy which governs their elite status in these decision systems.

Last, here, is the consequence of feeling hopeless because the human services have been "captured by outsiders." Where does this leave the practitioners?

BEHIND THE 'RHETORIC OF RECONCILIATION'

The practice of the rhetorical reconciliation helps to explain how American society escapes many of the strains that might be expected when its professed ideals conflict with many of its accepted practices. There is a tendency to describe accepted practices in ideologically acceptable terms, whether or not the terms are truly descriptive of practices . . . What is significant is the metamorphosis of the normal language shifts in American society into a rhetorical reconciliation between opposing practices and beliefs; a reconciliation that serves to perpetuate the conflict by rationalizing practices and delaying changes in ideology. (Skidmore, 1970).

There is always hope of a reconciliation between differing views - here, between practitioners, researchers, decision-makers of policy and funding, the "public" in general and organized constituencies in particular. This is a basic tenet of our political ideology and a basic perspective in the socio-political analysis of individual and group wants and needs.

There is a possibility, though, that the very success of the evaluation social movement has institutionalized words, notions, ideas, symbols and social processes which are at root inherently different, as different as art and science; and that at best there are but tangential ways in which these culture systems overlap and can be brought into mutual support on the level of workers and clients. It is our task to find these points, for without them, we will clothe our art in robes of scientific rhetoric and be no better for it. Our responsibility is to search for a fit between research and practice in the service of our ideals.
BIBLIOGRAPHY


Evaluation, all issues.


Haines, Nicholas, "Responsibility and Accountability", Philosophy, (1955), 30, No. 4: 141-163.


-287-


ANTI-POVERTY POLICIES AND EVALUATION:
A CRITIQUE OF THE PLURALIST CONCEPTION
OF POLITICS AND EVALUATION

Robert D. Herman

School of Administration

University of Missouri - Kansas City, Missouri 64111

The beginnings of U.S. federal evaluation research can, in descripti-
tive historical terms, be located in McNamara's Department of Defense and the
later requirement that all federal government agencies adopt a Planning,
Programming, Budgeting System. While the formal PPB system was discontinued
in 1971, the analytical or policy evaluation activities it required, still
live on, especially in agencies dealing with human resource development and/or
social welfare programs (Schick, 1973, Wholey, et al, 1970). Given the recent
advocacy of increasing and improving federal evaluation efforts, I think it
important to examine some of the assumptions and consequences of those
assumptions of the general evaluation research model, particularly in regard
to anti-poverty policies.

Evaluation research, whether it be of anti-poverty policies or court-
ordered school-desegregation programs, is much like other social science
research. The chief difference is that the hypothesis of evaluation research
is, on a very general level, that some governmental agency program has achieved
its intended goal. While this simple statement in reality often involves
dealing with a whole host of methodological and organizational problems, the
point is that evaluation research is intended to be useful to decision-makers.
Evaluation research is not primarily concerned with theory building and theory
testing, but with practical "real-world" affairs. This makes evaluation
research important and often controversial. It is controversial because it can
effect (or is perceived as such) the budgets, staff size and power of various
agencies and the distribution of public benefits to various classes or
categories of citizens. It is important because methodological and empirical
issues - such as sample size and representativeness, validity and reliability
of measures, accurate operationalization of program goals, and appropriate-
ness of causal models and inferences - now have (or are perceived as having)
immediate and direct political consequences.

Some evaluations, for instance, the Coleman report and the Westinghouse
Learning Corporation - Ohio University study of Head Start, have become centers
of controversy. It has been argued that criticisms of these evaluations, while
explicitly directed toward methodological issues, are basically politically and
ideologically motivated (Williams and Evans, 1969; Weiss, 1972). The implica-
tion seems to be that such criticism will severely hinder not only the utiliza-
tion of the given specific evaluation study, but will also retard the extension
and improvement of evaluation generally. When evaluation advocates call for
de-politicizing evaluation, they are not claiming the evaluation is not part of
the political process or that evaluation does not have political implications.
Rather, evaluation must be de-politicized so that it can help make political decisions more rational by supplying evidence (of an "objective or disinterested" character), which presumably is equally available to contending parties, on which to make decisions about maintaining, expanding, modifying or abandoning programs intended to achieve some general policy goal. That is, in order for evaluation to be useful, the design, implementation and data analysis phases of evaluation must be free from pressures that would or could bias the outcome(s) in one direction or another. Indeed, much of the recent literature about evaluation is concerned with such issues as quasi-experimental designs, causal models, the proper organizational locus of evaluation projects, and multi-variate statistical methods (see, for instance, Rossi and Williams, 1972; Wholey, et al, 1970; and Weiss, 1972). In order for these evaluation phases to be free of biasing pressures, this requires that parties interested in a policy area eschew political-ideological attacks upon evaluation. It is in this sense that evaluation research needs to be de-politicized and its a-political or disinterested character preserved, recognized and enhanced. This is, I submit, the typical frame of reference for considering the issue of politics and evaluation.

All the policy evaluations and discussions of politics and evaluation that I've seen assume or accept the dominant theory of the U.S. political system - pluralism*. Thus, policy goals are viewed as the result of bargaining among various coalitions of interest groups. Such coalitions are not stable, but shift as issues or classes of issues change. Each interest group has limited power and consents to the "rules of the game." Such bargaining results in piecemeal, incremental policies. In this view, all citizens are seen as essentially similar, each having equal access or opportunity to participate in the political arena. Figure I presents a highly simplified schematic diagram of the pluralist view of the political system and the place of evaluation in the system. Evaluation comes at the "end of the line" of a cycle and provides feedback in the (intended) form of "objective" or "disinterested" evidence about the outcomes of a policy decision. This evidence is then to be used in the following cycle.

*While some have distinguished different varieties of pluralism, for the purposes of this paper, all pluralisms are essentially the same.
Figure 1
Pluralist Model of Political System
(Including the location of evaluation)

Authoritative (governmental)

Decisions on policy goal(s)  ➔ Implementation  ➔ Evaluation

(administrative agencies)*

feedback of evidence

Interest groups

Access to active participation

Mass of Citizens

Legend
  0 -- -- interest group
  ( ) -- temporary coalition
  → → -- bargaining and influence attempt

* Administrative agencies also often behave as interest groups, and their location in this figure is not meant to suggest otherwise.
But does the pluralist perspective adequately account for policy decisions concerning the poor - such as the creation, passage and implementation of the Economic Opportunity Action and its provisions for Community Action Agencies and Community Action Programs? And if it does not, what does this imply for the appropriateness of the traditional evaluation approach? The fundamental axiom of pluralism is that policy decisions are the result of bargaining among interest groups that have some power, and that have something (symbolically but especially materially) to gain or lose from the decision. While the poor apparently had something to gain from the Economic Opportunity Action and from Community Action Programs, to what extent did the poor act as an interest group in framing or in securing approval of the Economic Opportunity Act? Levitan (1969) shows that various interest groups, mostly federal departments and agencies such as HEW, Labor, Commerce, and the Bureau of the Budget, did indeed bargain over, and arrive at compromises about substantive and administrative issues. However, nowhere is there any evidence that the poor were involved in the decision process. Sheppard (1969: 487) puts it succinctly "Unlike other class-related laws, in this case the law was only remotely - if at all - the result of lobbying or direct pressuring by the class intended to be positively affected."

Most analysts of the political history of the Economic Opportunity Act, though sometimes puzzled by it, are not much concerned with explaining this apparent anomaly, preferring to see governmental decisions as being not just the result of pluralistic bargaining, but also, on occasion, the result of a rational, more or less disinterested, professional solving orientation (see, for instance, Donovan, 1967; Sundquist, 1968; Rubin, 1969; Marris and Rein, 1967; Moynihan, 1965; and Moynihan, 1969). The pluralist-problem solving perspective, thus, sees the War on Poverty as originating in the recognition of poverty as an important social problem by professionally trained social scientists in foundations, universities, and government who were able to influence sympathetic politicians in the Democratic administration. Many of the specific programs, such as Community Action Program, are seen as attempts to apply, to some extent, social science concepts and theories, to important social problems. The War on Poverty programs, to the extent that they are based on social science theories at all, are based on or congruent with the cultural or sub-cultural theory of poverty (Valentine, 1968). While the cultural theory of poverty, and its chief rival, the situational theory are not well-developed theoretically (Spilerman and Elesh, 1971; Herman, 1974), the cultural perspective sees (hard-core) poverty as the result of socialization in a sub-culture that produces personality characteristics and orientations (such as fatalism, preference for immediate gratification, feelings of powerlessness, etc.) that prevent such people from taking "full advantage of the changing conditions or increased opportunities that may occur in their lifetime." (Lewis, 1968: 6) Thus, traditional evaluations of anti-poverty programs would attempt to assess to what extent such programs (in their various forms) resocialized, changed the personalities, of the poor who were, in some way, included in the program. This was certainly the case in Westinghouse-Ohio University evaluation of Head Start, and appears to be the case in evaluations of Community Action Programs.
This seems reasonable enough as far as it goes, but does it go far enough? It does not, if there is another perspective that more adequately accounts for the War on Poverty programs. An elite or limited pluralism perspective provides a more convincing explanation of the War on Poverty. As developed by Gamson (1968), in an important but much neglected article, the U.S. political system can best be described as one of limited pluralism. Gamson agrees that the political system contains multiple centers of power, but then argues that the political system normally operates to prevent incipient groups from achieving full entry into the political system, and thus such groups are unrepresented in the political system. Represented groups do indeed bargain and compete among themselves, but they also cooperate (whether intentionally or not is unimportant) to maintain the boundary (i.e., hinder entry) to the political system. For instance, the creation of political organization among unrepresented groups is discouraged by such general social conditions as (1) a lack of access to information about the effects of political decisions, (2) the lack of politically experienced and skilled leaders, and (3) the "culture of subordination". The political establishment can and has undertaken more specific actions that discourage the formation of interest groups and their inclusion in the political system. Such boundary maintaining activities include: (1) attempts to undermine the legitimacy of a developing interest group; (2) harassment of leaders; (3) co-optation of leaders; (4) going around leaders of a developing interest group and appealing to members and potential members to reject the interest group and its leaders as their representative; and (5) incrementalism in the resolution of major social issues. Gamson explicitly points out that all citizens are not the same. Social stratification of citizens affects access to and resources available for participation in the political system. Figure 2 presents, once again, a simplified and heuristic diagram of the elite or limited pluralism perspective.

Figure 2

Elite or Limited Pluralistic Model of Political System

Social-economic Stratification of Citizens

Increasing access to and resources for political participation

Authoritative (governmental) Decisions on policy goal(s)

boundary maintaining

Implementation

Unrepresented Groups

Legend

0 — — — interest group

0 — — temporary coalition

= bargaining and influence attempts
Since most research on the poor has been concerned with investigating the culture or subculture of poverty, the psychological characteristics of the poor, and the impact of programs designed to resocialize the poor, there is little direct data available to assess the extent to which the poor, or what and which proportion of the poor, are unrepresented in the political system. However, I think it safe to contend that for the most part, and in comparison to more privileged socio-economic classes, the poor are unrepresented.

The elite or limited pluralism perspective sees governmental policy decisions, especially those dealing with or "for" the poor (or for other unrepresented groups), not as exercises in consensual problem-solving, but rather as a means of winning elections without alienating established interests in (i.e., while maintaining the stability of) the political establishment. Though Piven and Cloward (1971) do not explicitly utilize Gamson's model of the political system, their analysis of the functions of welfare fits very nicely with and illustrates Gamson's views. Piven's and Cloward's (1971) primary thesis, supported by chronological data they present, is that expansive welfare policies (increasing expenditures, "loosening" requirements) function to dampen civil disorders, while restrictive policies (decreasing expenditures, "tightening" requirements) function to reinforce work norms. Betz (1974) recently provided additional empirical support for this basic thesis when he demonstrated with cross-sectional data that cities experiencing riots had significantly larger increases in welfare expenditures the year after the riot than non-riot cities. Thus, there is evidence that supports the proposition that welfare policies are not attempts to solve the poverty problem. (See also Goroff (1974).) Rather, the functions of welfare policies are to win elections (either by creating and appealing to the prejudices of working and middle-class people or by "buying" the votes of the poor) and maintain the stability of the political system (either by heading off changes in the political system that civil disorder might bring about or by emphasizing, to both the poor and non-poor alike, the necessity of work).

In the course of documenting this basic thesis, Piven and Cloward (1971) offer an explanation of the origins and functions of the War on Poverty that is consistent with their basic thesis and with Gamson's theoretical model of the political system, and at odds with the pluralist problem-solving accounts. Briefly, they argue that by the 1960s most local urban Democratic parties were unconcerned with and hence unresponsive to the black and poor in the cities. The national Democratic party and its candidates, though, were dependent on capturing and holding this constituency. In order to do this, the federal government intervened in local welfare arrangements, not directly through liberalizing legislation, but indirectly by creating programs, such as the Economic Opportunity Act and its provisions for Community Action Programs, that helped to mobilize pressure against local welfare policies. Regardless of the intentions of the key political actors, the War on Poverty and its Community Action Programs, by circumventing local Democratic establishments, functioned to overwhelmingly secure the inner city poor and black vote for the national Democratic party, and its candidates - Kennedy and Johnson. However, in some cities the conflict between the poor and the local political establishment (often Democratic) became
so intense that mayors and others contended that the stability of the political system was threatened. The opposition of big city Democrats and their allies can account for the retreat from citizen participation in Model Cities programs established by the Demonstration Cities and Metropolitan Development Act of 1966 (Cole, 1974); Strange, 1972). Unlike CAA programs, all federal funds for Model Cities programs go directly to the municipal government, thus reestablishing local political control over the programs.

It is, of course, unrealistic to expect evaluation studies conducted or sponsored by the government to depart from a reliance on the pluralistic problem-solving model of the political system or to go beyond investigating the extent to which formal policy goals (no matter how multiple, vague or broad they may be) are achieved, and at what cost-benefit ratio. However, the point of this exercise is to show that traditional evaluation can not be truly de-politicized. Evaluation, particularly of anti-poverty or other social welfare policies, can not but be political, even in its research phases, being based on the pluralist theory, evaluation research is constrained to raise only certain issues and hypotheses. Evaluation, in this view, functions (whether intentionally or not is unimportant) to reaffirm the pluralist view, and to focus analysis on specific programs rather than on the political system which produces them.

One sometimes senses that evaluators and evaluation advocates are uneasy about the narrowness of evaluation research. For instance, Weiss (1972: 334) in a section titled, "Implications of Evaluation: A Radical Critique" argues that "in a basic sense, the bent toward the negative that is characteristic of social action evaluations is not something to be masked or shunted aside. To the extent that null results are real and not an artifact of primitive methodology, they betoken serious weaknesses in social programming." Later she concludes that "it is time we recognized the failure of our moderate, piecemeal, cheap solutions to basic social problems. They have been tried and evaluation research has found them wanting. Bold experiments are called for. It is a fraud to perpetuate variations on outmoded solutions to problems that are rooted in our system of social stratification." (Weiss, 1972: 337)

If the elitist model presented here has any validity, and I think it has a great deal, it implies for those of us concerned with analyzing poverty and inequality and attempts to reduce it, that we not allow formal policy goals and their evaluations to define the entire field of investigation. Additional issues and questions need to be raised. It implies that relying exclusively on the government (or more broadly, the political system) to seriously do something about reducing poverty and inequality is foolhardy. Finally, it implies that poverty and inequality might be more effectively reduced by focusing on ways of promoting the formation of interest groups among the poor and of easing entry into and changing the established political system. That is, the problem of poverty might be more fully solved if we analyze and find ways of changing the policy process, instead of analyzing and changing only the policies.
REFERENCES

Betz, Michael

Cole, Michael L.

Donovan, John C.

Gamson, William A.

Goroff, Norman N.

Herman, Robert D.

Levitan, Sar A.

Lewis, Oscar

Marris, Peter and Martin Rein

Moynihan, Daniel P.
1965 "The professionalization of reform." Public Interest, Fall (no. 1): 6-16.

Moynihan, Daniel P.

Piven, Frances Fox and Richard A. Cloward
Rossi, Peter H. and Walter Williams (eds.)

Rubin, Lillian B.

Schick, Allen

Sheppard, Harold L.

Spilerman, Seymour and David Elesh

Strange, John H.

Sundquist, James L.

Valentine, Charles A.

Weiss, Carol H. (ed.)

Weiss, Carol H.

Wholey, Joseph S., John W. Scanlon, Hugh G. Duffy, James S. Fukumoto and Leona M. Vogt

Williams, Walter and John W. Evans

-797-
INTRODUCTION

As has been pointed out by a host of writers, a crucial issue in conducting evaluative or action type research is the problem of gearing an appropriate research design into the cogs of organizations, ensuring that once initiated the evaluation is in fact conducted according to the original design, and that the results of the study are utilized in policy formulations and ultimately in program development. The focus of this case study is on the major political and administrative problems and processes involved in planning, conducting, and utilizing the findings from a field experiment dealing with the relative effects of parole supervision as provided to a group of juvenile aged boys and girls.

An assumption of the paper is that many of the problems encountered in the process of conducting evaluation research in the field of Corrections will have relevance to a variety of projects in different substantive fields. Accordingly, there is a need for illustrative case material dealing with the problems involved in planning and conducting the research as well as in utilizing the findings derived from an evaluation project. The hope is that such material will accumulate to provide insight into the peculiar nature of the socio-political processes endemic to this work, help to illuminate issues, suggest general propositions and indicate avenues for conflict resolution.

The case study reported on here centers on an examination of outcomes resulting from the provision of differential parole supervision to juvenile aged boys and girls released from the two major state-level correctional institutions in Minnesota between August 1, 1970 and May 31, 1971. The aim was to determine whether juveniles released from institutional settings would adjust as well on parole without formal supervision from parole officers as a corresponding group receiving conventional parole supervision. Viewed in chronological sequence, five major steps were involved in implementing the research: (1) the identification and assignment of boys and girls from the
two institutions to a study pool, (2) random assignment to experimental or control status from within the study pool, (3) release on parole, (4) provision of either conventional parole supervision or no parole supervision to the respective groups, (5) evaluation of outcomes on parole. In short, the formal design of this research took the form of an "after-only" field experiment.

Prior to the actual implementation of the study design considerable time and effort were devoted to planning with the relevant administrative units of the Department of Corrections. While detailed attention to defining relationships with the various units involved in the conduct of the research at the outset was time consuming and frustrating, none the less it was to be an essential pre-condition to the successful conduct of the project. The assumption was that if the research could be fitted into the normal operations of the agencies concerned, resulting in only mild disruption in routines, the probability of receiving continued on-going administrative support and cooperation would be greatly enhanced.

The pertinent administrative units dealt with included the Parole Division of the Department of Corrections as the administrative unit responsible for providing parole supervision to all juveniles released from state-operated correctional institutions. This division was made up of parole officers, supervisors and administrators scattered throughout the state. The Youth Conservation Commission or parole board was a further key administrative unit as it was responsible for decisions on the placement and release of youth in state correctional institutions. Finally, the two correctional institutions from which samples were to be selected were obviously relevant.

Because the project involved some departures from agency "business as usual" detailed discussions of the general purpose as well as its implementation were carried out to help ensure that the research would be conducted with a minimum of confusion and in accordance with the originally agreed upon procedures.

NEGOTIATIONS WITH THE PAROLE BOARD

Once the members of the Parole Board had formally agreed to the conduct of a study that proposed to assess the relative impact of "regular parole supervision" and "no parole supervision" with juvenile boys and girls, two major questions had to be resolved. First, was it necessary for all youth to receive parole or could some youth be given flat discharges from the institutions? A further question concerned the kinds of restrictions to be placed upon the sample selection procedures by this releasing authority (i.e. what youth would not be eligible for no parole supervision)?

The latter issue generated heated discussions between the researchers and the paroling authorities. On grounds of greater potential generalizability of research findings, the investigators preferred that the study be conducted with no restrictions on the kinds of youth to be placed in the sample.
selection pool; the simple fact of imminent release on parole would constitute the only logical basis for random assignment. On the other hand, the releasing authority felt responsible for potentially endangering either communities or particular parolees as a result of releasing certain youth without the presumed benefits of parole supervision. Involved here was an overriding concern on the part of the Board members with what they saw as a distinct possibility of engendering negative reactions from community members due to the sampling requirements of the study. This possibility, moreover, had to be of concern to the investigators, for if a youth were released on parole without supervision and subsequently became involved in a serious criminal offense, a cause and effect link - no matter how spurious - might be imputed and lead out to pressures to terminate the study. As political appointees, members of the Parole Board could reasonably be expected to withstand only limited pressures from the community; particularly if those pressures were directed against research. This issue was finally resolved by the two parties agreeing that certain categories of youth would be kept out of the pool from which samples were to be drawn.

Given this constraint, it was then necessary to develop clear criteria on the basis of which such youth could be identified and excluded from the random selection pool. The researchers were concerned that too wide a set of invalidating criteria would be defined by the releasing authority; in contrast, the Board members were concerned that too narrow a set of invalidating criteria would be defined by the researchers, thus allowing an unacceptably broad spectrum of youth the chance of being included in the experimental or "no supervision" group. To avoid a stalemate over these differing perspectives it was necessary to negotiate a compromise.

From a number of criteria, five were finally filtered out by the Board members as being especially troublesome. These remaining criteria were then used to exclude youth from the study pool: (1) officially known involvement in an act of arson, (2) officially known involvement in an act of rape; (3) formal diagnosis of 'severe emotional disturbance', (4) parole to a living situation other than to the legal parental home, (5) a consistent record of assaultive behaviors. While each of these criteria reflect some degree of vagueness, they were seen by the researchers as reflecting the best compromise that could be developed under existing conditions.

A second major issue that had to be negotiated with the paroling authorities prior to the implementation of the project was that of determining the exact nature of parole for the experimental and control groups. The investigators preferred that the experimental group members be given flat release from the institutions to eliminate the possible confounding effect of parole status, itself. The design of the research would simply involve first assessing whether or not prospective parolees met the study criteria and second, given that these were met, randomly assigning them to either the conventional parole supervision group or to the experimental group - the members of which would be provided with outright discharges upon release from the institution. With the intruding factor of 'parole' eliminated for the experimental group, the potential effects of parole supervision as compared to no parole supervision could be accurately and fairly easily assessed.

-300-
Because standing policy dictated against flat release from institutions for juveniles, the members of the Youth Commission felt that they could not accept this proposal. As a result, parole had to remain a constant for both study groups with modifications made around the relative degree of supervision received while on this status. Because the conventional parole agreement stipulated conditions requiring the parolee and parole agent to maintain contact, a modified parole agreement form had to be developed by the investigators for those youth assigned to experimental status. The major modification made was to eliminate the requirements for parole agent-parolee contact. Clearly, however, ethical requirements demanded that in a somewhat subjectively defined 'crisis', the youth would have the option of contacting the parole officer. As a consequence, procedures had to be developed for monitoring parole contacts made by members of the experimental and control groups.

Once the issues of defining excluding criteria and maintaining the fact of parole as a constant were dealt with, the eventual size of samples had to be negotiated. These issues were all inter-related inasmuch as the specified criteria would directly affect the proportion of all released from the two institutions that could be assigned to either the experimental or control groups. Thus, if the five criteria eliminated a large proportion of the total number of released from possible inclusion in these study groups, the time period set for obtaining a sizeable sample would be considerably lengthened. Compounding this problem was a lack of information on the proportion of youth being released from the relevant institutions falling within the five study criteria.

For similar reasons the members of the Parole Board, the administrators of the parole division and the administrators of the institutions were concerned with this question. The Board members were not prepared to sanction a study which would take an interminable length of time to complete. Furthermore, agency administrators were not disposed to tolerate research-caused disruptions in the normal routine of their programs for any prolonged period of time. As a result of these considerations a ten-month follow-up period was agreed upon. While this follow-up was far too short to obtain information on all sample members who would eventually return to correctional settings, it did cover the average length of the parole period for juveniles in the state.

A more pervasive type of public relations issue had to be dealt with once the above questions had been resolved. During the initial planning stages, the Board members and the administrators of the parole division had raised persistent concerns about the manner in which local police and juvenile court officers were to be informed about the proposed investigation. In particular, concern was expressed about the possibility of generating negative reactions from these local officials if the proposed study were conducted without their explicit approval.

To handle this problem the Board members suggested that a description of the proposed study be provided to local police and court officials prior to
initiating the research effort. The rationale for proceeding in this way was to provide local authorities with an opportunity to react to the study before it was formally implemented. The investigators, on the other hand, were wary. They saw the suggestion as a catalyst stirring up negative sentiment among local officials and possibly undermining or even aborting the project before it got off the ground.

As a less disastrous alternative, the researchers proposed that following the actual implementation of the study, copies of a form letter describing the research be provided to each parole officer in the state who would then be free to use his own personal discretion in providing it to the local police and court authorities. Proceeding in this way had several advantages. First, because the local parole officers would be better able to assess potential reactions to such a letter, they would be in a better position to make an informed judgement about making it available. Also, informing local officials after the actual initiation of the project might discourage - if not eliminate - action to terminate it since they would be faced with a fait accompli. Furthermore, since the field agent was closely involved with these local officials, there would be a greater opportunity to explain on a more intimate basis those issues which were seen as requiring further clarification. Finally, the process of having the parole officer explain the nature of the study to these local officials was expected to diminish the possibility of the research being interpreted as a case of local youth being "victimized" through "depriving" some of them of parole services.

After frequent and lengthy discussions with the members of the Parole Board and the administrators of the parole division, this issue was resolved in the manner proposed by the investigators. After the first monthly group had been randomly assigned, the parole officers were equipped with form letters which at their own discretion they provided to local authorities. The low probability of field agents having an experimental group member assigned to their caseload was a point that was forcefully stressed in this material.

NEGOTIATIONS WITH THE INSTITUTIONS

Upon receiving the formal consent of the Parole Board for the conduct of the proposed study, the investigators contacted the administrators of the two institutions which were to function as sample selection sites and received tentative assurances of cooperation. A series of meetings were then conducted with the supervisory and administrative staff members at these institutions in order to deal with the kinds of demands that this study would place upon the staff. The content of these discussions ranged from very practical concerns such as the use of the study instruments and procedures to be followed in informing the field agent of a youth's assignment to the experimental group, to the more general question concerning the extent to which the research itself might conflict with the service orientation of the staff. Generally, staff members at both institutions could agree with the proposed conduct of the
study but not unreservedly. There were still major reservations about the possible compromise of service goals.

Specifically, the sets of priorities under which the staff at the institutions operated were viewed as conflicting with the research being proposed. The practitioner concern with what Wright has called a "strategy of activity" which was manifested in providing service to youth committed to their care, was seen by some staff members as antithetical to a "strategy of research". This concern was most graphically expressed by a few who indicated that they felt the researchers were "playing dice" with human lives by randomly assigning children to experimental and control groups. This statement seemed to reflect a nagging concern over the conflict between providing services to young people and evaluating the effects of such services; the choice presumably prioritized as one of helping young people against gathering information on them. From the researcher's perspective, however, the conflict was more apparent than real. For without adequate information and testing, there would seem to be little to ensure that the service being provided was really a "service" or that the help being offered was actually helping anyone - assumptions which the practitioners adhered to in defending the service priority.

Another problem was to develop ways to gear the random selection procedures of the study design in with the administrative procedures of the two institutions without either violating the design or grossly interfering with the normal administrative functioning of the institutions. By working closely with the pertinent administrative staff at the two institutions, this problem was resolved to the mutual satisfaction of both the researchers and the staff members.

NEGOTIATIONS WITH THE FIELD STAFF

Concurrent with the meetings held with the administrative staff at the pertinent institutions and those with the Parole Board over a six-week period prior to the formal initiation of the study, meetings were arranged with all supervisory and administrative staff members in the parole division. The initial meeting held with representatives from this Division involved the researchers and the four top administrative staff members. The general nature of the proposed study as well as the procedures to be followed as those related to parole supervision were presented and discussed. Once general approval was received at this meeting, sessions were scheduled with each supervisory staff member throughout the state.

The purpose of meeting with supervisory staff members was explicitly to provide information on the proposed research and solicit reactions to the kinds of problems that could be anticipated as arising during the conduct of the study. Questions relating to the general design of the study were not a major concern at these meetings; however, questions on the specific procedures as to how experimental group members were to be handled by the individual parole officers did arise. How, for example, would agents handle the common
school requirement that all returning parolees be accompanied to school by the parole officer? Now would placement plans formulated at the institution prior to release be affected by assignment to experimental status? How would "contacts" be defined? Not unexpectedly, field also wanted a clear understanding of how they were to be notified of the youth's placement in the experimental group as well as how the family was to be notified of such placements. Finally, questions were raised regarding the amount of extra work which would be involved for the agent in handling experimental cases and how to deal with the anticipated community pressures from law enforcement, court, or school officials over not actively supervising certain parolees.

These questions required clear responses on a practical level, an expression of understanding by the investigators of the kinds of problems the study would create for these staff members, along with a clearly expressed willingness to help whenever possible in resolving difficulties. The cooperation and active assistance of the field agents, which was crucial for the conduct of the research, was not likely to be forthcoming if the researchers were seen as callously shoving problems off on them without actively involving themselves in working toward mutually satisfying solutions.

Some of the questions presented at these meetings may have been prompted by an underlying concern. While not explicitly stated as such, the thrust of the proposed investigation was likely perceived as a test of the competency of the field staff. Similarly, the roles of the evaluators seemed to be viewed as those of inspectors or management spies. The fact that the Parole Board and top administrators of the parole division had given their sanction did little to minimize the problem, in fact, their interest in this study may have exacerbated it. To help reduce tension, the investigators found it necessary to explain repeatedly who they were, what their intentions were, and what the potential implications of the research might be in terms of the role and function of the field agent. Moreover, attempts were made to underscore the idea that the program's impact was the focus of evaluation and not the individual practitioner's ability or competence. While anonymity was assured, the mere fact of conducting the study nonetheless constituted a threat by calling into question a rather basic assumption: namely, does parole supervision have an effect on the further delinquent behavior of young people?

Of great help in overcoming these sources of resistance was one of the researcher's prior work experience in the Department of Corrections, and consequently, his familiarity with many of the parole staff. Perhaps because of this, he was not perceived so much as a stranger coming into the setting from the alien world of the university but more as a former colleague who could be expected to have some appreciation of, and sympathy for, the difficulties of the parole officer role.

INVolVEMENT OF THE PAROLE BOARD IN THE CONDUCT OF THE STUDY

A number of problems directly involving the parole board arose during the course of this project. A minor but pervasive problem concerned the on-
going articulation of the random selection procedures at the institutions
with the mutual tasks of the institutional staff and the parole board members
at the time of the parole hearings.

Other concerns developed around the availability of the random assign-
ment schedules to either the members of the parole board or institutional
staff and the possible slippage of the agreed-upon study criteria. To protect
against these concerns becoming serious problems, careful monitoring by the
researchers was required on an on-going basis.

As indicated previously, the conduct of this research tended to place
local parole officers in a difficult position relative to the local police
and court authorities. This was most obviously the case in rural communities
where the local police and court tried to remain closely involved in the
general progress of a youth released back to the community on parole. On a
number of occasions the fact that a parolee was returned on an unsupervised
basis seems to have been interpreted by these local officials as a threat to
their authority over the local county probation officer and, ultimately, over
the youth he was supposed to serve.

For example, over the course of the third and fourth months of the
study, a number of letters were received from county probation officers who
had experimental group members assigned to their caseloads. These letters
tended to express concerns around the methods used in assigning youth to the
experimental group as well as the lack of involvement on the part of the local
juvenile court judge in planning and the research. Reinforcing these letters
were similar types of concern communicated by local judges directly to parole
supervisors and administrators.

These individually communicated concerns were more forcefully brought
to our attention approximately six months after initiating the study. At a
regular meeting of Juvenile Judges from throughout the State, the Chairman of
the Parole Board who "happened" to be present was questioned about the nature
of the assignments made to the study and, in particular, why the Judges had not
been informed prior to beginning the research. As a result of apparent concern
over maintaining a good working relationship with the jurists, this official
proposed that the investigators meet with each juvenile judge in the State and
inform him of the nature and purpose of the study. The implication was clear
that if further concerns were raised by the Judges, modifications would have
to be made in the procedures used for assigning youth to the study groups.
While politically expedient, this suggestion was unrealistic in view of the
excessive travel time that would be required. The message, however, was
quite clear; this was the first real indication that the Parole Board was
prepared to terminate the study in response to pressure from influential
publics.

Never completely resolved, this issue was diminished somewhat by the
researchers agreeing to provide the Judges with copies of all information pertinent to the study. Because, in most cases, they had already received this material from the local parole officer, this procedure simply amounted to providing them with an additional copy.

THE INVOLVEMENT OF THE PAROLE STAFF AND INSTITUTIONAL STAFF IN THE CONDUCT OF THE STUDY

The investigators attempted to maintain continuous contact with parole staff over the life of the project. To accomplish this, reports providing general information on the conduct of the study were sent out to all staff on a monthly basis. These reports noted the kinds of problems various parole officers encountered and pointed out possible ways in which they might be handled.

The problems that arose at the institutions during the conduct of the research were minor. These included an occasional failure of the institutional staff members to contact the designated parole officers after the parole hearing to inform them of a particular youth's placement in the experimental group and an occasional clerical error in failing to notify the family of a youth's placement in the experimental group.

SUMMARY AND CONCLUSION

Planning and conduct. The process of planning, negotiating, and conducting this evaluative research study presented a number of problems that went beyond the purely technical considerations of research design and analysis. Among the most crucial and exacting of these were integrating the 'strategy of research' concerns of practitioners and the political and administrative concerns of the Parole Division, institutional settings, and the Parole Board. The most difficult and persistent problem was thus to maintain a balance between ensuring that the research would not be severely compromised and the functioning of the various administrative components unnecessarily disrupted. Herzog, for example, has described the process of blending scientific rigor with clinical and administrative practicality as resembling an act of love "... in the fact that it is vastly written about and yet when it happens to a person it feels new, unexpected, uncharted."

In order to help avoid some of the potential pitfalls of conducting a field experiment of this magnitude, it was crucial to have a clearly articulated initial working agreement between the researchers and the administrative and service components relevant to the study. A clear understanding of the kind of constraints bearing on the research in relation to what could and what could not be done had to be specified. In turn, the constraints bearing on the clinicians and administrators had to be acknowledged and handled within the limits of the research design. Specifically, it was necessary to deal with political constraints involving groups such as the Probate and Juvenile Court Judges and policy constraints relating to the institutions and field services.
components of the Department of Corrections and the Parole Board. Without deliberate planning along with a mutual sense of trust between the relevant bodies and the investigators, the project may well have been aborted on several different occasions. Involved here was the importance of the investigators providing administrators and practitioners with assurances of their technical competency, their understanding of the agency setting, as well as their personal integrity.

Findings and utilization. A major finding of this research was that assigning cases on the basis of providing differential exposure to parole supervision does not have a significant positive effect upon the parole adjustment of juvenile-aged boys and girls. Within the group of boys, significantly more control group members as compared to their experimental counterparts had parole revoked during the ten-month follow-up. Furthermore, with regard to girls, a similar but not statistically significant tendency was apparent.

Unplanned for and unexpected findings included sex-related differentials in treatment afforded juveniles as well as the amount and perceived quality of help sought out and received by juvenile parolees in the community. With regard to the former, it was found that girls as compared to boys experienced fewer opportunities to remain within the local community either on probation status or in county-operated institutional settings and were committed to the state-level Youth Commission and subsequently placed in state operated correctional settings at a younger age and for the commission of less serious types of offenses. In terms of the provision of help within the community context, it was found that close relatives and peers as compared to more professionally defined 'helping sources' were most frequently indicated as people with whom problems were discussed and the fact of experimental status did not significantly relate to either the type of people sought out for help, the frequency they were sought out, or the extent to which they were defined as helpful.

The major conclusions were included in a summary of the project which was distributed to all parole officers, supervisors and administrators, as well as to the members of the Parole Board. In conjunction with this summary of the project findings, the investigators communicated their willingness to meet with and discuss the programmatic implications of the research. Unfortunately, no response to this offer was ever made by either the practitioners, supervisors or administrators of the parole division or by the members of the parole board. Almost totally ignored, the project became in effect, a "non-study" - left to gather dust on a shelf.

Approximately a year after the completion of the study a new administration was appointed to the Department of Corrections. An attempt was made by a new Deputy Commissioner to have the findings of the project discussed and the implications for the delivery of parole services examined. The major focus of discussion by parole supervisors soon came to be the adequacy of the methodology used. Consequently, the Deputy Commissioner contracted with an outside researcher to write a critical appraisal of the study including the methods used and the plausibility of the findings. Faced with a positive
assessment of the study, the supervisors were left with the recourse of making veiled charges of unethical behavior on the part of the researchers along with further methodological nitpicking.

Even with the strong support of the Deputy Commissioner, little more came of using the study findings for the more effective and efficient delivery of parole supervision.

Why did the findings of this project have so little direct impact on the relevant administrative units? This question is particularly intriguing in view of the excellent cooperation and high degree of trust maintained between the investigators and the administrative, supervisory and line-level parole staff during the course of the project. A primary reason may have been the initial failure on the part of the researchers and the pertinent organizational units to explore and make explicit what each party saw the purpose of the study to be, particularly in relation to functional programmatic decisions. While the researchers viewed the major purpose of the project to be a rigorous test of the validity assumption parole supervision, the relevant organizational units seemed to view the study as one of documenting what they already 'knew' to be true about the program. Negative results were neither seriously entertained by the Parole Board nor raised as a highly probable outcome by the researchers. In turn, the program implications were never dealt with. Consequently, when the negative results were demonstrated, there was no commitment on the part of the parole staff to make modifications in the manner in which services were being delivered. Added to this lack of commitment was a state of uncertainty over what changes would be required in order to more effectively achieve organizational goals. Thus, while the researchers clearly demonstrated that parole efforts were largely ineffective, the effort was of little help in charting clear alternative courses of action. Attention had not been paid in advance to contingency planning of the different ways the results conceivable could turn out as well as the ways in which each set of alternative findings could lead to modifications in the delivery of parole services. Thus, while program personnel might have been able to accept the general accuracy of the findings the substantive implications for using the results to modify service delivery patterns were ambiguous. Program people were left in the untenable position of either concluding that parole services as presently being provided were totally ineffective or that the research was in error. Being unwilling to conclude the former, they chose the latter course and, as a consequence, disregarded the conclusions which were obtained.

This research was supported by grant NI-052 provided by the Law Enforcement Assistance Administration.

A three year follow-up period in correctional research is generally held to be necessary in order to obtain a high percentage of recidivists. See; Gene Kassebaum, David Ward, Daniel Wilner, *Prison Treatment and Parole Survival*, New York, John Wiley, 1971, esp. p. 212.


Elizabeth Herzog, *Some Guide Lines for Evaluative Research*, Washington, U.S. Department of Health, Education and Welfare, 1959, p. 88. More specifically on this point with reference to correctional settings, Cressey has clearly described the conflicts faced by correctional workers in formulating their attitudes toward evaluation research and has noted that while these workers are supposed to be in favor of such research they recognize that the findings may threaten the continued existence of their program. See: Donald R. Cressey, "The Nature and Effectiveness of Correctional Techniques," *Law and Contemporary Problems*, Autumn, 1958, pp. 754-771.

10. This finding is supported by a wide variety of studies using different methodologies, designs, and involving different populations. See, for reviews of this literature; Stuart Adams, "Some Findings From Correctional Caseload Research," *Federal Probation*, December, 1967, pp. 48-57; H. J. Vetter and Reed Adams, "Effectiveness of Probation Caseload Sizes: A Review of the Empirical Literature; *Criminology*, February, 1971, pp. 333-343.

Manuscripts for the Journal should be submitted to Dr. Ralph Segalman, Department of Sociology, California State University, Northridge, California, 91324. Please submit three copies of the manuscript. The Journal welcomes a broad range of articles concerned with empirical research on social welfare, analysis of problems in social welfare, "think pieces" and others.

SPECIAL ISSUE - SPRING 1975

THEME: THE SOCIOLOGY OF ETHNICITY AND SOCIAL WELFARE

SPECIAL ISSUE EDITORS: JOAN WALLACE, HOWARD UNIVERSITY
                  SELIG RUBINROTT, UNIVERSITY OF CONNECTICUT

SEND THREE COPIES OF MANUSCRIPT TO:
JOURNAL OF SOCIOLOGY AND SOCIAL WELFARE
UNIVERSITY OF CONNECTICUT
SCHOOL OF SOCIAL WORK
1800 ASYLUM AVENUE
WEST HARTFORD, CONNECTICUT 06117

SUBSCRIPTIONS

Subscriptions to the Journal may be obtained by writing to:

Journal of Sociology and Social Welfare
University of Connecticut
School of Social Work
1800 Asylum Avenue
West Hartford, Connecticut 06117

Individual subscriptions are $7.00 per volume (4 issues)
Institutional or Library $10.00 per volume (4 issues)
Student $5.00 per volume (4 issues)

Individual copies of the Journal may be purchased at $2.50 per copy (include postage and handling).

Please include check with order.