8-1971

Evaluative Research: Toward the Implementation of a Model for Research in Law Enforcement

Thomas V. Schade
Western Michigan University

EVALUATIVE RESEARCH: TOWARD THE IMPLEMENTATION OF A MODEL FOR RESEARCH IN LAW ENFORCEMENT

by

Thomas V. Schade

A Dissertation
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Doctor of Philosophy

Western Michigan University
Kalamazoo, Michigan
August, 1971

ACKNOWLEDGEMENTS

For their encouragement of and confidence in this writer I wish to thank my entire doctoral committee, especially Dr. Charles Keely and Dr. Robert Betz. Committee Chairman Dr. Lewis Walker gave unstintingly of his time and is largely responsible for that which is noteworthy in this dissertation.

Appreciation is also expressed to Dean Fox, Chief, Kalamazoo Police Department and his administrative assistant, Paul G. Aldo. Their cooperation in completing the training school evaluation was all that one could ask.

For her meticulous care in typing and editing I am indebted to my wife, Jane. In spite of all the assistance I have received, any shortcomings in the dissertation remain entirely my own responsibility.

Thomas V. Schade

PLEASE NOTE:

Some Pages have indistinct print. Filmed as received.

UNIVERSITY MICROFILMS

TABLE OF CONTENTS

LIST OF TABLES

INTRODUCTION

The purpose of this study is to examine the nature of evaluative research and its relationship to social action programs. A case study involving the evaluation of a police training program is used to illustrate some of the problems encountered in evaluation. An existing model for the evaluation of educational programs (the CIPP model) is then applied specifically to the substantive area of law enforcement.

Law enforcement is chosen as the focus for this discussion of evaluative research due to increasing public interest in police practices. Contingent upon this interest has been the outpouring of federal and state monies for the upgrading of programs and personnel in law enforcement agencies. An integral part of these new programs is the systematic evaluation of the extent to which they meet their objectives. Current problems in the relationship between the police and the community they serve attest to the need for innovative training and practices within law enforcement. Such innovations require consistent and conclusive evaluation. To the extent that the other areas share the characteristics and problems of law enforcement, suggestions for the implementation of evaluative research in law enforcement may be useful in these areas.

Several questions are raised in the development of this topic.

Some of these questions have reference to the relationship of evaluative to non-evaluative research. Other questions focus on the choice of appropriate research designs within evaluation. Still others refer to the problems of implementing comprehensive evaluative strategies, especially in law enforcement. All of these questions are designed to highlight the nature, function and implementation of evaluative research as it relates to decision-making in social action programs.

CHAPTER I

THE NATURE AND FUNCTION OF EVALUATIVE RESEARCH

In its broadest, though not necessarily most useful meaning, evaluation can be defined as a judgmental process. All research, whether in the social or physical sciences, involves a judgmental process. Any time a theory is posited and hypotheses derived from it are tested, a judgment can be made as to whether or not that theory is supported. In this broad sense all scientific research can be said to involve evaluation and therefore have some claim to be included under the heading of evaluative research. More restricting and limiting definitions of evaluation appear in the literature, however.

An early definition equated evaluation with measurement. (See Thorndike and Hagen, 1961, and Ebel, 1965) No one would argue that measurement and evaluation are unrelated, yet it is quite a different matter to regard them as identical concepts. Such a definition gives evaluation a mechanistic focus that tends to obscure the effect of value judgments and limits evaluation to only those areas in which we have developed tools for measurement. Equating measurement with evaluation has, however, led to the use of statistics in analyzing large sample data from which broad generalizations can occasionally be made.

A second alternative definition involves the congruence between performance and objectives. (See Tyler, 1950, and Furst, 1958) Evaluation in this context is seen as the method by which one determines the extent to which the outcome of a program or innovation has met the established objectives. Objectives, therefore, are seen to be useless unless they were in measureable form or operationalized. Although the advantages of this congruence model hold at a conceptual level, they have been put into practice only with great difficulty and sometimes reluctance. The evaluation of the entire change process continued to be largely ignored, and evaluation remained a basically static, post factum technique.

The difficulty of operationalizing some objectives led to a third definition of evaluation, one which equated evaluation with professional judgment. By virtue of experience and expertise, assessment by professionals is accepted as necessary and sufficient evaluation. Accreditation associations in education most often rely on this definition.

A fourth definition has recently been suggested by Daniel Stufflebeam who defines evaluation as ". . . the process of delineating, obtaining and providing useful information for judging decision alternatives." (Stufflebeam, 1970:60) This definition stresses the judgmental aspect of evaluation and is similar to positions of other methodologists of educational evaluation. Scriven (1967) sees

evaluation as a methodological activity which combines performance data with a goal scale while Glass (1971) views evaluation as an attempt to assess the worth or social utility of a given program.

These varying definitions are difficult to sort through unless one examines the place of the judgmental process in scientific research in general. Within sociology as well as other disciplines, many types and variations of the research procedure have been embraced. Throughout the literature, a recurring theme seems to focus on the formation of a continuum to represent two polar types of research. One extreme is often referred to as basic or pure research. Although its meaning varies somewhat with the writer discussing it, there seems to be agreement that it emphasizes explicit theoretical concerns and is directed at ascertaining facts and the causal relations between facts. (Madge, 1967) At the other extreme of the continuum is the applied or action oriented research type. Here the use of theory and its attendent hypothesized relationships is less explicit. The focus is more on practical issues, with emphasis on the development of efficient means to specified ends.

It would be inappropriate to think of basic research as being either more or less evaluative in nature than applied research, if we define evaluation as simply a judgmental procedure. In basic research a theory is evaluated with respect to its ability to explain causal relationships between facts. Applied research involves the

evaluation of the efficacy and the efficiency of a strategy in achieving a specific goal. Most research in sociology probably would fall somewhere between the two extremes of basic and applied research with respect to its evaluative component. Judgments as to the explanatory power of a particular theory often have direct implications for an action program that could be structured to make use of that theory. Conversely, the evaluation of the efficiency of a particular action program often involves the implicit evaluation of the causal relationship between the facts which form the basis for the program.

Since the evaluative or judgmental process cannot be seen as the exclusive property of either basic or applied research, the justification for its treatment as a distinct and separable research form must be found elsewhere. This justification stems from the difference in the purpose and implication of judgments made. Returning to the basic/applied research continuum, judgments made in research emphasizing explicit theory testing tend to have few direct and immediate political implications. Whether or not a theory is supported is of interest primarily to the originator of the theory and perhaps his academic community. Shifting political power does not usually result from such a judgment except in a limited sense within a particular university department or professional organization. For example, if a new theory regarding the development of self-concept was supported through basic research, its author might receive a

promotion but the effects on the total community would be minimal. When, however, judgments as to the theory's effectiveness in improving academic performance in the public schools are made, the political ramifications of such an evaluation could be considerable. Those supporting old, less effective procedures in the schools might lose previously held power. New leaders might arise on the basis of their early support for the now established self-concept approach. Pressures from proponents of alternative approaches of education to minimize the effectiveness of the self-concept approach might be felt by those engaged in continuing evaluation. (See Cohen, 1970 for a more complete discussion of these problems.)

Another distinction between the nature of judgments made in basic versus applied research is suggested by Wrightstone. (1969) Judgments in basic research are determined by a flexible timetable. A judgment is made only when the most sophisticated of available techniques of analysis show the datum to support the theory. This may take several months or many years to accomplish. In applied research the judgments as to effectiveness of programs must often be made according to a rigid timetable. The datum may be incomplete and there may be insufficient time for complete analysis. Although the researcher may still be concerned with the basic theory, his primary function is often to appraise a practical activity while meeting an administrative deadline.

A third distinction regarding the nature of judgments in basic and applied research is the impetus behind the need for a judgment to be made. Cherns (1969) suggests that the need for judgments in basic research often arises from within the discipline out of which the theory originates. The desire to establish laws explaining the relationship among variables appropriate to a particular discipline requires that judgments as to the support for these theories be made. In applied research, judgments are stimulated by needs within some organizational framework. The purpose of these judgments is to make decisions regarding planned change within the organization. Both situations involve judgment or evaluation, but for different reasons. In basic research the potential for generality is usually great, but the potential for utilization in practical situations is usually limited. In applied research the exact opposite is most frequently the case.

A fourth distinction in the purpose and implication of judgments embodied in applied as opposed to basic research is the role played by the organization in which the research is taking place. In most basic research settings that involve the testing of an explicit theory in an outside organization, the concern of the researcher would be determined by his theory rather than the needs of the organization for information. The organization becomes an arena for the testing of the researcher's hypothesis. In applied research, with its

emphasis on the development of efficient means to achieve specified ends, the needs or objectives of the organization would be of greater importance. Both types of research utilize theory and hypotheses, but in differing forms. In applied research the theory and hypothesis are implicit. The objectives of the program or organization being studied provide the major focus for applied research. This focus is provided for basic research through the use of explicit theory and hypotheses. The judgmental process engaged in by the applied researcher actively involves some level of participation by the administrators of the organization being studied. The basic researcher will, of course, require a certain amount of cooperation from the administrators, but will probably not be as much concerned with their participation in the judgmental process.

The purpose and implication of the judgmental process when undertaken within the framework of applied research as specified above begins to delineate that type of research known as evaluative research. This is not to say that all applied research is evaluative research. Other forms of applied research may focus on the cause or incidence of a particular behavior that need not involve either the purpose or implication of the judgmental process as discussed above. For example, Caro (1971) suggests that applied research may contribute to social action without assessing the effects of specific interventions. If one conceives of basic and applied

research as being simply opposite extremes of a single continuum, then evaluative research may sometimes also share the characteristics of basic research. Generally, however, evaluative research is closely related to applied research.

It was previously stated that the justification for separating evaluative from non-evaluative research rests on the differing purpose and implication of the judgmental process. Throughout the four distinctions discussed above, the judgmental process corresponding with applied research can be seen to have a distinctive purpose--to facilitate the decision-making process engaged in by someone other than the researcher. Both basic and applied researchers are constantly involved in decision making. But decision making by the researcher himself does not distinguish evaluative from non-evaluative research. The type of evaluative research that this study is concerned with is that research which has the intended purpose of facilitating an administrative decision apart from internal research decisions. The implications of the judgmental process in evaluative research will be more likely to involve external political consequences, operations based on fixed timetables, questions of specific utility and the active participation by outside administrators in all phases of the research.

The discussion of both the purpose and implications of the judgmental process in evaluative research is not intended to suggest

that a set of rigid criteria have been established for evaluative research. It is more accurate to think of a specific research design as being either more or less evaluative as determined by the purpose and implications of its judgmental process.

Before exploring alternative models for research characterized by evaluative considerations several aspects of evaluation should be considered. First, the issue of inside as opposed to outside evaluation (or "in-house" and "out-house" evaluation) raises important operational questions for the researcher. What are the advantages of one approach over the other? According to Caro (1971), outside evaluation will tend to insure greater objectivity on the part of the researcher. Although subject to political and administrative pressures, the outside researcher will not have as great a personal stake in the outcome as would a member of the organization. The outside researcher is also more likely to suggest criteria for evaluation that question organizational premises. Simply raising these questions could be beneficial to the organization. In addition, the outside researcher is in a better position to moderate any internal conflicts that may arise during the evaluation. He is better able to avoid problems of status incongruity and can remain free from unwanted and often time consuming tasks. On the other hand, it can be argued that the inside researcher will have several advantages including a more detailed knowledge of the organization and its programs than

an outsider. This would be an increasing advantage as the program becomes more complex. Also, the inside researcher is in a better position to do continuing evaluation of a program. Each research situation will have to be examined individually with regard to the best approach for its particular objectives.

Many frustrations are likely to be experienced by those engaged in evaluative research. One may feel that agreement to work for an organization implies that the researcher accepts all the goals of that organization. He may, in fact, question or be in disagreement with the goals but continue to assist the organization in evaluation. This dilemma must be resolved by each researcher in each evaluational setting.

The dissemination of results may present problems to the researcher. An organization may wish to restrict the public knowledge of its programs. An inside evaluator would no doubt have to abide by such restrictions or look for work elsewhere. An outsider has the option open to him of refusing to carry out the research if the restrictions are seen as unwarranted. Many researchers, including the present writer, would be very reluctant to engage in evaluation in which the dissemination of findings was in any way restricted.

Methodological frustrations might also face the evaluator. As mentioned previously, the restricted time in which an evaluation

might have to be undertaken could rule out the use of the most effective research design. For example, longitudinal studies are often preferable in assessing the impact of some planned change. Organizational decisions, however, often need to be made "yesterday" and sometimes restrict the use of extended designs. Putting complex statistical interpretations into language understandable to administrators and useful in their decision-making functions can also be a difficult job for the researcher.

Because of the above considerations, several models for the structuring of evaluative research appear in the literature.[1] Scriven (1967), although not suggesting an actual model, has made a useful distinction within evaluative research which has had implications for other models. He suggests that evaluation is either formative or summative. Formative evaluation is designed to improve a program while it is still active. It would involve an on-going process that would lead to modifications of an existing program to insure that it would better meet its objectives. Summative evaluation is designed to appraise a product after it has been established. Summative evaluation has been the form traditionally chosen by researchers.

---

[1]Evaluative models outside of education such as PERT (Program Evaluation and Review Technique, Quade, 1967) used in the development of military armament and the various management and production models from industry will not be discussed as their utility in social action programs is highly specialized.

Such a distinction raises again the question of purpose within evaluative research. Does either summative or formative evaluation serve decision-making functions only following the completion of a program, or should it also serve other functions?

Caro (1971) suggests that there is a planning-action-evaluation cycle. This would seem to imply that he views evaluation as being appropriate only after the action phase and therefore not concerned with the planning or the implementation of a program. Such a view would seem to follow Scriven's conception of summative evaluation.

Suchman (1966), however, begins to combine formative with summative evaluation as he suggests a four component model for evaluation involving effort, effect, process and efficiency. Effort refers to the amount of action undertaken in a program. Effect concerns itself with the results of effort and is similar to a summative phase. Process is concerned with how the effect was achieved, and is similar to a formative phase. Efficiency takes into consideration the cost factor. Suchman's approach, although perhaps not technically a model in that it does not provide specific steps to be followed, does begin to give direction to the design of evaluative research.

Based on the work of Greenberg (1968), Suchman (1967) and Scriven (1967) a model for evaluative research has been developed by Stufflebeam (1970). It is his contention that evaluation has, in essence, a single purpose--to facilitate the administrative decision-

making process. This writer is in general agreement with Stufflebeam's position with the clarification that the decision-making process being served is in addition to the internal decision-making aspect characteristic of most research. Therefore, when the decision-making process is referred to in connection with evaluative research, it will denote an external administrative process.

Stufflebeam's model combines the summative and formative aspects of evaluation discussed earlier, and introduces several new dimensions. The model is comprised of four stages. The first is context evaluation to be used when a program is in its planning stages. Second is input evaluation to determine the resources and procedural designs available to achieve the program goals. Third is process evaluation to provide continuous feedback to program managers in identifying weaknesses in procedure and implementation. Fourth is product evaluation to determine the overall effectiveness of the program with respect to the program's goals as previously specified through context evaluation. It should be made clear that the evaluator does not, by himself, decide upon the goals to be achieved through a particular program. He assists in their formation in order that the extent to which they are achieved might be better evaluated.

Each of the four stages of evaluation is based on a stage of the decision-making process. Context evaluation serves the planning stage of the decision-making process. In the planning stage the

objectives of the program being evaluated are arrived at jointly by the evaluator and administrator. It is the administrator's primary responsibility to set the program's goals. The evaluator's primary responsibility is to see that these goals are operationalized. These two functions may overlap, however, with each contributing to the other's area of responsibility. Context evaluation also serves the planning function by identifying both the organizational needs and the "status quo" in order to establish a baseline for future stages of evaluation.

Input evaluation considers the resources available to the organization for achieving their objectives. Input evaluation serves the structuring stage of the decision-making process. Financial considerations, talent within the organization and other sources of input to facilitate the structuring process would be provided. The selection of an appropriate research design for the process and product phases of evaluation would also be undertaken to provide information for structuring purposes.

The process phase of evaluation serves the implementing stage of the decision-making process. If the preceding phases of evaluation have been carefully carried out, then process evaluation need only record the implementation of the program and the immediate reactions of the participants. Sometimes important administrative decisions need to be made regarding changes in the structure of the

program. Part of the function of process evaluation would be to provide administrators with the relevant information to make such decisions.

The product phase of evaluation serves the recycling stage of the decision-making process. This is the stage which has received the most attention in the past, and to which many attempts at evaluation are limited. This is not surprising since this is the most useful piece of information from the administrator's perspective. The utility and accuracy of product evaluation depend upon the care with which previous stages were completed. Assuming that such care was taken, what is then involved in product evaluation? Stufflebeam is quite specific in suggesting the use of two types of criteria against which the product can be judged. These are instrumental and consequential criteria.

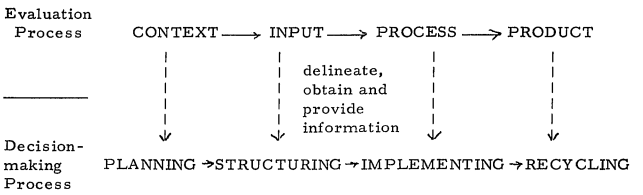
The first, instrumental criteria, are found at intermediate stages of the program. The instrumental criteria roughly correspond with the formative aspect of evaluation as suggested earlier by Scriven (1967). The consequential criteria are the fundamental baselines established in the context phase against which the final objectives of the program are judged. The consequential criteria correspond with Scriven's summative aspect of evaluation. Both of these sets of criteria are evaluated as products of the program, and serve recycling decisions.

Remembering that Stufflebeam's definition of evaluation emphasized the processes of delineating, obtaining and providing useful information to administrators, these steps are to be followed within each of the four types of evaluation (context, input, process and product). Based on this model, the total process would therefore potentially involve twelve steps--from delineating information in the context stage to providing information in the product stage. Following the initial context evaluation, however, the consideration of the degree of change desired with respect to the level of relevant information available would determine the need for consequent input, process and product evaluation. Without specifying all of the change settings suggested by Stufflebeam (1970:155-163), their common premise is that the greater the extent of desired change, the greater the need for input, process and product evaluation--especially when the level of relevant information required by the decision makers is maximal.

In order to more clearly show the relationships among the processes and concepts introduced in the preceding pages, the following information is presented in Figure 1.

The evaluation process is seen as serving the various decision-making functions. As the planning decisions eventuate in recycling decisions, so too, the context evaluation eventuates in produce evaluation which must then be evaluated as to its impact on the altered system. The key to understanding this rather complex evaluation

FIGURE 1



model is simply that all evaluation processes are designed to delineate, obtain and provide relevant information at various stages in the decision-making process.

Within this basic evaluational model (CIPP) a variety of specific techniques may be employed. Experimental design, for example, has occupied a hallowed place in the methodology of both traditional and evaluative research. As circumstances may prohibit precise control over the variables involved, quasi-experimental research designs have been developed to approximate the "ideal" classical experimental design. (Campbell and Stanley, 1966) Such factors as reliability, validity, statistical regression, instrumentalization. etc. are intermediate designs that are appropriately dealt with following the establishment of the external components of the CIPP model. The internal factors (the choice of specific statistics for use within the model) are one step further removed from the general CIPP model being review-

ed. This is not to say that the choice of one's statistics can be made capriciously, but simply that all the contingencies which affect such technical factors are not within the purview of this discussion. To reiterate, the major concern here is to consider the appropriate designs to serve the decision-making needs.

In serving the decision-makers, the evaluator observes and records events (quantifying where possible) logically organizes this information according to a conceptual framework and at least implicitly predicts that certain regularities of behavior will exist under specified conditions. These predictions take the form of statements as to whether the behavior under study will be appropriate to the established goals. Throughout the context, input and process phases of evaluation suggestions as to modification of the program can be made to further insure that the observed behavior will eventuate in the desired goal. Simple product evaluation is terminal and post factum in nature and therefore of limited feedback use. This does not negate its effectiveness, but does limit its utility for the specific program from which it emerges.

The thrust of all evaluative research is not in the direction of traditional hypothesis testing and theory verification, but rather it is utilized as a necessary adjunct to the decision-making process. As such, it is predictive in the sense that suggestions for modifications in a plan of action are implicitly predictive of specific outcomes at

various stages in the program. As these outcomes are achieved, the program continues as planned; when the expected (predicted) outcomes are not achieved, alternative approaches are suggested to and implemented by the appropriate decision-makers. Consequently there are three options available to the decision-makers as evaluation reports are continually being received; first, the program can be continued as originally conceived, second, it can be modified to better achieve its stated goals, and third, it can be discontinued due to the low probability of its ever achieving its stated goals. Therefore, besides the determination of effectiveness through product evaluation, evaluative research (the CIPP model in particular) can and should have equally important utility by means of immediate and direct feedback to an on-going program in order to facilitate administrative decisions as to the continuance, modification or discontinuance of that program.

In summary, evaluative research should be of utility to administrators in that it serves each stage of the decision-making process from planning decisions through recycling decisions. Evaluative research will tend to be characterized by specific utility for and active participation by staff administrators, operations based on fixed timetables and the likelihood of generating external political controversy.

In the next chapter the use of evaluative research within the substantive area of law enforcement will be examined. The

structural characteristics of law enforcement agencies as they might effect evaluation design and the evaluation process will be delineated.

CHAPTER II

THE IMPLEMENTATION OF EVALUATIVE RESEARCH IN LAW ENFORCEMENT AGENCIES

Within law enforcement agencies in the United States (and, indeed, around the world) evaluation of police practices is, at best, rudimentary. In many instances they are non-existent, especially in small, sparsely populated areas. There are no doubt a variety of factors which have singularly or corporately combined to inhibit the growth of sophistication of law enforcement evaluation. It appears that some of the factors include the mutual apprehension and distrust between academia and law enforcement, the confusion as to the objectives of law enforcement, and the history of reliance on intuition, experience and professional judgment. These and other factors will be examined more extensively in Chapter III, but are introduced here to help establish the background against which evaluative research among law enforcement agencies will be examined.

Although varying sources, forms of reporting, and contradictory interpretations are sometimes confusing, there is a consensus among law enforcement personnel that problems and confusion in all spheres of criminal justice are increasing. Not only law enforcement, but the judicial and correctional fields are also experiencing

a kind of anomie or norm confusion with respect to public expectations. The entire spectrum of law enforcement from federal agencies to township constables is being pushed and pulled in opposing directions. One element of the community focuses its concern on the accused--his rights and constitutional guarantees of freedom from punishment until proven guilty. The entire investigative and apprehension process is judged on the criterion of respect for the individual rights of all concerned. Another part of the community, although not necessarily ignoring or rejecting the sanctity of individual rights, chooses to emphasize the investigative and apprehension phases of law enforcement with respect to efficiency, speed and the success of eventual prosecution. These criteria have popularly been combined and labeled as "law and order" concerns. These often opposing standards by which one can judge law enforcement philosophies and techniques have led to the confusion surrounding any attempt to evaluate specific programs.

Essentially, this study is an attempt to focus on the problems involved in the evaluation of community relations training programs within law enforcement as one example of a social action program. With respect to the evaluative process several questions arise that are of particular relevance to police-community relations. Are the objectives of training programs similarly defined by the police department and the community? Is the most appropriate criterion

for assessing change attitudinal, behavioral, or both? Are current approaches and research designs sufficient for the evaluation of such training programs? Are there structural factors found in law enforcement agencies that require specifically constructed designs for evaluation? What evaluation model will provide the necessary information for administrative decision-making. Finally, how is evaluation in law enforcement related to broader issues in both traditional and evaluative research?

In attempting any form of research within a law enforcement agency, certain characteristics of that agency may present unique problems to the researcher. The research design is supposed to facilitate the research process as a route guide aids the motorist. The route drawn on the map may be misleading, however. If time is a primary consideration, the quickest route will probably be assumed to follow the shortest distance between the point of origin and the destination. This would not be the case, however, if the most direct route was a crowded, two-lane secondary road under construction. Perhaps the four-lane expressway by-pass, although less direct and ten miles longer, would be the most efficient route. Therefore, prior to embarking on a trip, the motorist should be aware of the specific terrain to be covered and the condition of the local roads to be travelled. Similarly, the naive adherence to a general research design without consideration of the particular properties of the

population and setting under study could present serious roadblocks to the researcher.

By the very nature of its function in the community, the law enforcement agency must keep most aspects of its operation closed to the public. Some departments do allow tours through the facilities, run public school educational programs or participate in question and answer sessions at public meetings. Such activities do not, however, permit the public to have much access to the internal workings of the agency. Also, researchers have traditionally found it difficult to gain the confidence of law enforcement officials. An outsider is typically seen as a threat to the system. As stated by Niederhoffer and Blumberg, "the police, prosecution and criminal courts represent a closed community manifesting a defensive attitude toward 'outsiders' and 'critics'." (Niederhoffer and Blumberg, 1970:6)

What has led to the development of a closed community and its accompanying suspicion of most non-members? In discussing secrecy and the police, William Westley has suggested a plausible explanation. (Westley, 1950) Group identity and solidarity is maintained through the establishment of confidentiality as a part of the role expectations held by fellow officers. Those who are privy to certain secrets are bound together through their common possession of restricted information. This then serves the function of differentiating them from those who do not hold or have access to such

information. As a result, an in-group/out-group dichotomy is formed, and any leak of information to the out-group (the civilian public) is seen as a breach of security by the in-group (the police). A positive, collective image is thereby formed for police officers, and any deviation from the secrecy norms is both personally and collectively threatening. The result is a subculture which Westley characterizes as follows:

> The exigencies of the occupation form the police into a social group which tends to be in conflict with and isolated from the community; and in which the norms are independent of the community. (Westley, 1951:294)

The tendency toward the establishment of a subculture and its attendant norms of secrecy takes on added significance when officers' perceptions of community attitudes toward them are analyzed.

The perception of police officers as to the attitudes toward them held by members of the community may influence the extent of officer suspicion and hostility toward researchers and other outsiders. Research into the accuracy of police officer perceptions of attitudes toward them and their role in the community as held by inner-city youth in a mid-western city, has shown a strong tendency toward negative misperception. (Schade, 1970) In almost every case, the youth attitudes toward the police were demonstrated to be much more positive with respect to the officer and his job than the officers expected. The officers were consistently inaccurate in

their perception of youth attitudes, and with few exceptions believed the youth to be far more hostile and negative toward the police role than they actually were. Believing, however accurately, that the public is negatively evaluating them and their work, the police officer sees secrecy as necessary to the successful performance of his duty. Any civilian who either figuratively or literally passes beyond the "Authorized Personnel Only" barrier is perceived to be hostile and thereby a potential threat to the in-group and their secrets.

This suspicion of outsiders is not without some foundation. No doubt some researchers who have gained the confidence of particular officers and departments have violated their established trusts. It would serve to further reinforce the belief that civilians are not to be trusted. The cloak of integrity that may have once permitted a professor or other professional researcher to gain access to police secrets is losing its power to gain entrance. A researcher must depend more and more on personal friendships with influential officials. Being contingent upon friendships, however, results in an often tenuous relationship with the agency.

The preceding discussion has been based on the assumption that the agency is being approached by the researcher and asked to give up much while receiving little in return. This has been a frequent relationship established between the researcher and law enforce-

ment (and other) agencies. In the last few years, this relationship has begun to shift. The rapid increase in public funding of law enforcement programs has brought with it a need for systematic evaluation for accountability. Police agencies are not usually staffed with personnel capable of carrying out complex evaluation programs. Even if they were, some state and federally funded programs require external evaluation for the expressed purpose of better insuring objectivity. Consequently, the evaluation researcher is now being sought after by the law enforcement agency as has been the case in education for the past decade.

Due to the structural and subcultural characteristics of the police, the researcher, in being asked to enter a particular organization, should require a clear statement of restrictions on access to certain information, personnel, records or even areas of a building. The extent of these restrictions should be determined not by the whim of a commanding officer, but by the freedom necessary for the evaluator to effectively provide his service to the department. In the role of invited researcher rather than that of interloper, the evaluator is more likely to meet with willing cooperation from administrators.

Although access to sensitive areas may be greater and administrative restrictions fewer, does the evaluator compromise his academic or scientific freedoms through his "acquiescence" to

administrators? By being in the employ of a particular agency, the evaluator is expected to produce specific information that will be of use to that agency in deciding whether to continue, modify or discard the program under study. In the context, input and process phases of evaluation as earlier discussed, there is almost unlimited opportunity to call to the attention of administrators factors which they might not be aware of relating to their program. This would seem to be an important contribution that could be made by the sociologist as evaluator. Individuals trained only in the methodology of evaluation would have little to contribute of a substantive nature. In addition, the sociologist has the opportunity to make use of his observations of on-going social processes through contributions to his discipline within the "context of discovery." (Phillips, 1966:56) His specific research focus would be determined by administrative decision-makers, but this need not involve compromise in its negative connotations. Compromise can also result in gains to both parties. Each situation would have to be assessed individually in order to weigh the administrative restrictions on the research focus against the opportunities for broader contributions to both the agency and the discipline.

Related to these potential restrictions are those involving dissemination of the research findings. A particular agency may require that the results of any research involving its personnel and

practices be kept within the agency. Publication of results in professional journals, even in disguised form, might also be prohibited. In law enforcement such factors as the chief's job security, the extent of professional provincialism, the state of local police-community relations and, of course, the overall light in which the department is cast would probably have to be considered. Other bureaucratic organizations may present similar problems. Each instance of administrative restriction has to be considered in the context of its own set of circumstances. No formula could possibly account for the variety of factors one might encounter.

In the last decade the subject of law enforcement in general and the policeman in particular has become of increasing importance to journalists, sociologists and policemen themselves. Whittemore spent many hours with the police of three cities trying to get a better understanding of the feelings of individual officers as they react to various situations. (Whittemore, 1969) Such a perspective has not been forthcoming from the academic community. It might be argued that it took a young journalist to do what sociologists should have been doing many years ago. When sociologists have become involved with the functioning and problems of law enforcement they have done a creditable job. Skolnick, as reported in Justice Without Trial: Law Enforcement in Democratic Society, spent several years utilizing questionnaires, interviews, participant observation and many

other research techniques to bring to bear the socio-legal perspective on the day-to-day activities of several metropolitan police departments in the far West. (Skolnick, 1966) In his empirical study of police cynicism, Arthur Niederhoffer has drawn upon his twenty-one years of police experience together with formal training in sociology to look at specific consequences of recruit training and continuing peer socialization. (Niederhoffer, 1967)

The broader questions of justice and crime are shown to be of vital interest to Americans through the popularity of books ranging from The Godfather by Mario Puzo, to Crime in America by Ramsey Clark.

All of this interest in law enforcement has not, however, generated the quantity or sophistication of evaluative research within law enforcement that one would expect. Although this relatively recent national interest in the problems of law enforcement (as highlighted by the Omnibus Crime Bill passed by the U. S. Congress in 1968) has generated many innovative police programs, the relative success or failure of these programs has generally been left to an educated guess. It can be argued that our techniques of evaluation are all of the "guess" variety which are simply more or less educated than other forms of guesses. Such a semantic argument is of little value. The important question is whether or not we are using all that our collective education can provide with respect to insightful perspectives and pragmatic

techniques.

This does not seem to have been the case in the past. In looking through issues of professional police journals such as Police Chief and Police, the past five years have produced only a few articles related to the evaluation of training programs. In 1966, a two week delinquency control training program was given through the Institute of Governmental Affairs--University of Wisconsin Extension. (Owens and Hamann, 1966) A written exam covering course content was administered both before and after the course. The pre-test was used to determine what needed to be emphasized in the program to follow. The post-test was designed to reveal the content areas that were insufficiently covered. Presumably the instruction or curriculum would consequently be revised in future programs to more effectively cover the areas demonstrated through the post-test to be weak. The diagnostic and feedback functions of the research model, though rudimentary, are admirable, if they were in fact utilized. It is often the case, however, that such diagnoses are misinterpreted and feedback is ignored. The research is nonetheless important given the dearth of competing models in use.

In a more recent article entitled "Federal Funding for Police Staff Development and Training Programs," Albert Roberts suggests that three major questions must be considered by evaluators of law enforcement training programs. (Roberts, 1970)

1. Do trainees learn?
2. Do trainees put their learning into practice?
3. Do trainees who put their learning into practice have better results in both crises intervention and repression of crime?

At first these questions may seem very elemental. The first question is elemental in that the objective of any training program or educational experience certainly is to facilitate learning. But the second and third questions indicate what is often assumed to be a consequence of learning which is too often untested. Is learning put into practice? Is practice effective in achieving the desired goals in the community? The second question extends the research interest from the covert attitudinal dimension to the overt behavioral extension of those attitudes. The third question then asks if the attainment of two specific goals, crisis intervention and repression of crime, are facilitated by an officer's acquisition of the desired behavior patterns. It is a major improvement over the simple pre-test post-test design limited to the attitudinal dimension.

The extent to which Roberts' ideas have been put into practice by law enforcement agencies is minimal. In addition, the attention given to the area of police-community relations is even more discouraging. As Gallati reported in 1964, the types of research engaged in by most departments are of questionable value.

> A recent issue of a police journal indicated some of the types of research that are being done by various departments--such as the length of a nightstick; the use of ripple-soled shoes; the quality of decals for police car insignia. (Brandstatter and Radelet, 1968:399)

Given the energy devoted to these kinds of problems, it is not surprising that more has not been accomplished in the development of more sophisticated approaches to substantive problem areas such as community relations. In a sourcebook of police-community relations literature, two out of sixty-five articles deal with planning and research. (Brandstatter and Radelet, 1968) The thrust of both of these articles is that, even in major cities, little, if anything, is being done in an attempt to evaluate the current state of police-community relations. In a book devoted entirely to community relations as they relate to riot prevention, Momboisse examines every conceivable aspect of the police role as it relates to the cause and prevention of riots. (Momboisse, 1967) Even the weather is discussed at some length as an important initial cause of riots. Throughout the book, however, there is no mention whatsoever of any evaluation of the effectiveness of the preventative programs discussed.

The absence of evaluation within law enforcement is particularly noticeable in community relations training. In a program involving officers from the city of Detroit as reported in a document prepared for the President's Commission on Law Enforcement and Administration of Justice, the following statement is made:

> An evaluation scheme per se was not built into the project, and therefore, classification as a success or failure is not feasible . . . . (Survey V, 1967:317)

Such a statement is almost ludicrous in its indictment of those

administering the program. Of what value is any project if its success or failure cannot be determined? It is no wonder that the cry for accountability has arisen. As has been stated in the preceding paragraphs, this lack of evaluation is widespread, and results at least in part from ignorance of appropriate techniques for evaluation. With reference to national and local institutes on community relations, the Commission report asks and answers the vital question:

> How can those various types and levels of Institutes be evaluated from a training standpoint? In no case have sophisticated, scientific techniques for evaluation been employed, not even in the National Institute, e.g., as to degree of attitude change among participants, etc. (Survey V, 1967:293)

Should, then, these programs be abandoned? If we cannot or will not evaluate them, perhaps the resources expended on them would be better used elsewhere. If we are, in fact, unable to establish success criteria and determine the extent of their achievement, then the programs should be dropped. However, this is not the case. We do have the knowledge and it is simply a matter of applying it. The President's Task Force in police-community relations takes the same position. It states:

> These programs should then continue, but greater care should be taken of the Institute-type program, subject to the conditions and especially with respect to the need for more systematic scientific evaluations at a fundamental level of police performance. (Survey V, 1967:296)

In summary, law enforcement agencies present an imposing and only semi-penetrable exterior which is difficult for non-police

personnel to pass through. The secrecy, suspicion and restrictions facing an evaluator in many law enforcement agencies should be understood as to their origins and functions for the police. The lack of any substantial evaluation of police training, especially in the area of community relations, may be a function of both the difficulties encountered in entering the police establishment and the lack of attention to research techniques appropriate to police problems. The growing national concern over police-community relations has not been followed by systematic attempts at evaluating innovative programs designed to improve relations. The following chapters will view a community relations training school as a case in point to discuss what attempts at evaluation were undertaken, and additional procedures which might have been used.

CHAPTER III

THE EVALUATION OF A COMMUNITY RELATIONS TRAINING SCHOOL: A CASE STUDY

The research to be reported on in this chapter was an attempt to evaluate an in-service community relations training school for selected members of a municipal police department located in a moderate sized mid-western city. The opportunity to conduct this research came about at the request of the police chief in whose department the training school was held. In consultation with the training school director (a sociology professor at a nearby university) the chief outlined seven major objectives of the school. They were as follows:

1. That officers should be aware of both the biological and social meanings of "race."
2. That officers should be aware of the historical background and contributions of the black man in America.
3. That officers should be aware of the relationship of poverty to crime.
4. That officers should be aware of the basis for prejudice and discrimination.
5. That officers should be aware of the meanings and implica-

tions of "black power" and militancy.

6. That officers should be aware of the meaning and extent of racism.

7. That officers should be aware of the legal and social bases for the protection of individual rights.

Given the objectives of the program, the school was divided into two groups of twenty officers each.[1] Officers were randomly assigned to one of the two groups after taking into account work schedules, vacations, etc. In each group of twenty officers, fourteen were from the patrol division, four from the traffic division and two from the criminal division. Of all the officers participating, thirty-eight held the rank of patrolman while two were detectives. No supervisory or auxiliary personnel were included. Those selected were from divisions whose work demanded extensive public contact including traffic and district (neighborhood) patrolmen. Most of the trainees had between one and five years of experience in the department. It was felt by the police chief that less experienced and younger men would be more receptive to the goals of the school than older officers.

---

[1]Several months prior to the operation of the school under study, a similar demonstration school was held in the same department involving twenty additional officers. The data from this school are incomplete and have not been utilized in this report..

The two groups, hereafter known as School A and School B, participated in a six-week in-service training program. Two and one-half hour meetings were held twice a week for two weeks. This was followed by two weeks of independent assigned reading. For the final two weeks the twice weekly meetings were resumed. School A began two weeks prior to School B, allowing for a continuous eight weeks of presentation by the director and his staff.

For the first half of each evening session a formal presentation was made to the trainees by an expert from the community. These presentations were each focused on one of the seven major objectives specified previously. The speakers, both black and white, represented the fields of law, education, employment and civil rights, among others. Following the initial presentation, an open "round-table discussion" took place. It usually focused on the evening's speaker, but often was broadened to include topics of immediate concern to the trainees.

At the beginning of both schools, three questionnaires were administered to and completed by the patrolmen. (See Appendix B, pp. 107-113. Following the schools, the same questionnaires were administered, with only the order of the items having been changed. The data derived from these questionnaires formed the major basis for the evaluation of the training program. Before discussing the rationale used to construct the questionnaires and establish the

operational specifications for the department's objectives, several of the impressions gathered by the researcher will be presented. It is hoped that through these impressions the reader will be able to gain "a feel" for both the tone of the sessions and some of the trainees' intense reactions.

The researcher attended all of the night meetings of both schools, and met informally with some of the officers at a near-by bar on several occasions. During the first few meetings of each school all of the trainees were highly suspicious of the activities. When asked to complete the pre-test questionnaire, the officers responded with a great deal of good natured yet genuine concern regarding the guarantee of anonymity of respondents. A rather elaborate numbering system had been devised whereby each trainee was confidentially assigned a number by a command officer. The trainee was to use this number on all questionnaires. This procedure would enable the researcher to identify each individual's pre-test and post-test responses yet not know the respondents identity by name. Moreover, the department would not have access to the data. It took hours of talking to convince the trainees that their responses would not be available to the department for use in considering promotions, discipline, etc. The institutionalized suspicion discussed earlier was abundantly evident in this context.

Throughout the training school, open discussion of departmental

and personal grievances was welcomed. Such an opportunity seemed to have a cathartic effect on the officers. After the initial fears of surveillance were allayed (which at one point included an officer's personal inspection of a two-way mirror which faced the classroom) dissatisfaction with the methods of supervision and distrust of the civilian public, especially minority groups, were freely discussed. These patrolmen were acutely aware of the lack of means for upward communication within the department. Hiring and promotion procedures were questioned. Department discipline was seen as arbitrary and inconsistent.

General public apathy complicated by open hostility from the black and student communities was perceived to be directed against the officers personally. During the bar room discussions, which would last until two or three o'clock in the morning, some frightening stories were related to the researcher. One officer was told he would receive a commendation for a particularly heroic act. Three years later he had received no further word about it. The juggling of promotional exam scores to satisfy political pressures, the recording by command officers of misconduct that an officer was unaware of until it was used against him, and other equally disturbing practices were also reported to the evaluator.

The area of police-community relations seemed to be of great interest to the officers apart from their required participation in the

training school. A point repeatedly made by the officers was the lack of any formal program designed to inform or "train" the public in the functions and responsibilities of the police department. If, it was reasoned, police-community relations was a "two-way street," then the community also needed to be educated and advised of the rules under which the game was being played. An officer told of an encounter with a young, bearded man from the university in the city. After listening to several speakers talk about the unfairness of prejudging and stereotyping, the officer decided to try and overcome the animosity he had developed toward "hippies" over the years. The next day he stopped this young man's car for a minor moving violation. Upon noting his appearance, his first instinct was to ticket him and try to discover any other violations for which the fellow could be written up. Instead, remembering the training sessions, he simply gave the young man a good natured warning, and they parted friends. Two hours later the officer observed the same driver involved in the same violation and the results were predictable. These kinds of frustrations were reported numerous times, and need to be recognized when attemptying to understand a situation in its entirety.

It is interesting and worthy to note that, concerning black Americans, the most commonly made statements were to the effect that blacks in this city had an equal if not better opportunity to get work. It was assumed by most trainees that a black family could live any-

where in the city that their income would allow them to live. The admitted sins of the past committed by the white society in their treatment of blacks were seen as irrelevant to the complaints of blacks today in this city. "Why do they (blacks) take it out on us?" asked many officers. "We're not the police in Selma, Alabama, or Detroit, Michigan." Neither the historical perspective or the national implications of current racial confrontations were perceived as applicable to their own experience.

The preceding paragraphs are not intended to be definitive of the officers' attitudes regarding their community. They are included as an important adjunct to the highly structured yet limited information gained from the formal questionnaires. Although adaptable to more sophisticated analysis, a total reliance on impersonal data gathering techniques cannot help but ignore the subtleties of human communication revealed through direct interaction. This aspect of data collection as it relates to evaluative research will be discussed more specifically in Chapter Four.

To return to the seven school objectives mentioned previously, the specification of these objectives is perhaps the single most critical procedure in evaluative research. It was agreed by the police chief, the school director and evaluator that the simple retention of material presented at the training sessions was not the objective. "Awareness" as used in the statement of each objective was opera-

tionalized to include three levels of meaning. These levels are the cognitive (beliefs), affective (emotional) and behavioral (action tendency). Oppenheim, 1966:106) The cognitive dimension embodies the rational organization of knowledge about a particular idea, event or other stimulus into a belief pertaining to that stimulus. This belief is perceived to have been arrived at through reason by the individual holding it. The affective dimension of attitudes embodies the more emotional response to the stimulus. The behavioral dimension (or more accurately an extension of the cognitive and affective components) is one step closer to the overt behavioral response to a stimulus. It is the individual's conception of how he would in fact overtly behave in response to the perceived stimulus. In a sense, these three levels can be seen as forming the transition between that which is known but uninterpreted and consequent overt behavior. The attitudinal complex therefore gives meaning to knowledge and direction to behavior.

Awareness, as used in the statement of objectives, has therefore been specified as being comprised of this three part attitudinal complex. The retention of interpreted knowledge of the material presented during the training school was not measured directly. The form of presentation, as discussed previously, stressed interpretation and meaning rather than rote memorization of facts. The items chosen to measure the attitudinal complex are, of course,

based on the material presented in class, and use that information to call forth the cognitive, affective and behavioral tendency responses. Overt behavior in response to natural (rather than contrived) stimuli in the community is, of course, the most important test of the program's effectiveness. It is also the most difficult dimension to observe and evaluate. Given time, cost and administrative considerations, such a direct observation of a patrolman's behavior in the community could not be undertaken as a part of this evaluation. It's inclusion in future evaluative research will be discussed in Chapter Five.

The questionnaires used in this evaluation were all specifically developed for this purpose. The affective and behavioral extension items were pretested on the trainees at the demonstration school mentioned earlier. Questionnaire #1 is divided into seven sections of four items each. Each section is treated as a single unidimensional scale. Questionnaires #2 and #3 are each treated in their entirety as single unidimensional scales.

As suggested by Phillips (1966:159-162), no known validation process is completely satisfactory. All validation processes ultimately revert to a form of face validation as used in this research. The reliability of the instruments can be noted by the similarity of the mean scores between School A and School B. In addition, if our assessment of validation is accurate, then reliability must also

follow. The cognitive dimension instrument will be used as both a single scale and a series of seven scales comprised of four items each. The seven scales will correspond to the seven objectives of the school. Since the other two instruments contain only one or two items relative to each objective, it was decided to treat them as unidimensional.

The administrative decision to assign all eligible patrolmen who met the research requirements to either School A or School B prevented the establishment of a control group. Those patrolmen who were not assigned to one of these two schools had either participated in the prior demonstration school or did not meet the guidelines as to length of service. Therefore, although desirable for most researches utilizing a pre-test/post-test design, an independent control group could not be formed. As with most scales an assumption of equal intervals (interval data) is doubtful. (Oppenheim, 1966) Consequently only statistics appropriate for use with nominal and ordinal data will be utilized. Since there is little reason to believe the population to be normally distributed, we are restricted to the use of non-parametric statistics.

CHAPTER IV

A SUMMARY AND CRITIQUE OF THE DATA

The data presented and analyzed in this chapter are the results of a product evaluation of the Police-Community Relations Training School described in Chapter Three. The effects of the training program will be analyzed for the two sections of the training school (Schools A and B) both jointly and separately. The presentation begins with the most general form of analysis and moves to the more specific. These data and their analysis are intended to be illustrative of an attempt to carry out one phase of the Stufflebeam model for evaluative research (product evaluation) and are therefore not to be interpreted definitely. A complete discussion of all four stages of the model as they can be applied to the law enforcement area will be presented in Chapter Five.

The most general statement that could be made in evaluating the Police-Community Relations Training School project (hereafter known as the CRS) would be whether or not it was an overall success or failure. Such a determination would have to be based on the extent to which the CRS accomplished its objectives as specified in Chapter Three.[1]

---

[1]Problems involved in operationalizing these objectives will be discussed further on in this chapter.

One way of making such a determination is to dichotomize the combined results of trainee responses to the three questionnaires measuring the cognitive, affective and behavioral extension dimensions of awareness. Any movement in the desired direction (increased awareness), from whatever base or pre-school level, could be classified as indicative of the school's success relative to that trainee. Any movement in the undesired direction (decreased awareness) could indicate the school's failure with respect to that trainee. All instances where no movement or change between pre-school and post-school scores is indicated will be dropped from the analysis according to the procedures suggested by Siegel for use of the Sign Test and Wilcox on Test. (1956) For each of the seven sub-dimensions of cognitive awareness (as measured by separate four-item scales included in Questionnaire #1) together with the affective and behavioral extension dimensions (as measured by ten and eight item scales in Questionnaires #2 and #3 respectively) an evaluation of success (+) or failure (-) is recorded for each trainee. These data are presented separately for schools A and B of the CRS project in the tables in Appendix A.

When the ratio of positive to negative individual changes (regardless of magnitude) is computed for each of the seven cognitive sub-dimensions, the affective dimension and the behavioral extension dimension, a ratio of greater than one could be interpreted as indi-

cating the overall success of the CRS project on that dimension. A ratio of less than one could be interpreted as failure to accomplish the objective represented by that dimension. The distribution of ratio scores is thereby dichotomized with values greater than one being assigned a plus to indicate success, and values of less than one assigned a minus to indicate failure. We would predict that the frequency of successes would exceed the frequency of failures. The implied null hypothesis is that there is no difference in these frequencies. The one-tailed sign test is used to determine the level of statistical significance. (Siegel, 1956) Values significant at the .05 level could be considered sufficient to conclude that the CRS project was successful. The following table presents this datum.

TABLE 1.--The Frequency of Successes and Failures on Each of the Nine Dimensions and Sub-Dimensions for Schools A and B Combined

Frequency of Successes	Frequency of Failures	Total N	Percentage of Successes
13	5	18	72% $p < .05$

The datum in Table 1 can be interpreted as establishing that the frequency of successes (positive changes) exceeds the frequency of failures (negative changes) and that this difference is statistically significant beyond the .05 level. Therefore, since the frequency of

successes exceeds the frequency of failures with statistical significance, we might conclude that the CRS project was a success. Obviously, although such a conclusion would be justified if this were the only datum available, a more thorough examination of this and other data might modify this initial evaluation of substantive significance. For example, when schools A and B are analyzed separately as to the ratio of success to failures a different picture develops.

TABLE 2. --The Frequency of Successes and Failures on Each of the Nine Dimensions and Sub-Dimensions for Schools A and B Separately

	Frequency of Successes	Frequency of Failures	Percentage of Successes
School A (N=9)	4	5	44%
School B (N=9)	9	--	100%

This table suggests that there was an extensive difference between the overall success of schools A and B. A sign test was not applied to this datum since school A, with a less than fifty percent success rate was not even in the predicted direction while school B was successful across each dimension measured. With this new information indicating a discrepancy in the success of the two components of the CRS project, the overall evaluation of success, although statistically possible, is substantively misleading. One does not know which

objectives of the project were achieved or which of the two schools was successful in achieving each objective.

Table 3 presents the frequency of successes and failures and the percentage of successes, by school, for the first of the cognitive sub-dimensions--the biological and social meanings of race. The one-tailed sign tests, predicting a greater than fifty percent success rate, were used again since the CRS project is predicted to be successful. Success rates of less than fifty percent are obviously not in the hypothesized direction.

TABLE 3. --Percentage of Successes on the Cognitive Dimension of Awareness (the Biological and Social Meanings of Race)

	Frequency of Successes	Frequency of Failures	Percentage of Successes
School A (N=18)	4	14	22%
School B (N=14)	10	4	71%

In this table we note that while school A was successful in bringing about positive change in trainee awareness with respect to the meanings of race in only twenty-two percent of its trainees, school B was successful for seventy-one percent of its trainees. Since neither finding is statistically significant at the .05 level, we must accept the implied null hypothesis and conclude that the CRS project

was unsuccessful in increasing the awareness of the biological and social meanings of race for the trainees in both schools.

Although this information is considerably more specific than what was learned from Tables 1 and 2, this same datum can be analyzed with a more powerful statistic. When each of the instances of demonstrated change in awareness between pre-school and post-school measures are ranked according to the magnitude of change, a higher level of analysis is possible. The Wilcoxon signed-ranks matched-pairs test can be applied to the data when it is assumed to be ordinal data. In reality, the Wilcoxon test assumes the data to be of an ordered metric level. (Siegel, 1956) The researcher can make the judgment of "greater than" between any two difference scores arising from any two pairs as well as that judgment between any pair's two performances. By this method a more powerful test of significance can be applied to the data when each difference score is ranked according to its magnitude regardless of sign. The power of the Wilcoxon test is near 95 percent of its parametric parallel, the t test, when the assumptions of normality are met. In order to follow the procedures established for the Wilcoxon test, the pairs demonstrating no change between pre and post tests must, unfortunately, be dropped from this analysis.

When the Wilcoxon test is applied to each of the sub-dimensions of cognitive awareness, and the resulting T values in the desired

(positive) direction are statistically significant at the .05 level, then these values could be interpreted as indicating the success of the CRS project on that sub-dimension. The following table presents by school, the T values and their direction for each of the cognitive sub-dimensions.

Table 4 shows that in only two instances did either school A or B give evidence of a statistically significant change in the awareness of the trainees with respect to the seven sub-dimensions of cognitive awareness. In both instances this occurred in school A. One such change was in the desired direction (increasing awareness) while one was in the undesired direction (decreasing awareness). Our original evaluation of the CRS project as successful in achieving its objectives is made even more doubtful through this finding.

On the other hand, although not statistically significant, the direction of change in school B for each of the sub-dimensions is in the desired direction. Our selection of the .05 significance level prevents one from concluding that school B was successful along these sub-dimensions, but the information does further confuse our attempt at evaluation.

The second dimension of awareness, the affective component, similarly analyzed by school, is presented in Table 5.

Table 5 indicates that although trainees in both schools evidenced a change in awareness in the desired direction, only school

TABLE 4. --The Extent and Direction of Effectiveness of the Seven Sub-Dimensions of Cognitive Awareness for Schools A and B*

Cognitive Sub-Dimensions	T Value	Direction
Biological and Social Meanings of Race		
School A (N=18)	32.5 $p < .05$	(-)
School B (N=14)	25.0	(+)
Contributions of Black Americans		
School A (N=16)	48.0	(-)
School B (N=15)	59.5	(+)
Relationship of Poverty to Crime		
School A (N=17)	77.0	(-)
School B (N=14)	44.0	(+)
Basis for Prejudice and Discrimination		
School A (N=17)	55.5	(-)
School B (N=13)	22.5	(+)
Meaning and Implication of Black Power		
School A (N=16)	22.0 $p < .05$	(+)
School B (N=14)	37.0	(+)
Meaning and Extent of Racism		
School A (N=18)	48.5	(+)
School B (N=13)	37.0	(+)
Legal and Social Basis for Protection of Individual Rights		
School A (N=17)	77.0	(-)
School B (N=14)	22.0	(+)

* All incomplete response pairs and those indicating no change have been omitted from the analysis.

TABLE 5. --The Extent and Direction of Effectiveness of the Affective Dimension of Awareness for Schools A and B*

	T Value	Direction
School A (N=19)	103.0	(+)
School B (N=16)	18.5 $p < .05$	(+)

*All incomplete response pairs and those indicating no change have been omitted from the analysis.

B gives evidence of a statistically significant change having occurred. At best it could only be concluded that the CRS project was partially successful in increasing affective awareness relating to the specified objectives. The discrepancy in effectiveness between schools A and B persists.

Table 6 presents a similar analysis of the behavioral extension component of awareness.

TABLE 6. --The Extent and Direction of Effectiveness of the Behavioral Extension Dimension of Awareness for Schools A and B*

	T Value	Direction
School A (N=16)	18.0 $p < .05$	(+)
School B (N=16)	12.5 $p < .05$	(+)

*All incomplete response pairs and those indicating no change have been omitted from the analysis.

For the first time in our analysis both schools A and B appear to have been successful. The behavioral extension dimension of awareness is shown in Table 6 to have significantly increased in both schools. This dimension, when coupled with the analysis of the cognitive and affective dimensions of awareness further complicates our overall evaluation. How does one account for the apparently different success rates of the three dimensions of awareness? Why the discrepancy between the effectiveness of schools A and B? How could these problems have been at least minimized?

First, the seven school objectives were never carefully operationalized. As stated in Chapter Three, the term "awareness" was simply divided into three components or dimensions. These dimensions should have been operationally defined prior to their use by the evaluator as criteria against which the success of the CRS project was to be evaluated. The objectives as stated would be difficult to operationalize in behavioral terms.

Second, the measurement of the extent to which a trainee was "aware" of the seven substantive areas (meanings of race, black power, etc.), rested on assumptions as to the validity of the instruments. Although imperfect, tests for validity, using established criteria, might have shown the measures to be inadequate. The unidimensionality of the scales was also assumed to exist and may have proven otherwise had factor analytic techniques been employed

for verification.

Third, the specification of what would constitute a successful project was never made clear. The administrator (police chief) and the evaluator were both unaware of what the operational definition was of success. If all but one trainee demonstrated increased awareness, would the project be considered successful? All but two? All but ten? It would have been possible that everyone increased in awareness, but only at a minimal level. What level would be sufficient to evaluate the project as achieving its objectives? An additional dilemma seems to have been posed by the high frequency of reversals of the desired awareness of trainees. That is to say, the number of trainees who gave evidence of a negative change in awareness following the training school was in several instances greater than those demonstrating positive change or no change. To what extent can negative change as the result of the CRS project be tolerated and still evaluate the overall project as successful?

Obviously the criteria which were to be met and which would constitute success were never specified. Without these criteria the selection of evaluative instruments, the statistics used and the level of statistical significance chosen was without much meaning for the administrator in the case in point described above.

Another question that arises with respect to the data presented above concerns the discrepancy between the effectiveness of schools

A and B. As stated previously in Chapter Three, the trainees for both schools were chosen according to the same criteria. As a result, comparability of rank, length of service and work assignment was to be expected. In fact, the distribution of rank and work assignment was identical in each school. There were nineteen patrolmen and one detective assigned to each school. Of the twenty officers in each school, one was assigned to the youth bureau, one to the detective division, four to the traffic division and the remaining fourteen to the patrol division. Table Seven presents a comparison of additional demographic characteristics of the trainees.

TABLE 7. --Selected Demographic Characteristics of Officers Participating in Training Schools A and B

	Length of Service in Years	College Training in Years	Frequency of Suspensions	Frequency of Awards
School A (N=20)	$\bar{X} = 4.20$	$\bar{X} = .25$	$\bar{X} = .15$	$\bar{X} = .10$
School B (N=20)	$\bar{X} = 4.18$	$\bar{X} = .50$	$\bar{X} = .13$	$\bar{X} = .25$

Table 7 shows the differences on the dimensions measured between trainees in schools A and B to be minimal. In addition, there is no reason to believe that other characteristics such as marital status, family size, residence or income were other than randomly

distributed. The high degree of demographic comparability between the two schools casts further doubt on the validity and reliability of the instruments used in the evaluation given the contradictory evidence of success above.

It can be safely stated that the attempted product evaluation presented above as a case study raised more questions than it answered. Although the analysis of the data from the CRS project is far from complete, the weaknesses in operationalization of terms, instrument development and specification of objectives and success criteria would not seem to justify further analysis. The important question raised by this abortive attempt at evaluative research within law enforcement is what steps and procedures could have been undertaken to effectively evaluate such a program? It is suggested by this writer that had Stufflebeam's CIPP model been followed throughout the evaluative process, the quality of the evaluation would have been much higher. Also, had the CIPP model been followed, the utility of the evaluation for the administrator would have been greater. The CIPP model, it will be recalled, was designed primarily for use in educational settings. Law enforcement presents different problems to a researcher as discussed in Chapter Two. It is therefore suggested that the CIPP model, with special reference to research problems encountered in the area of law enforcement, would be a helpful guide for future evaluations in law enforcement.

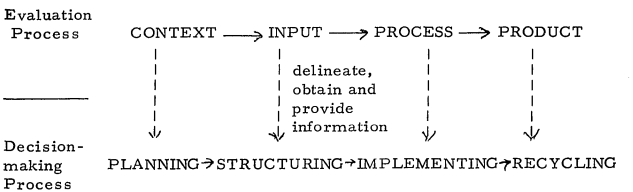
Chapter Five will present each step of the CIPP model as it might be applied in the area of law enforcement with illustrations, when appropriate, from the case in point described in the preceding chapters.

CHAPTER V

THE IMPLEMENTATION OF AN EVALUATIVE RESEARCH MODEL IN LAW ENFORCEMENT

The preceding chapter has presented, with criticism, the data gathered for the purpose of evaluating the Police-Community Relations Training School (CRS) project. Interpretation of these data is made extremely difficult for the reasons specified in Chapter Four. Had a comprehensive evaluation model adapted for the area of law enforcement been utilized, many of the difficulties in evaluating the CRS project could have been avoided or overcome. Stufflebeam's CIPP model seems to be adaptable for such use. For the sake of easy reference, this model as represented in Chapter One will be reproduced here as Figure 2.

FIGURE 2



The CIPP model is to be interpreted as a twelve step model

comprised of four basic stages--the context, input, process and product stages. Within each of these four stages are three sequential processes. The model indicates that the evaluator should first delineate, then obtain and finally provide relevant information for the decision-making process that corresponds with the appropriate stage of the evaluation process. Therefore, when undertaking an evaluative research project, the evaluator is to begin by delineating relevant information in the context stage in order to facilitate the planning stage of the decision-making process. Following this delineating process, information is then obtained and finally provided for the planning stage of the decision-making process. After the planning stage is served, the same three processes are utilized in the input stage of evaluation to serve the structuring process of decision-making. The same series of steps is carried out in the process and product stages of evaluation for their respective decision-making stages. The final step is therefore the providing of information for recycling decisions. This twelfth step may then lead to the need for new planning decisions as the cycle is repeated. For our purposes a single series of the twelve steps will constitute the basic evaluation model.

Before specifying what is involved in each of the twelve steps of the CIPP model, the four stages of the evaluation process will be reviewed with reference to the CRS project. As noted several times in

the previous chapter, the attempt to evaluate the CRS project was essentially a failure. It is not the intent of this section to specify what should have been done to improve the evaluation at each of its weak points. Since the CRS project has been completed this would serve no useful function. The general references made to the project in the following pages are intended to be illustrative of situations likely to be encountered in similar settings.

Context evaluation is designed to identify what is in order to establish a baseline for future stages of evaluation. In law enforcement, it is most important that this stage include an assessment of the total community and its understanding of what law enforcement is and should be doing. In the CRS project, there was little opportunity to determine what the community as a whole and the black community in particular expected of their police department. It would be equally important to know what the program administrators and participants think to be the needs and objectives of their department. If some consensus as to objectives cannot be arrived at among all concerned parties, then agreement as to what constitutes a successful CRS cannot be achieved. The importance of such agreement on goals is precisely why it is mandatory for the evaluator to point out to administrators the context in which their program is taking place. Total evaluation must begin with a consideration of all relevant implications of the program being evaluated.

It is possible that the personnel in a police department see themselves as operating on the "social work" premise of law enforcement. In other words, they may see their role as that of public servant--providing service and help for the whole community. The community, on the other hand, may view the police as agents of arbitrary social control. Were such a disparity to exist, the CRS might be seen by the community as a program to increase the effectiveness of social control. A final product evaluation of the CRS would probably be seen by the community as evidence of the further tightening of unjust social controls. Such an interpretation would be certain to work against any chance for an improvement in police-community relations. This and other kinds of contingencies should be discovered through an adequate context evaluation. School objectives could then be formulated and publicized with the support and cooperation of the entire community.

Another aspect of context evaluation involves an assessment of the congruence between what a system or organization is supposed to accomplish and what it in fact accomplishes. Without the knowledge of what a system is supposed to accomplish, it is impossible to formulate objectives for a program designed to improve the functioning of that system. In addition, desired objectives should be explicit and clearly stated. For example, say that a program's objective is to increase officer "awareness" of certain factors but does not

reveal what the overt behavioral objective of the training program is supposed to be. Without such an explicitly stated objective, an evaluation of the behavioral consequences of any planned change is impossible.

This was one of the problems encountered in the evaluation of the CRS. Even if financial and administrative considerations had permitted a true behavioral evaluation, the lack of specific behavioral objectives made such an evaluation impossible. This situation may reflect a more widespread lack of concrete goals in law enforcement. Confusion within a department may arise over whether to follow the "social work" approach or the "law and order" approach. Context evaluation should be able to detect such ambiguity of purpose in the organization. In so doing it can permit the evaluator to advise the administrators of the problem and its impact on the specific program. This may result in a refocusing of training objectives and procedures.

In short, context evaluation is the process of explicitly specifying where one is and where one is going. The evaluation process therefore begins well before the test program gets under way. It helps to determine what is needed, how to accomplish it and what the "side effects" might be. In law enforcement, context evaluation must include consideration of the community to be served as well as the department providing the service.

The input phase of evaluation considers the resources available to the organization for achieving their objectives as specified in the context phase. The first of these resources to consider is financial. If an organization has its own funds for research and evaluation then the question is simply one of adequacy. It is usually the case, however, that at least partial outside support is needed. Such outside sources of funds and the related requirements may affect the nature and scope of both programs and their evaluation.

For law enforcement agencies the primary funding source during the 1970's will no doubt continue to be the agencies created by the Law Enforcement Assistance Administration and the Omnibus Crime Control and Safe Streets Act of 1968. This federal act and its bureaucratic offspring have their counterparts in each state. These agencies, operating under the block-grant concept, have monies available for planning and action programs, both requiring evaluation.

Private sources should also be considered. National or local foundations may have an interest in certain projects being undertaken and be willing to underwrite their cost. Again, the implications of such funding sources should be clarified at this stage.

Problems of organization should be considered during the input phase of evaluation. The question of internal versus external evaluation mentioned previously is of crucial strategic importance. In

assessing resources, input evaluation should determine if the agency in which the project is undertaken has the personnel to conduct the project themselves and whether such personnel are able to gather and evaluate valid information. In the CRS project it was decided, by administrators, that outside personnel would be better equipped to run the training program. An additional consideration was the nature of the program itself. Since openness and honesty were believed to be important components of teacher-trainee rapport, the absence of departmental supervisors was deemed essential. Such a decision, although requiring much time to be spent in the gaining of officer trust, eventually proved to be a wise one. Although not documented, it seemed as though feelings were exposed and confidences gained that could never have taken place in the presence of command officers or other departmental supervisory personnel. This advantage which accompanied the use of external staff should not be overlooked in its implications for other projects in law enforcement. Input evaluation can provide the vehicle through which determinations such as this can be made.

The selection of an appropriate research design is also the responsibility of the input phase of evaluation. Many factors have to be taken into account when selecting a particular design. Time, cost, competence and availability of personnel, and the form in which the project's objectives are stated are only a few of the considerations.

If, as with the CRS project, overall effectiveness is stressed, then the research should be designed accordingly. Perhaps a project would be structured to identify the relationship between personal characteristics and various outcomes. In this case, appropriate measurement strategies and analytic techniques need to be developed to discriminate among the various types and outcome modes.

Flexibility of design is an important factor also. During the project unanticipated outside events may suggest new objectives or procedures. During the CRS a violent confrontation between the police officers of the city and high school students tock place over a period of several days. The reaction of the trainees to their involvement in this incident was quite varied and often emotional. Had such an occurrence been anticipated, additional procedures for the evaluation of the impact of this event on the officers could have been utilized. Even if unanticipated, a more flexible design together with the sensitivity to interpret and catalog reactions to this event would have added to the overall evaluation program.

In addition to the issue of flexibility, more specific design considerations should be explored in both the context and input phases of evaluation. For example, the construction of questionnaires and their pre-testing would be undertaken in both the context and input phases. If interview techniques are to be employed then schedules must be constructed and pre-tested. If participant observation techniques

are utilized then arrangements for entry into the program as a participant must be made. Steps to insure the validity and reliability of any instruments used should also be taken during context and input evaluation.

In short, whatever procedures are required to serve administrative structuring decisions should be completed by the end of the input phase of evaluation. The structuring of the action program should parallel and be served by the development of evaluative procedures.

Following the input phase of evaluation is process evaluation. The necessity for this phase is largely dependent on the care with which the context and input phases have been carried out. During the course of the program being evaluated a number of problems such as the high school incident could arise. If the previous phases of evaluation have built into the overall design enough freedom to adequately respond to these situations, then process evaluation need only record these changes and their effect. Sometimes important administrative decisions need to be made regarding changes in the implementation of the project. Part of the function of process evaluation would be to provide administrators with the relevant information to make such a decision. A simple example would be the change in scheduling that occurred during the CRS project. It was originally scheduled for Tuesday and Friday evening sessions. After the first meeting officers complained of having to lose their precious

weekend nights off. In this case a change in meeting dates was suggested to the administrators and was agreed upon by them. More serious questions such as the requirement of an unexpected interim report might arise, in which case an on-going process evaluation would be invaluable.

A major instrument in process evaluation is the maintaining of a diary. This record of the day-to-day activities as they occur is of great help in locating areas of program weakness. When possible, an evaluator should be present at most of the program's activities in order to record the nuances of interaction. If his presence would be disruptive to the program, perhaps a recording system (with the knowledge of the participants) could be arranged. Such a procedure may be seen to produce reams of useless data. The organization and editing of such data would be similar to that required in the use of participant observation data. (See, for example, Phillips, 1966) Its utility would be realized, however, when trying to account for failure within the program. The discrepancy in the extent of attitude change demonstrated by schools A and B in the CRS project might have been partially explained had such a record been kept. As it is, this lack of process data leaves this finding unexplained.

Process evaluation could also be helpful in plotting the course and effects of the program being undertaken. The stages of development in attitude change could be examined if periodic evaluations of

attitudes were made. Perhaps most of the desired change takes place after the first few sessions. If this is the case, then an administrative decision to shorten the program would be aided by such information.

Recalling that instrumental criteria are found at intermediate stages of the overall program, what criteria would one expect to use when evaluating law enforcement programs? In law enforcement these criteria might include judgments by superiors as to an officer's performance during the course of his training. This may not constitute a direct test of program objectives, but would be a part of the total picture of its success or failure. Unfortunately, such intermediate evaluations were not made during the CRS project. As with the similar process evaluations, they would have been helpful in explaining variations in schools A and B.

The final stage of the CIPP model involves product evaluation. It was this aspect of evaluation with which the assessment of the CRS project was mainly concerned. Although some aspects of the first three stages of evaluation were sometimes incorporated into the design, this was done unwittingly. The major emphasis of the CRS evaluation was focused on the extent to which the school's objectives were achieved. This emphasis requires, as do all product evaluations, that criteria against which achievement or success is evaluated are carefully and explicitly established. They must then be

utilized in the product stage as the final basis for making recycling decisions. The establishment and use of appropriate criteria was a major weakness in the CRS evaluation.

Consequential criteria are the fundamental baselines established in the context phase against which the basic objectives of the program are evaluated. The evaluation of the CRS project utilized the three dimensions of attitude (or more specifically "awareness") in the form of pre-school and post-school responses. As noted earlier, perhaps the most important consequential criterion was omitted from the CRS project. This is, of course, the overt behavioral criterion of performance on the job. An attempt was made to utilize the frequency and nature of community complaints against officers. Data were gathered in an attempt to demonstrate any change in acquisition of complaints during comparable periods of time preceding and following the training school. Unfortunately, the form in which the data were gathered did not permit any meaningful interpretation. The use of such a criterion would have been helpful in focusing both the school's objectives and its consequent evaluation.

There is a great need for innovative techniques for observing and recording the everyday interactions of the police with the community. Forms of physiological reactions to specified individuals of situations will have to be utilized. Galvanic skin responses and other autonomic and chemical changes in the body should provide important

consequential criteria for evaluation. (See, for example, Summers, 1970:481-554)

The product stage, therefore, will be composed of gathering the data on the measures of objectives and comparing them with the baseline data of the context stage. The variety of techniques involved, and there is a need for innovative techniques in law enforcement evaluation, will vary considerably depending on the criteria of success.

Derived from the four basic stages of the CIPP model is the sequential twelve step process. These steps will be discussed in the order of their suggested implementation. It should be realized, however, that there may be situations in which one or more of the steps has to be omitted. This would no doubt weaken the effectiveness of the evaluation undertaken but would not completely invalidate the model. In order to maximize its utility, the model is intended to be somewhat flexible.

Some of the situations likely to be encountered when implementing each stage of the CIPP model have been discussed earlier. The following pages will attempt to specify more clearly how each of the twelve steps of the model are to be implemented when evaluating a situation in the area of law enforcement. The attention given to this process as it applies to law enforcement specifically is justified by the extent to which law enforcement differs, in degree, from other institutions or quasi-institutions. As discussed in Chapter Two, an

important factor is the powerful norm of secrecy and the attendant suspicion and distrust of outsiders. With this in mind, the following specification of the steps to be taken in implementing the CIPP model within law enforcement is presented.

The twelve steps are specified in a manner that will allow one without sociological or research training to understand them. As is true with many research models, Stufflebeam's discussion of his own model is somewhat complex and confusing. Those without sophistication in research procedures would have difficulty making sense of the model and its suggested implementation. For this reason Stufflebeam's model for evaluative research has been broken down into twelve sequential steps which will hopefully be comprehensible to administrators as well as to those engaged in evaluative research.

Step One--Delineate relevant information in the context stage to facilitate planning decisions.

The sketching out of program objectives is not an easy task in law enforcement. There are likely to be competing philosophies of law enforcement within a single department. The social work/social control issue has already been mentioned. It is one of the functions of Step One to point out to administrators any disagreements among administrative staff and between staff and community as to the assumptions underlying the program's objectives. Perhaps several

days spent in circulating among staff personnel in order to get an outsider's "feel" for the organization would be well spent. In addition, line personnel (the patrolmen on the street) probably have much to offer with respect to the utility of program objectives. It is likely that such information will have been seldom solicited by staff personnel in the past. If access to the patrolmen can be gained, such information should be included in the delineation process.

Step One should also include the clarification of the current status of the literature on the area under study. It is vital to know where one begins in order to determine how far he has traveled. The preprogram status of available information should be recorded in terms that are relevant to the program's objectives.

Step Two--Obtain relevant information in the context stage to facilitate planning decisions.

The information delineated in Step One can be obtained from various sources. The administrators and line personnel could be interviewed. Information questionnaires could be distributed to key personnel. Records and documents could be reviewed. All of these techniques would probably require permission or clearance from the appropriate authorities. As stated earlier, this may be a particularly difficult problem in law enforcement. The honesty and trustworthiness of the evaluator would have to be determined by the administra-

tors given the strong secrecy norm prevalent in police departments. One must be extremely careful not to violate any confidences and not agree to any restrictions if he is doubtful that he can abide by them.

Outside sources will also be necessary to collect information about the department under study. Members of the governmental organizations provide different perspectives on situations involving the law enforcement agency. Private citizens may have important views relating to the issues in the community involving the agency.

The impressions gained by the evaluator himself as the result of informal contacts with personnel should be formalized and as such may be the source of information useful in the planning stage.

Step Three--Providing relevant information in the context stage to facilitate planning decisions.

The step of the model in which the obtained information is given to the administrators seems to be self-explanatory. However, a word of caution is required. The form in which the information is given to the administrators is crucial to the effective use of that information. If the information is so complex in form that the administrators have difficulty interpreting it, the usefulness of the information will be limited. A clear and concise presentation (terse but pithy) is a necessity to facilitate planning decisions by administrators. Information should be in summary form. The use of complex tables and

charts should be avoided when possible. These forms of summary data can be misused and defeat their original purpose. Above all, the information provided should be the information desired. Unwanted or excess information simply confuses and detracts from that which is pertinent. The importance of the information provided is that it will not only serve as the basis for some planning decisions but it will also be the basis for interaction between evaluator and administrator and will thus set the stage for their relationship during the entire program evaluation.

Step Four--Delineate relevant information in the input stage to facilitate structuring decisions.

Information necessary for structuring decisions would include a specification of the resources available to the agency. Perhaps within its own staff the agency may have the people, equipment and talents required for the implementation of its plan of action. Outside sources of talent, etc., should be identified and contacted as to their availability. The fiscal requirements need to be itemized. The source and extent of funds to finance the program also need to be determined.

The general research design (as well as specific research techniques) to be incorporated in the process and product stages of evaluation need specifying. The specific situation under study would,

to a large extent, determine the design to be employed. Cost and time factors would also have to be considered. The major factor determining choice of design would be the administrative needs and objectives as specified in the context stage. In sum, the areas in which information will generally be needed for the input stage are: time available, money available, staff available and research designs and techniques.

Step Five--Obtain relevant information in the input stage to facilitate structuring decisions.

Information as to available resources can be obtained from both the agency and outside specialists. If there is a university or college in the area, faculty members may provide a pool of talent from which to draw, when available, experts in related fields such as courts and corrections should be consulted. Management consultant firms may be a source of expertise in the area of research design. The research designs used in market and production analysis may be helpful in suggesting certain applications of decision-making statistics that would be relevant for program decisions in law enforcement. Decisions based on the maximization of gains or minimization of losses could be aided by such designs.

The evaluator should probably not rely exclusively on his own knowledge of what is available, but should consult others in the field.

Highly specialized designs, developed in other fields, may be appropriate for some programs in law enforcement.

Step Six--Provide relevant information in the input stage to facilitate structuring decisions.

Again, the success of this step, the providing of information, rests on the clarity and appropriateness of that which is provided. The strengths and weaknesses of alternative designs should be made clear to administrators. Technical descriptions of the functions of certain research designs and statistics may only further confuse administrators.

The structuring decisions in the law enforcement area may be restricted by the need to maintain confidentiality of records and investigations. The evaluator can take this into account when providing the administrator with information to be used in structuring the program. Designs and techniques which can be implemented with minimal information can be suggested. Data collection techniques which assure anonymity of respondents can also be suggested.

Structuring decisions are of great importance to the overall decision-making process. They form the basis on which the final evaluation of a program rests. Providing the administrator with all of the relevant information with which to make these decisions is necessary for the final two stages to be successfully completed. The

administrative decisions at this stage identify for the evaluative researcher the possibility and perimeters for his work. Thus, decisions at this point affect not only the program, but also the possibility and ease of evaluating the program.

Step Seven--Delineate relevant information in the process stage to facilitate implementing decisions.

Once the program or plan of action has been undertaken there are a variety of implementing decisions which may need to be made. Research designs or aspects of the program as originally conceived, may, for reasons beyond the control of the evaluator or staff, no longer be feasible to implement. If such a situation arises, contingency plans or modification of the original plan will have to be provided. Evaluations as to the success of each phase of the program will probably be helpful to the administrator. Perhaps all phases appear to the casual observer to be going smoothly. The evaluator should keep in constant contact with key personnel both administering and participating in the program. When any "malfunction" develops the reasons for it should be recorded and examined.

While the program is in process there may be new information available through evaluation that could require changes in the program's objectives. Some internal or external antagonism toward the program and its goals could develop. Pressures may be brought to

bear on the agency from civic groups or government officials. Reactions by the community may bring into question the credibility of the department. Information will be needed which can either refute or substantiate such criticisms and which will allow the program to be flexible enough to continue.

Step Eight--Obtain relevant information in the process stage to facilitate implementing decisions.

To obtain information while the program is in process may involve some participant observation. That is, the evaluator should sit in on at least some aspects of the program to learn first hand what is taking place. Again, the necessity for keeping some information out of the hands of unauthorized personnel may make this difficult. The program may, for example, involve the discussion of suspected police graft. It is possible that an outsider would not be a welcome participant in such a discussion. Such instances would seem to be in the minority and not cause a serious problem for the use of the participant observation technique.

Simply gaining the confidence and respect of participants may present more of a problem in obtaining process information. Care should be taken to maintain personal integrity with respect to the discussion of controversial issues. This should be accomplished by openmindedness and a genuine concern for others' feelings. The

evaluator should be careful, however, that his presence does not significantly affect the proceedings to the point of distorting any product evaluations. Any information or proceedings to which the evaluator or program staff cannot have access should be specified and clearly understood at the outset of the evaluative process. As suggested above, the adaptation of participant observation techniques to this research situation provides a likely way to successfully collect and codify information.

Step Nine--Provide relevant information in the process stage to facilitate implementing decisions.

The administrators will need to have prompt reports of any disruptions in the original plans. The evaluator should be ready to provide these reports. They should be in a form that will facilitate a quick decision. The process stage of evaluation is not the place for extensive presentations that attempt to look at a problem from every angle. The relevancy of information provided is of utmost importance at this stage.

Channels of direct communication should be established between the administrator and the evaluator. In the previous stages communication could be scheduled in advance and follow a convenient timetable. In the process stage this is not always possible and requires that steps be taken to insure the often vital communication of unanti-

cipated developments in the program.

Step Ten--Delineate relevant information in the product stage to facilitate recycling decisions.

In this final stage of evaluation the information to be delineated relates mainly to the accomplishment of the program's objectives. These objectives, as delineated in the context stage, need to be reviewed to see if any changes have taken place during process evaluation. If this is the case, the necessary additions to the information prepared for product evaluation should be made.

All other information required should have been outlined in the previous stages. Since the program will generally have been completed when product evaluation is begun, it is a post factum evaluation. This means that product evaluation can rely on only the data collected throughout the course of the program. The product stage is therefore only as good as the steps which have preceded it. The sketching out of the post-test information to be gathered is the primary task of the delineating process in this final stage.

Step Eleven--Obtain relevant information in the product stage to facilitate recycling decisions.

The information to be obtained in this step constitutes the crux of evaluative research. It is on the basis of this information that

the administrator must make his decisions regarding the continuation, modification or discontinuation of the program. Great care should be taken to see that all safeguards are taken to insure accuracy in obtaining these data. No amount of statistical sophistication can overcome "sloppy" or inaccurate data. The methods of collection used in obtaining the baseline or pre-test data (if that design is used) should be repeated when obtaining post-test data. The same problems of data contamination encountered when obtaining pre-test data should be avoided.

Step Twelve--Provide relevant information in the product stage to facilitate recycling decisions.

This final step of the evaluative process is the point at which the importance of careful attention to the other eleven steps becomes apparent. If these steps have been followed, then all that remains is to synthesize the collected information and present it to the administrator. The information should be in a mutually agreed upon form which will allow for easy application to recycling decisions.

Such decisions may not rest entirely on the results of the evaluation. Cost, political interests, public acceptance and other factors may make unrealistic the continuation of a program evaluated as successful. For this reason, alternative programs that would accomplish the same objectives as the original program but which would

also satisfy outside critics might be suggested to the administrator.

What is most important in this step is that the information provided have utility for the administrator. It must provide the answers to the questions asked by the administrator and seen as necessary for the making of recycling decisions.

The utility of the twelve steps of the evaluative research model outlined above has yet to be demonstrated within the area of law enforcement. Obviously, the steps were not followed in the evaluation of the CRS project. The writer was not aware of the existence of the model until the training school had been completed. It is suggested, however, that had the model been followed, the CRS project could have been much more effectively evaluated. This contention will remain unsupported until a functional test of the model is undertaken.

Even without the aid of Stufflebeam's model for evaluation, the shortcomings of the attempted evaluation of the CRS project might have been overcome. The source of the problems encountered in evaluating the CRS project can probably be found present in most other unsuccessful evaluations. It is the lack of meaningful communication between the evaluator and the administrator. If such communication is not maintained, the interlocking of the information provided by the evaluator with the informational needs of the administrator will not be accomplished. The final responsibility

for such a failure can rest only with the evaluator. He is the one whose responsibility it is to provide information to administrators which is relevant to their decision-making needs.

It is intended that each of the discussions following the twelve steps in evaluative research listed above has as its common theme the importance of meaningful communication between evaluator and administrator. With each working at differing purposes, there is little hope for the implementation of knowledge gained through evaluation. For this reason many evaluations in the past, especially in the area of social action programs, have remained unused (and perhaps unread) rather than facilitating the decision-making process. Evaluators will complain of the insensitivity and incompetence demonstrated by administrators as they ignore or inhibit the evaluation of their programs. Regardless of the truth contained in these accusations in any given situation, the responsibility for overcoming these hurdles remains with the evaluator.

In addition to questions of utility, problems also may arise concerning the extent of an evaluator's involvement in the social action program and his influence on its outcome. When following the CIPP model, the evaluator is often a participant in the program from its inception through to its conclusion. The evaluator may find himself mediating disputes between administrators and participants. If a lack of congruence between administrators and participants as to

program goals is discovered by the evaluator, he should make such discrepancies known and attempt to resolve them. In this way, the evaluator will have an impact on the selection of program objectives. Such an influence is not necessarily desirable or undesirable but should at least be recognized as occurring by the evaluator and all others concerned.

During the process stage, the evaluator is likely to interact with participants. In so doing, the evaluator may be paralleling the role of the participant observer in data collection. As such, the evaluator must be aware of his impact on the program. Any comments he may make, or even his presence in the room will have some effect, however slight, on the outcome of the program. The evaluator should try to minimize his effect on the accomplishment of program objectives by remaining unobtrusive and non-critical during his participation in the on-going program. The evaluator's presence is intended to help him better evaluate the processes as they take place, but not to directly influence the achievement of program goals.

The need for models in evaluative research should be apparent. Most problems encountered in designing and interpreting evaluative research can be traced to the lack of explicit and comprehensible models to be followed. Social scientists should be devoting more of their time to the development of new, more effective models or the development of the steps required to implement existing models.

Such models should, in addition to serving decision-making needs, be of value to the social sciences in theory construction and theory testing. Even though their impact may be indirect, evaluative research models may aid in the inductive acquisition of knowledge. In the usually controlled settings for evaluative research, previously unspecified relationships among variables may be discernable which would be useful to the social sciences in theory construction. In his role as an evaluator, the social scientist is in an advantageous position to contribute to the development of his discipline. His familiarity with fundamental research procedures together with his knowledge in substantive areas should make the social scientist a valuable asset to any evaluative research undertaking, especially those involving social action programs. Thus, the needs of both an academic discipline and the organization in which the evaluation takes place can be served by social scientists enacting the role of evaluator.

APPENDIX A

TABLE 8. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Biological and Social Meanings of Race, School A*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
1	4.0	5.5	+1.5	14.5
2	2.5	2.0	- .5	6.0
3	3.0	3.3	+ .3	2.5
4	2.8	2.0	- .8	9.5
5	5.0	---	---	---
6	5.8	5.3	- .5	6.0
7	4.0	3.0	-1.0	11.5
8	2.5	4.0	+1.5	14.5
9	5.5	3.0	-2.5	17.0
10	3.3	2.8	- .5	6.0
11	6.0	2.8	-3.2	18.0
12	4.8	2.6	-1.8	16.0
13	5.0	3.8	-1.2	13.0
14	5.8	6.0	+ .2	1.0
15	5.5	4.5	-1.0	11.5
16	4.8	4.8	N.C.	
17	6.0	5.5	- .5	6.0
18	3.3	3.0	- .3	2.5
19	4.5	4.0	- .5	6.0
20	4.3	3.5	- .8	9.5

* $\bar{X}$ Score range 1-6

TABLE 9. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Biological and Social Meanings of Race, School B*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
21	3.3	- - -	- - -	
22	3.8	3.3	- .5	5.0
23	4.8	3.5	-1.3	11.0
24	---	---	---	
25	---	3.0	---	
26	3.5	5.0	+1.5	13.0
27	3.5	3.8	+ .3	2.5
28	4.0	3.8	+ .2	1.0
29	5.0	5.5	+ .5	5.0
30	---	6.0	---	
31	3.0	3.5	+ .5	5.0
32	4.8	4.0	- .8	8.0
33	4.8	5.5	+ .7	7.0
34	4.5	6.0	+1.5	13.0
35	3.0	3.3	+ .3	2.5
36	3.8	4.8	+1.0	9.0
37	4.5	4.5	N.C.	
38	---	5.5	---	
39	2.3	3.5	+1.2	10.0
40	4.5	6.0	+1.5	13.0

* $\bar{X}$ Score range 1-6

TABLE 10. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Contributions of Black Americans, School A*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
1	2.0	3.5	+1.5	15
2	3.3	2.3	-1.0	12
3	2.5	3.5	+1.0	12
4	2.0	2.0	N. C.	
5	4.3	---	---	
6	5.3	4.5	- .8	9.5
7	3.0	2.3	- .7	7.5
8	3.0	3.8	+ .8	9.5
9	4.0	2.3	-1.7	16
10	4.0	4.0	N. C.	
11	3.3	2.0	-1.3	14
12	2.0	2.5	+ .5	5
13	3.3	3.5	+ .2	1.5
14	6.0	5.3	- .7	7.5
15	3.5	4.0	-1.0	5
16	5.0	4.0	-1.0	12
17	5.3	5.0	- .3	3
18	4.0	3.5	- .5	5
19	3.5	3.3	- .2	1.5
20	2.8	2.8	N. C.	

* $\bar{X}$ Score range 1-6

TABLE 11. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Contributions of Black Americans, School B*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
21	3.5	---	---	
22	3.8	2.8	-1.0	12.5
23	5.0	3.5	-1.5	14
24	---	---	---	
25	---	4.0	---	
26	4.0	4.3	+ .3	5
27	2.8	2.5	- .3	5
28	3.8	4.0	+ .2	1.5
29	4.3	5.0	+ .7	11
30	---	6.0	---	
31	3.3	2.8	- .5	9
32	3.8	4.0	+ .2	1.5
33	3.0	5.8	+2.8	15
34	4.5	4.0	- .5	9
35	3.0	3.5	+ .5	9
36	4.8	4.5	- .3	5
37	3.8	3.5	- .3	5
38	---	5.0	---	
39	4.0	4.3	+ .3	5
40	4.8	5.8	+1.0	12.5

* $\bar{X}$ Score range 1-6

TABLE 12.--Cognitive Dimension of Individual Trainee Attitudes Regarding the Relationship of Poverty to Crime, School A*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
1	3.0	3.0	N.C.	
2	2.5	3.0	+ .5	6
3	3.3	5.0	+1.7	7
4	4.3	2.8	-1.5	16
5	4.5	---	---	
6	6.0	5.0	-1.0	14
7	2.8	3.8	+1.0	14
8	3.5	4.3	+ .8	11
9	3.5	3.5	N.C.	
10	4.0	4.3	+ .3	3
11	3.0	3.8	+ .8	11
12	4.8	3.8	-1.0	14.
13	4.0	3.8	- .2	1.5
14	4.8	5.5	+ .7	9
15	4.5	4.0	- .5	6
16	3.8	4.3	+ .5	6
17	5.8	5.0	- .8	11
18	4.3	3.8	- .5	6
19	4.0	3.8	- .2	1.5
20	3.3	2.8	- .5	6

* $\bar{X}$ Score range 1-6

TABLE 13. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Relationship of Poverty to Crime, School B*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
21	3.8	---	---	
22	3.8	4.8	+1.0	11
23	6.0	3.8	-2.2	14
24	---	---	---	
25	---	6.0	---	
26	2.8	3.8	+1.0	11
27	3.0	3.8	+ .8	9
28	3.5	3.8	+ .3	3
29	6.0	5.0	-1.0	11
30	---	4.8	---	
31	3.8	3.5	- .3	3
32	4.5	4.0	- .5	6.5
33	3.5	5.3	+1.8	13
34	3.8	3.8	N.C.	
35	3.8	4.3	+ .5	6.5
36	4.8	5.0	+ .2	1
37	3.5	4.0	+ .5	6.5
38	---	5.8	---	
39	4.3	4.0	- .3	3
40	5.0	4.5	- .5	6.5

* $\bar{X}$ Score range 1-6

TABLE 14. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Basis for Prejudice and Discrimination, School A*

Trainees Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
1	1.5	3.5	+2.0	16.5
2	3.5	2.8	- .7	9
3	3.3	3.3	N.C.	
4	3.8	2.8	-1.0	12.5
5	3.5	---	---	
6	5.0	4.0	-1.0	12.5
7	1.8	3.0	+1.2	14
8	3.0	1.5	-1.5	15
9	2.5	2.3	- .2	2
10	3.7	3.5	- .2	2
11	3.0	1.0	-2.0	16.5
12	3.5	3.8	+ .3	4
13	3.0	3.5	+ .5	6
14	4.8	4.8	N.C.	
15	4.3	3.5	- .8	11
16	5.0	4.3	- .7	9
17	5.0	4.8	- .2	2
18	3.3	3.8	+ .5	6
19	3.3	4.0	+ .7	9
20	3.5	3.0	- .5	6

* $\bar{X}$ Score range 1-6

TABLE 15. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Basis for Prejudice and Discrimination, School B*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
21	3.3	---	---	
22	4.5	4.0	- .5	5.5
23	4.0	2.8	-1.2	11.5
24	---	---	---	
25	---	3.5	---	
26	3.0	3.3	= .3	2
27	2.0	3.0	+1.0	9.5
28	2.5	3.5	+1.0	9.5
29	4.0	4.3	+ .3	2
30	---	6.0	---	
31	2.0	2.5	+ .5	5.5
32	3.5	3.8	+ .3	2
33	3.3	5.5	+2.2	13
34	3.8	3.3	- .5	5.5
35	3.3	3.8	+ .5	5.5
36	3.8	4.5	+ .7	8
37	4.0	4.0	N.C.	
38	---	3.5	---	
39	---	3.5	---	
40	3.8	5.0	+1.2	11.5

* $\bar{X}$ Score range 1-6

TABLE 16. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Meaning and Implication of Black Power, School A*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
1	1.5	2.0	+ .5	5.5
2	3.0	2.8	- .2	1.5
3	1.0	4.0	+3.0	16
4	2.5	3.0	+ .5	5.5
5	4.8	---	---	
6	4.5	4.5	N.C.	
7	2.0	3.3	+1.3	12
8	1.8	2.5	+ .7	8
9	2.5	2.8	+ .3	3.5
10	3.7	2.8	+ .3	3.5
11	2.5	1.6	- .9	11
12	2.5	2.3	- .2	1.5
13	3.3	3.3	N.C.	
14	6.0	6.0	N.C.	
15	2.5	3.3	+ .8	10
16	1.7	3.5	+1.8	15
17	5.5	4.8	- .7	8
18	3.3	4.0	+ .7	8
19	2.0	3.0	+1.5	13.5
20	1.3	2.8	+1.5	13.5

* $\bar{X}$ Score range 1-6

TABLE 17. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Meaning and Implication of Black Power, School B*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
21	3.3	---	---	
22	1.8	1.5	- .3	3
23	3.3	2.5	- .8	10.5
24	---	---	---	
25	---	3.5	---	
26	3.0	3.8	+ .8	10.5
27	2.0	1.3	- .7	8
28	3.8	4.0	+ .2	1.5
29	5.0	4.3	- .7	8
30	---	4.8	---	
31	1.8	3.3	+1.5	13
32	3.3	3.8	+ .5	6
33	2.8	4.8	+2.0	14
34	3.8	4.5	+ .7	8
35	2.5	2.3	- .2	1.5
36	4.0	3.5	- .5	6
37	2.8	3.3	+ .5	6
38	---	4.0	---	
39	---	3.3	---	
40	3.0	4.3	+1.3	12

* $\bar{X}$ Score range 1-6

TABLE 18. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Meaning and Extent of Racism, School A*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
1	2.5	2.5	N.C.	
2	2.3	2.5	+ .2	4
3	3.5	3.8	+ .3	8.5
4	3.5	2.8	- .7	14
5	3.8	---	---	
6	4.8	5.0	+ .2	4
7	2.8	3.3	+ .5	11
8	3.8	3.0	- .8	16
9	2.0	2.5	+ .5	11
10	2.0	1.8	- .2	4
11	2.0	1.3	- .7	14
12	2.3	2.8	+ .5	11
13	2.8	3.0	+ .2	4
14	5.5	5.3	- .2	4
15	4.0	3.3	- .7	14
16	1.8	3.8	+2.0	18
17	4.3	4.0	- .3	8.5
18	2.3	3.5	+1.2	17
19	3.3	3.5	+ .2	4
20	3.5	3.3	- .2	4

* $\bar{X}$ Score range 1-6

TABLE 19. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Meaning and Extent of Racism, School B*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
21	3.8	---	---	
22	2.5	3.3	+ .8	8.5
23	5.8	3.3	-2.5	13
24	---	---	---	
25	---	3.5	---	
26	3.3	3.3	N.C.	
27	2.3	2.0	- .3	4
28	3.3	4.3	+1.0	10.5
29	4.3	3.8	- .5	6
30	---	4.8	---	
31	2.0	2.5	+ .5	6
32	3.5	4.5	+1.0	10.5
33	3.5	4.3	+ .8	8.5
34	3.0	4.5	+1.5	12
35	3.0	2.8	- .2	2
36	4.0	3.8	- .2	2
37	2.8	3.0	+ .2	2
38	---	3.8	---	
39	---	3.8	---	
40	4.5	4.0	- .5	6

* $\bar{X}$ Score range 1-6

TABLE 20. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Legal and Social Basis for Protection of Individual Rights, School A*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
1	4.3	5.8	+2.5	17
2	2.8	2.5	- .3	4.5
3	3.0	4.3	+1.3	14
4	2.7	3.8	+1.1	13
5	4.0	---	---	
6	5.0	4.3	- .7	10
7	3.0	2.8	- .2	2
8	3.8	3.5	- .3	4.5
9	3.3	1.8	-1.5	15.5
10	3.0	2.3	- .7	10
11	2.3	3.3	+1.0	12
12	3.3	2.8	- .5	7
13	3.3	3.5	+ .2	2
14	5.8	5.3	- .5	7
15	5.3	6.0	+ .7	10
16	4.5	5.0	+ .5	7
17	6.0	4.5	-1.5	15.5
18	3.0	3.0	N.C.	
19	3.3	3.5	+ .2	2
20	2.8	2.8	N.C.	

* $\bar{X}$ Score range 1-6

TABLE 21. --Cognitive Dimension of Individual Trainee Attitudes Regarding the Legal and Social Basis for Protection of Individual Rights, School B*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
21	2.3	---	---	
22	3.5	4.0	+ .5	6
23	4.5	3.8	- .7	7.5
24	---	---	---	
25	---	5.3	---	
26	3.5	3.8	+ .3	4
27	2.0	3.0	+1.0	10
28	3.8	3.0	- .8	9
29	5.5	5.3	- .2	1.5
30	---	4.8	---	
31	2.8	4.8	+2.0	11.5
32	4.8	5.0	+ .2	1.5
33	2.3	5.5	+3.2	14
34	3.0	5.0	+2.0	11.5
35	3.8	4.5	+ .7	7.5
36	5.5	5.8	+ .3	4
37	2.5	5.0	+2.5	13
38	---	5.5	---	
39	---	4.3	---	
40	5.8	5.5	- .3	4

* $\bar{X}$ Score range 1-6

TABLE 22. --Affective Dimension of Individual Trainee Attitudes as Related to School Objectives, School A*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
1	3.4	2.5	- .9	19
2	2.9	2.7	- .2	4
3	2.4	2.9	+ .5	14
4	2.7	3.1	+ .4	11.5
5	3.5	---	---	
6	3.8	4.1	+ .3	8.5
7	2.6	2.2	- .4	11.5
8	3.0	3.1	+ .1	1
9	3.3	2.5	- .8	17.5
10	2.8	3.0	+ .2	4
11	1.8	1.5	- .3	8.5
12	2.5	3.0	+ .5	14
13	2.7	3.0	+ .3	8.5
14	4.2	3.7	- .5	14
15	3.2	3.4	+ .2	4
16	3.3	3.0	- .3	8.5
17	3.8	3.6	- .2	4
18	2.5	2.7	+ .2	4
19	2.6	3.4	+ .8	17.5
20	3.0	2.3	- .7	16

* $\bar{X}$ Score range 1-5

TABLE 23.--Affective Dimension of Individual Trainee Attitudes As Related to School Objectives, School B*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
21	2.2	3.1	+ .9	13
22	2.5	3.5	+1.0	14.5
23	4.1	3.1	-1.0	14.5
24	---	---	---	
25	---	3.1	---	
26	2.7	3.4	+ .7	11
27	2.5	2.7	+ .2	2
28	3.0	3.5	+ .5	7
29	3.4	3.9	+ .5	7
30	---	3.7	---	
31	2.5	2.8	+ .3	4
32	3.0	3.5	+ .5	7
33	2.8	4.0	+1.2	16
34	3.3	4.0	+ .7	11
35	2.4	3.0	+ .6	9
36	2.8	3.1	+ .3	4
37	3.7	3.4	- .3	4
38	---	3.4	---	
39	3.2	3.3	+ .1	1
40	3.8	4.5	+ .7	11

* $\bar{X}$ Score range 1-5

TABLE 24. --Behavioral Extension of Individual Trainee Attitudes As Related to School Objectives, School A*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
1	3.6	3.5	- .1	3
2	3.0	3.0	N. C.	
3	3.1	4.3	+1.2	14.5
4	3.4	5.0	+1.6	16
5	4.3	---	---	
6	4.5	4.8	+ .3	7
7	2.5	3.0	+ .5	10.5
8	2.6	3.8	+1.2	14.5
9	3.6	3.6	N. C.	
10	3.4	3.5	+ .1	3
11	2.6	3.0	+ .4	8
12	4.0	3.9	- .1	3
13	3.8	3.6	- .2	6
14	4.3	4.3	N. C.	
15	3.6	4.5	+ .9	12.5
16	3.9	4.8	+ .9	12.5
17	4.0	3.9	- .1	3
18	3.4	3.9	+ .5	10.5
19	3.5	3.9	+ .4	9
20	3.9	3.8	- .1	3

* $\bar{X}$ Score range 1-5

TABLE 25. --Behavioral Extension of Individual Trainee Attitudes As Related to School Objectives, School B*

Trainee Number	Pre-Test $\bar{X}$	Post-Test $\bar{X}$	Change or d	Rank of d
21	3.6	4.6	+1.0	16
22	3.4	4.0	+ .6	10.5
23	4.8	4.3	- .5	8
24	---	---	---	
25	---	4.8	---	
26	3.6	3.9	+ .3	6
27	3.3	3.8	+ .5	8
28	3.6	4.3	+ .7	13
29	4.4	4.5	+ .1	2
30	---	5.0	---	
31	3.6	3.4	- .2	4.5
32	3.5	3.6	+ .1	2
33	4.4	4.5	+ .1	2
34	3.5	4.4	+ .9	15
35	3.6	4.3	+ .7	13
36	4.0	4.6	+ .6	10.5
37	3.6	4.1	+ .5	8
38	---	3.9	---	
39	3.4	4.1	+ .7	13
40	4.4	4.6	+ .2	4.5

* $\bar{X}$ Score range 1-5

APPENDIX B

QUESTIONNAIRE NO. 1

Cognitive Dimension of Trainee Attitudes Toward School Objectives

In the space to the left of each statement indicate the extent to which you agree with that statement by filling in the appropriate number from the key below.

1 - Strongly agree
2 - Moderately agree
3 - Mildly agree
4 - Mildly disagree
5 - Moderately disagree
6 - Strongly disagree

____ 1. One can tell how a person will behave by knowing something about his racial group.

____ 2. Black Americans do not hold many responsible positions because they do not have the ability necessary to do a good job.

____ 3. Interracial marriages are bad for America.

____ 4. Because of their race, some people are born with very little ambition.

____ 5. Black Americans are not as bad off as we are led to believe.

____ 6. White people are almost solely responsible for the development of this country.

____ 7. The fact that black Americans were suppressed in the past has very little effect on their performance today.

____ 8. Contributions of blacks are accurately depicted in the study of American history.

____ 9. If the court would give stiffer sentences to those poor people who continually violate the law then they would change their criminal behavior.

____ 10. Most poor people should be subject to more frequent questioning and searching procedures by the police because of their typical involvement in crime.

____ 11. Most of the poor people who turn to crime do so because of their deprived circumstances.

____ 12. Low-income housing projects are usually responsible for high crime rates.

____ 13. Taken as a group, the "native intelligence" of blacks is somewhat lower than that of whites.

____ 14. The ability of Black Americans is not recognized due to their ghetto environment.

____ 15. A person who is prejudiced often has a valid reason for his feelings.

____ 16. In our society there are, proportionately, many more intelligent whites in the white population than there are blacks in the black population.

____ 17. Problems in the black neighborhoods would work themselves out naturally if the black militants could be effectively silenced.

____ 18. Peaceful and non-violent means have gained very little for the Black American.

____ 19. The demand for "black power" by blacks in our society is simply a way of asking for full citizenship.

____ 20. Black power is simply a slogan used to justify black violence.

____ 21. Most of the charges regarding racism in this city are unfounded.

____ 22. Racism is used all too often as an excuse by those who cannot make their own way in life.

____ 23. Racism is such a major force in our society that no one can escape being responsible for it.

____ 24. In all reality, the understanding of racism has little to do with a police officer satisfactorily performing his job.

____ 25. Whenever legislation increases individual rights, it gives

more freedom to the criminal rather than protecting the rights of the innocent.

____ 26. SDS, Black Panthers, and other such groups involved in violent disturbances should be given the same rights by an officer as those given to the "average" law abiding citizen.

____ 27. Those rights guaranteed to people by the constitution must always be upheld, even if a known criminal is allowed to remain free.

____ 28. Because of their actions and attitudes some people in this country don't deserve to have their individual rights protected.

Note: The post-test questionnaire consists of the same items in an altered sequence.

QUESTIONNAIRE NO. 2

Affective Dimension of Trainee Attitudes Toward School Objectives

In the space to the left of each statement indicate the extent to which you agree with that statement by filling in the appropriate number from the key below.

1 - Strongly agree
2 - Agree
3 - Undecided
4 - Disagree
5 - Strongly disagree

____ 1. If more black people understood the job of the police they would not be so hostile toward them.

____ 2. An officer's attitude toward minority groups would be greatly improved if he understood the cultural differences involved.

____ 3. Minority groups, especially blacks, are hostile toward the police because they are influenced by the Communists.

____ 4. Black people in general are not pushing hard and fast enough to bring about social change.

____ 5. Black Action groups are doing more harm than good.

____ 6. A policeman has to be tougher when dealing with black youth since they are more likely to be troublemakers.

____ 7. If a black person really wanted to work there is no problem finding a job in this city.

____ 8. America was built, for the most part, on the ingenuity, courage and skills of white Americans.

____ 9. Black youth feel they can violate the laws of society without being punished.

____ 10. This is a city that treats all ethnic and racial groups alike.

Note: The post-test questionnaire consists of the same items in an altered sequence.

QUESTIONNAIRE NO. 3

Behavioral Extension of Trainee Attitudes Toward School Objectives

Below are several situations involving Police Officer Poc. In the space to the left of each statement indicate the extent to which you agree with Officer Poc's behavior by filling in the appropriate number from the key below.

1 - Strongly agree
2 - Agree
3 - Undecided
4 - Disagree
5 - Strongly disagree

____ 1. Officer Poc., talking to a fellow officer, said, "If the Negro is not kept in his place, he becomes arrogant, overbearing and disagreeable."

____ 2. When asked about neighborhood integration, Officer Poc remarked, "I am convinced that even if Negroes had the same living conditions as white people, most Negroes would still have lower morals than whites."

____ 3. In a heated argument with four fellow officers, Poc said that recent Supreme Court decisions regarding police procedures made the job in the inner city more difficult and that these decisions should be reversed by the current Supreme Court.

____ 4. Last week Officer Poc observed an eight year old girl pick up a candy bar and walk out of the store without paying for it. She was stopped and reprimanded by him. Two days later he saw an eighteen year old boy stuff a bag of potato chips in his pocket; he was on his way out of the store when Poc arrested him.

____ 5. Martinez Sanchez, a migrant from a remote rural area in Texas, was observed by Officer Poc as he crossed the street on a "NO WALK" sign. Poc stopped Sanchez and gave him a citation in spite of Sanchez's efforts to explain, in broken English, his ignorance of the "NO WALK" sign.

____ 6. Officer Poc was instructed to investigate a fight involving a black family on the north end of town. Before going to the scene, Officer Poc decided to stop at a drive-in restaurant for a cup of coffee.

____ 7. Officer Poc stopped a black male driver for exceeding the speed limit. Poc recognized the youth because he had stopped him several times during the last six months. Without using any profanity, Poc gave the youth a good tongue lashing and a citation.

____ 8. When Officer Poc patrolled the "wealthy section" in town and came upon youngsters playing in the street he would slow down and give a cheerful wave as he passed, but in the "poor section" he would not slow down for youngsters playing in the street. Instead of a cheerful wave, he had a stern no-nonsense look on his face.

Note: The post-test questionnaire consists of the same items in an altered sequence.

SELECTED BIBLIOGRAPHY

Banton, Michael
1964 The Policeman in the Community. New York: Basic Books.

Bayley, David and Arnold Mendelsohn
1969 Minorities and the Police: Confrontation in America. New York: The Free Press.

Becker, Harold K.
1968 Law Enforcement: A Selected Bibliography. Metuchen: Scarecrow.

Berkley, George E.
1969 The Democratic Policeman. Boston: Beacon Press.

Black, Algernon D.
1968 The People and the Police. New York: McGraw-Hill.

Bouma, Donald H.
1969 Kids and Cops. Grand Rapids: Eerdmans.

Bouma, Donald H., et al.
1970 An Evaluation of a Police-School Liaison Program as a Factor in Changing Student Attitudes Toward Police and Law Enforcement. Washington: National Institute of Law Enforcement and Criminal Justice. (Grant Award: NI-068).

Brandstatter, A. F. and Louis A. Radelet (eds.)
1968 Police and Community Relations: A Sourcebook. Beverly Hills: Glencoe.

Brooks, M.
1965 "The Community Action Program as a Setting for Applied Research." Journal of Social Issues 21:29-40.

Campbell, Donald T. and Julian C. Stanley
1966 Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally.

Caro, Francis G.
1971 "Issues in the Evaluation of Social Programs." Review of Educational Research 41 (2) : 87-114.

Cherns, A.
1969 "Social Research and Its Diffusion." Human Relations 29: 209-218.

Clark, Ramsey
1970 Crime in America. New York: Simon and Schuster.

Cohen, David K.
1970 "Politics and Research: Evaluation of Social Action Programs in Education." Review of Educational Research 40 (2) : 213-238.

DeFleur, Melvin L., et al.
1971 Sociology: Man in Society. Glenview: Scott Foresman.

Dressel, Paul L., Director
1961 Evaluation in Higher Education. Boston: Houghton Mifflin.

Earle, Howard H.
1967 Police-Community Relations. Springfield: Thomas.

Ebel, Robert L.
1965 Measuring Educational Achievement. Englewood Cliffs: Prentice-Hall.

Furst, Edward J.
1958 Constructing Evaluation Instruments. New York: Longman, Green.

Gage, N. L. (ed.)
1963 Handbook of Research on Teaching. Chicago: Rand McNally.

Germann, A. C.
1962 "Police Planning and Research as it Affects Police-Community Relations." Police 6 (3) : 36-39.

Glass, G. V.
1971 "The Growth of Evaluation Methodology." AERA Monograph Series on Curriculum Evaluation No. 7.

Green, John A.
1970 Introduction to Measurement and Evaluation. New York: Dodd Mead.

Greenberg, B. G.
1968 "Evaluation of Social Programs." Review of the International Statistical Institute 36: 260-277.

Momboisse, Raymond M.
1967 Community Relations and Riot Prevention. Springfield: Thomas.

Niederhoffer, Arthur
1967 Behind the Shield. Garden City: Doubleday.

Niederhoffer, Arthur and Abraham S. Blumberg (ed.)
1970 The Ambivalent Force: Perspectives on the Police. Waltham: Ginn.

Oppenheim, A. N.
1966 Questionnaire Design and Attitude Measurement. New York: Basic Books.

Owens, Robert C. and Albert D. Hamann
1966 "Evaluating the Training Program." Police Chief 33 (February) : 22-24.

Phillips, Bernard S.
1966 Social Research Strategy and Tactics. New York: MacMillan

President's Commission on Law Enforcement and Administration of Justice
1967a A National Survey of Police and Community Relations. Field Surveys V. East Lansing: Michigan State University.

1967b Task Force Report: The Police. Washington: United States Government Printing Office.

Puzo, Mario
1969 The Godfather. New York: Putnam.

Quade, Edward S. (ed.)
1967 Analysis for Military Decisions. Chicago: Rand McNally.

Roberts, Albert R.
1970 "Federal Funding for Police Staff Development and Training Programs." Police 15 (2) : 14-19.

Rossi, P.
1967 "Evaluating Social Action Programs." Transaction 4 : 51-53.

Schade, Thomas V.
1970 "Police Perceptions of Inner City Youths Attitude Toward Law Enforcement." Unpublished Master's Thesis. Kalamazoo: Western Michigan University.

Scriven, Michael S.
1967 "The Methodology of Evaluation." AERA Monograph Series on Curriculum Evaluation. Book 1. Chicago: Rand McNally.

1969 "Evaluating Educational Programs: A Symposium." The Urban Review 3 (4) : 20-22.

Siegel, S.
1956 Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill.

Skolnick, Jerome H.
1966 Justice Without Trial: Law Enforcement in a Democratic Society. New York: Wiley.

Stufflebeam, Daniel L., Chairman, Phi Delta Kappa Commission on Evaluation
1970 Educational Evaluation and Decision-Making. Columbus: Ohio State University Evaluation Center.

Suchman, Edward A.
1966 "A Model for Research and Evaluation on Rehabilitation." In M. Sussman (ed.), Sociology and Rehabilitation. Washington: American Sociological Association.

1967 Evaluative Research. New York: Russell Sage.

Summers, Gene F. (ed.)
1970 Attitude Measurement. Chicago: Rand McNally.

Thorndike, Robert L. and Elizabeth Hagen
1961 Measurement and Evaluation in Psychology and Education. New York: Wiley.

Tyler, Ralph W.
1950 Basic Principles of Curriculum and Instruction, Syllabus for Education. 360. Chicago: University of Chicago Press.

Weirman, Charles
1970 "A Critical Analysis of a Police-School Liaison Program to Implement Attitudinal Changes in Junior High Students." Unpublished Master's Thesis. East Lansing: Michigan State University.

Weiss, R. and M. Rein
1969 "The Evaluation of Broad-Aim Programs: A Cautionary Case and a Moral." Annals of the American Academy of Political and Social Science 385: 133-142.

Westley, William
1951 "The Police: A Sociological Study of Law, Custom and Morality." Unpublished Doctor's Dissertation. Chicago: University of Chicago.

Whittemore, L. H.
1969 Cop!. Greenwich: Fawcett Crest.

Wholey, Joseph S. et al.
1970 Federal Evaluation Policy. Washington: Urban Institute.

Wrightstone, J. W.
1969 "Evaluating Educational Programs: A Symposium." The Urban Review 3 (4) : 5-6.