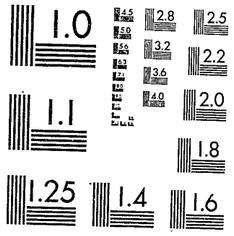

CR Jan
5-26-82

The Evaluation of 'Operational' Social Service Programs: Major Issues and Implications for Juvenile Delinquency Prevention Programs

Province of British Columbia
Ministry of Attorney General

POLICY PLANNING DIVISION
RESEARCH AND EVALUATION UNIT

REPORTS IN

THE JUVENILE DELINQUENCY PREVENTION SERIES

Research and Evaluation Unit

1. A Progress Evaluation of a Juvenile Delinquency Prevention Program : Youth and Family Services. (Wendy Rowe) February 1981

2. Exploratory Survey : School-based Delinquency Prevention Programs in British Columbia. (Maureen Donovan) April 1981

3. The Evaluation of 'Operational' Social Service Programs : Major Issues and Implications for Juvenile Delinquency Prevention Programs

Juvenile Delinquency Prevention : A Review of the Literature Forthcoming

A Systematic Evaluation of a Police-based Youth and Family Counselling Service : Does it Prevent Juvenile Delinquency? (Wendy Rowe) Forthcoming

✓THE EVALUATION OF 'OPERATIONAL' SOCIAL SERVICE PROGRAMS: MAJOR ISSUES AND IMPLICATIONS FOR JUVENILE DELINQUENCY PREVENTION PROGRAMS

by

Wendy E. Rowe

May 1981

Research and Evaluation Unit
Policy Planning Division
Ministry of Attorney General
Vancouver, B.C.

Canadian Cataloguing in Publication Data

Rowe, Wendy E.
The evaluation of 'operational' social service programs

(Juvenile delinquency prevention series ; 3)

Bibliography: p.
ISBN 0-7719-8767-6

1. Juvenile delinquency - Prevention. 2. Rehabilitation of juvenile delinquents. 3. Evaluation research (Social action programs) I. British Columbia. Ministry of Attorney General. Policy Planning Division. Research and Evaluation Unit. II. Title. III. Series.

HV9069.R68 364.3'6 C81-092396-3

TABLE OF CONTENTS

LIST OF TABLES

PREFACE

The Research and Evaluation Unit, Ministry of Attorney General, began a concerted effort during 1980 to address a number of research issues related to juvenile delinquency and its prevention. The research projects undertaken are intended to complement this Ministry's commitment to the prevention of delinquency throughout the Province of British Columbia.

Key areas of concentration identified by a Ministry Committee, which met between 1979 and 1980, included school programs, community development programs, and youth and family counselling programs. Accordingly, the Unit initiated studies to acquire background information for a state-of-the-art position on delinquency prevention. This included a literature review on delinquency, a survey of existing school-based delinquency programs, and evaluations of youth and family counselling programs in the Province. Most of these studies are completed or near completion.

The following report focuses on evaluation design issues which arose during investigation of the efficiency and effectiveness of the delinquency prevention counselling programs. Thus, this report acts as both an introduction to the issues considered in the actual evaluation of these youth and family counselling programs and as a condensed technical guide to the complexities inherent in the design of evaluations for social service programs in general. As such, this report can be read solely as an introduction to program evaluation for

the manager, the funder, and the evaluator, or it can be read as a precursor to the specific evaluation reports on delinquency prevention counselling programs.

Sandra Edelman

Research Officer

A. INTRODUCTION

Increases in juvenile crime rates and the failure of the juvenile justice system to deal with these offenders has prompted the development of numerous juvenile delinquency intervention and treatment programs. These programs can be described in terms of types of intervention and in terms of how early they intervene to prevent first-time or recurrent delinquent behaviour. Brantingham & Faust (1976) identify three levels or entry points of intervention as 'primary', 'secondary' or 'tertiary' prevention. Primary crime prevention is concerned with identifying the physical and social environmental conditions that provide opportunities for or precipitate delinquent acts, and with changing these conditions so that delinquent activity cannot occur. Secondary prevention aims at early identification of potential offenders and seeks to intervene in their lives in such a way that they are prevented or inhibited from future involvement in delinquent activities (e.g. educational, recreational and therapeutic prevention programs for pre-delinquents). Tertiary prevention deals with hard-core criminal offenders and involves a type of intervention in their lives such that they will cease to commit further offences. Fines and probation are examples of 'tertiary' prevention.

Intervention can also be 'punitive', 'mechanical' or 'corrective' (Lejins, 1967). In punitive prevention, a threat of punishment is used as a method to inhibit further delinquent activity (e.g. fines, probation). Mechanical prevention refers to the placing of obstacles

in the way of the potential offender to make it difficult or impossible for him or her to commit an offence. Corrective prevention is based on the assumption that there are certain underlying causes precipitating the youth's involvement in delinquent activities. Intervention is aimed at the elimination or resolution of these problems in order to prevent or reduce the youth's involvement in delinquent activity. Thus, corrective prevention is directed at specific individuals who display behavioural symptoms indicative of a potential for delinquent behaviour.

Punitive and mechanical intervention techniques aimed at hard-core delinquent offenders have generally been ineffective (California Youth Authority, 1976; Romig, 1978). Corrective intervention or treatment programs appear to have more encouraging results as some of these programs are successfully treating or inhibiting the delinquent or pre-delinquent youth from further involvement in juvenile crime. Analysis of the success and failings of existing delinquency prevention programs has resulted in some insight on what types of programs are most successful. It is the conclusion of Romig (1978) and Wright & Dixon (1978) that recreational programs, behaviour modification programs, social case work, individual psychotherapy, group counselling, detached worker/gang projects and therapeutic camping do not reduce the likelihood of recidivism in treated youths as compared with youths who have received no treatment. Intervention strategies that appear to have some promise are: (1) educational programs geared to the individual's specific educational deficiencies, (2) vocational training

programs that provide job advancement, skill training and planning (with follow-up help provided after job placement), and (3) family counselling programs which specifically focus on improving family communications and interaction patterns.

Youth service bureaus are a new approach for dealing with the delinquent and pre-delinquent offender, providing 'service brokerage', 'resource development', 'system modification', and direct community services to youth and family. While the initial evaluations of these programs have not been too favorable, the evaluations have not necessarily been properly conducted, and they have not determined why the youth service bureaus are ineffective.

Many evaluators and reviewers (Selke, 1977; Lewis & Davidson, 1977; Fumilietti, 1980) point out a program (e.g. a youth service bureau) may be unsuccessful because it has not been implemented in the community in the manner that it was so designed. It may be that the 'evaluation study' of the program has judged the effectiveness of the program on the basis of inappropriate criteria, and that, in fact, the program is successfully meeting some of the needs of the community, the families, and the juvenile offenders. Also, it may be that the conceptual or theoretical foundation of the program as a prevention or intervention strategy is invalid, and therefore the services provided within this model are not theoretically viable methods to meet the program objectives.

Evaluation research provides means to determine what types of intervention or treatment strategies are most effective to prevent or

reduce a youth's involvement in delinquent activities and to determine the best and most practical point in time to intervene (e.g. before the youth has committed delinquent acts, after a first offence or after multiple offences). Evaluation research provides a method to determine why a program is not successfully operating as planned or why it is not reaching its objectives. This information enables program directors and administrators to make necessary modification to ensure more efficient and effective intervention or treatment.

In the past, evaluation has been conceptualized typically as an experimental or quasi-experimental test of the effectiveness of a treatment or intervention program. When experimental evaluation methods are designed as an integral part of the testing and implementation of a demonstration intervention strategy, then it is possible to carry out a controlled, successful study. Unfortunately, attempts are made frequently to apply experimental evaluation methods to the post-hoc evaluation of multi-faceted programs that have been operating for a number of years. When this happens, the evaluation studies usually fail, either because of political and social resistance to the implementation of experimental controls, or because data is collected that is unrelated or invalid as an indicator of whether the program is operating effectively. Furthermore, information identifying factors hindering effectiveness generally is not available. However, the fault of these unsuccessful evaluation studies is not with the 'experimental evaluation methodology', but rather that experimental methodology is inappropriate for the evaluation of operational social

service programs. It is the opinion of this author that the experimental and quasi-experimental evaluation methods should be reserved for the testing of experimental and demonstration treatment and intervention programs. Different, but equally valid, process evaluation methods should be employed to monitor and test the successfulness of operational, ill-defined social service programs, such as the Langley Youth and Family Services delinquency prevention program.

This distinction between experimental and process evaluation strategies and between demonstration and operational programs has tremendous implications for many program managers and evaluation researchers involved in the evaluation of operational juvenile delinquency prevention programs. It is important that evaluation methods be employed that will be appropriate for the type of questions being asked.

Since, in the past, experimental and quasi-experimental methods have been employed frequently in the evaluation of operational social service programs, this report will focus initially on the limitations of this action. Other issues surrounding the evaluation of operational programs will be discussed also. An attempt will be made to determine what the role of 'evaluation' should be in the operation of social service programs; what 'evaluation' means to administrators, to funders, to the program staff, and to the clients; and how social service programs such as delinquency prevention programs can be evaluated so that information is obtained which is useful and relevant

to the goal of improving effective delivery of services to youths with problems. All of these issues will be considered with reference to the juvenile justice system and existing models of juvenile delinquency prevention programs.

The first part of this report consists of a brief description of several evaluation research models. The second part of the report will deal with the problems that occur when a quasi-experimental evaluation model is applied to 'operational' programs. The third section will consider the question: How (if at all) should 'operational' programs be evaluated? Some practical guidelines on how to initiate the evaluation of operational programs will be presented. Specifically, practical guidelines will be provided in two areas: (1) how to conduct an assessment of the evaluability of a program (a type of feasibility study), and (2) what type of information system is needed for a program which will facilitate efficient program monitoring and evaluation. The final section of the report will summarize the main issues and conclusions of the report and discuss the implications of some of these general issues for the evaluation of juvenile delinquency programs.

This report acts as an introductory preview of evaluation issues that were considered during the evaluation of "Langley Youth and Family Services", a juvenile delinquency prevention program that provides family counselling to potentially delinquent and first-time delinquent offenders and their families. Since this program has been operating for over three years, the problems of conducting a meaningful and useful evaluation of its services were multifold. The consideration of these issues, the design and results of the evaluation study, and the implications of the findings for research and evaluation of delinquency programs are available in a B.C. Ministry of Attorney General report (Rowe, W., 1981, A Process Evaluation of a Juvenile Delinquency Prevention Program: Youth and Family Services).

B. EVALUATION RESEARCH MODELS

Judgements about the value or success of most juvenile delinquency prevention programs have generally been based on a minimum of factual information. They are heavily influenced by subjective attitudes and expectations, by isolated events, and by clients' and staff's feelings of overall satisfaction. Evaluation research, therefore, has been promoted as a technique for obtaining reliable and valid objective information on the actual outcome effects of a program so that value-judgements about success can be made more rationally.

Ideally, evaluation research should measure the effects of a given program but also provide answers to other questions. Normandeau & Hoasenpusch (1980) give a good summary of the types of information that a traditionally comprehensive evaluation should include:

"1) a detailed description of the type, quality and quantity of the activities (program effort);

2) an accounting of the program's expenditures and the sources of funding;

3) a subjective assessment of the program's desired and undesired consequences as seen by the staff, the clients and other parties directly involved;

4) a scientific study on the question of whether the program's goals have in fact been achieved and whether it may be assumed that they have been achieved as a result of the program (program effect);

5) an analysis of the program effects in relation to their cost and in relation to the cost and effect of other, alternate programs; and,

6) an indication of the program's acceptance by the public, of the opinions of groups supporting and opposing the program, and similar politically relevant information, e.g., the public's perception of the program's outcome (p. 31)."

There are three general, but interrelated, types of evaluation strategies; process evaluations, cost-effectiveness efficiency studies, and experimental (or quasi-experimental) impact evaluations. An experimental impact evaluation study primarily focuses on the question, "Does the program achieve the desired results?" Process evaluation is concerned primarily with answering the question, "What happened?" Efficiency evaluation studies are primarily concerned with the question, "Is this the cheapest strategy to achieve the results we want?" A process evaluation is thought of as a study of the means whereby a program produces its results (Suchman, 1967). One attempts to describe and critically analyze the mechanisms by which 'effect is translated into outcome'. 'Effect' as outlined by Suchman refers to the quantity and quality of all environmental, financial, administrative, and psychological inputs that have gone into the operation of the program. This includes the number and types of clients, the

attitudes and efforts of staff, the amount of money and other resources required, and the expectations and contributions of other community services. This type of evaluation strategy will be discussed in greater detail later in this report.

Cost-effectiveness evaluation studies attempt to assess program efficiency. Suchman describes 'efficiency' as the degree to which one can minimize the amount of effort necessary to operate the program and to maximize the benefits of the program relative to how much it costs. "Efficiency is concerned with the evaluation of alternative paths or methods in terms of costs - in money, time, personnel and public convenience. In a sense it represents a ratio between effort and performance (outputs) - output divided by input." (Suchman, 1967, p.64). Since this type of evaluation is not appropriate unless both a process and experimental impact evaluation have been conducted, it will not be discussed further in this report.

Experimental 'impact' evaluations primarily test the attainment of specific predicted program outcomes or of specific client objectives. Many impact evaluations, however, also include some attempt to describe the program inputs and to describe the mechanisms by which the program objectives and goals are to be reached.

The best method for conducting an effectiveness impact evaluation employs the true experimental design. In this design, half the subjects or clients in need of the service are randomly assigned to an experimental group or a control group. Both groups of clients are assessed in terms of the variables that are expected to change as a

result of the program activities. The experimental group receives the services; that is, the treatment approach. The control group(s) receive no treatment or possibly an alternative treatment. Following treatment, both treatment and non-treatment groups are reassessed to measure any changes. If both groups are very similar at the pre-treatment assessment, but at the post-treatment assessment only the treatment group has improved, then client changes can be attributed to the treatment. Follow-up assessments conducted on both treatment and control groups at a later time (even up to 2 years) tests whether any observed effects or changes of program termination remain stable over time. Campbell & Stanley (1966) provide a detailed account of the benefits and increased validity of using this type of experimental approach.

The true experimental design is nearly impossible to implement in the areas of community research. The norm is the quasi-experimental design. One type of quasi-experimental design, which is a slightly less reliable but acceptable experimental evaluation approach, involves random assignment of clients into a treatment and control group with only one assessment or measurement taken at the end of the treatment period. For example, a group of juveniles that has received treatment can be observed or measured in terms of the number of repeat delinquent offences. This can be compared with another random sample of juvenile delinquents who have not received treatment. The problem with this approach is that one has to infer both groups had similarly high scores on delinquency measures prior to

treatment, and that the now observed post-treatment score for the experimental group is lower than the pre-treatment condition. The non-treatment or control group is assumed to have remained the same (i.e. high) on measures of delinquency.

Other quasi-experimental designs involve the use of a non-random treatment and non-treatment group of clients; that is, a comparison group from another population source. Although, typically, the groups have been matched on some variables, they nevertheless represent two different subject populations and are not similar subjects sampled from the same population. As in the experimental designs, pre and post measurements can be taken on both experimental and comparison groups, or just a post-measure (or several post-measures) can be taken on both the experimental and control groups. Testing of a non-treatment comparison group allows one to ascertain that client changes in the treatment group occurred as a result of the program and were not random or maturational changes.

The least reliable and valid approach for measuring program effects is a one-shot case study. In this approach, a single group of subjects is measured prior to treatment and then after treatment to assess the amount of changes that have occurred during treatment. This type of research design may be subject to considerable error. Pre-post changes could be due to maturational changes, other intervening variables, historical characteristics, regression to the mean, etc. (Campbell & Stanley 1966). This experimentally-weak approach becomes more unreliable when an experimental group of clients is

measured or observed only at the termination of treatment and the pre-treatment behaviour is assumed, or inferred.

Ideally, evaluation research is perceived to accomplish several purposes. By determining how effectively and efficiently a program is operating, program funders and administrators can decide whether to continue the program or whether to redirect the funds into another type of program that is more effective or more efficient. Ongoing monitoring and feedback of information on the effectiveness and efficiency of a program at specific stages can be useful to program staff as a procedure for continually modifying and improving the program. Evaluation studies also assist academics and administrators in increasing their state-of-the-art knowledge of juvenile crime prevention programs so that they can discard the ineffective approaches to juvenile crime prevention and concentrate on developing and investigating the approaches that have the most potential.

C. LIMITATIONS OF EXPERIMENTAL METHODOLOGY FOR EVALUATION OF OPERATIONAL PROGRAMS

1. Introduction

In reality few experimental (or quasi-experimental) impact evaluation studies are completed properly. Logan (1972) found that most evaluation studies suffered from one, or more of the following problems:

1) no clear set of program procedures,

2) no measure was taken of the behaviour that was expected to change,

3) there was no objective criteria of what amount of change indicates 'success',

4) there was no non-treatment group used to control for the possibility that change is not due to the treatment, or

5) there were no follow-ups conducted.

Even when a 'proper' quasi-experimental comprehensive study is carried out (e.g., Cambridge-Somerville Study, McCord & McCord 1959), often it is the case that the criteria by which 'effectiveness' is measured is found to be improper.

The evaluation results may fail to serve the purposes of the administrator, the funder, or the program staff and so are simply discarded. Also, the results may be disappointingly negative and, because of the limitations of the scope of the evaluation, they fail to provide any clues as to why the program is ineffective (Hackler, 1978, 1979). Successful, useful evaluation studies are extremely difficult to conduct.

2. Socio-political Difficulties

One of the primary sources of problems for evaluation studies of operational programs concerns the different vested interests and concerns of administrators, funders, program directors and evaluation researchers. In order to decide whether to continue funding a program, funders may expect the evaluator to carry out a quick, inexpensive assessment of the program's effectiveness relative to its costs. In contrast, the program director or program administrator may wish to utilize evaluation techniques as a management tool to provide comprehensive, ongoing information about the process and accomplishments of

the program's operation, and to provide immediate feedback for continuously improving the program. However, the program director may be concerned that the evaluation procedure does not interfere with the program's normal operation. Staff may be completely opposed to evaluation, considering it to be a potential threat to their employment or professional self-esteem. Clients may look on an evaluation study as personally intrusive and possibly as endangering their access to resources.

Finally, the researcher may add to the problems perceived by the program director by attempting to subject a program to the constraints of a quasi-experimental design. Under such constraints there may be a real or perceived disruption of program activies, unwanted intrusions into the personal lives of clients, excessive criticism of staff, difficulties in carrying out program activities, and reporting of final results, which may prove embarrassing for the program, without any positive consideration or recommendations for improvement.

However, evaluators are often completely dependent on the program staff to conduct the pre and post program assessments of clients' behaviour and to provide other information on the program's operations. Conflicts arising from the different vested interests of the evaluator and the staff may result in incomplete, unreliable or biased data collection. Failure on the part of the researcher to understand the decisions that need to be made by the administrators and funders, and the difficulties involved in completing the study on time, making the report accessible and intelligible to all individuals, and in

keeping the cost of the evaluation as low as possible, may result in the study not being completed at all, or, at best, may result in the report being shelved without consideration of the recommendations.

3. Technical or Methodological Difficulties

There are a number of technical difficulties involved in carrying out an experimental or quasi-experimental evaluation of an operational juvenile delinquency prevention program. First, program people question whether it is ethical to randomly assign some youths to a non-treatment control group. Clients are referred to the program because they have clearly observable problems that need to be treated. To deny these people a service that would be available to them, if not for the demands of an experimental evaluation design, is difficult to enforce. Selecting a matched comparison group of youths from another community where the service is not yet available may resolve the "ethicality" problem in the minds of program people, but this leads to other problems. It is difficult (if not impossible) to match a group of youths from another community to the youths in the treatment groups. Aside from observable differences in characteristics such as age, sex, and severity of problems, there are also differences between the communities in such things as level of community cohesiveness, type and number of other social services available, quality of schools, size of families, socio-economic status, etc. These variables may be significant factors influencing the 'effectiveness' of a program. A second major problem in selecting a matched group of youths from another community has to do with persuading families and

youths to participate in a study for which they receive no direct benefit. Families may express considerable reluctance to be scrutinized as 'guinea pigs' for research concerning a program in which they have no involvement. These families, also, may resent being identified as having family problems or potential delinquency problems with their child.

Another technical problem concerns the difficulty of making reliable and valid measurements of a client's behaviour or characteristics prior to receiving treatment and after treatment is terminated. Sensitive and accurate measurement of behavioural patterns that are likely to be affected by the program may require extensive standardized psychological testing, standardized counsellor observation and ratings, and require self-report data from the clients themselves. Generally, program staff are quite reluctant to subject their clients to this barrage and are also reluctant to spend time, or simply do not have the time available, to conduct these assessments themselves. Subsequently, most social service programs operate informally. Generally, assessment and diagnosis of the client's progress is conducted subjectively by the counsellor or staff member. Also, specific program strategies are designed subjectively and somewhat arbitrarily. Given these typical conditions in most social service programs, post-treatment assessments and follow-up assessments are likely to be viewed by program staff to be time consuming and intrusive.

In addition, it is unlikely that program staff will allow researchers direct access to clients. It is argued frequently that the anonymity of a client must be guaranteed and that all information about a client is confidential. Often this is just an excuse manufactured to deny researchers access to program files. Programs which have been operating for some time are particularly sensitive to possible intrusiveness from the experimental researchers. Very few of these programs have systematically collected objective data on their clients prior to receiving treatment or after treatment. Also, typically, these programs are unclear about the specific measurable objectives of their program. Often a program may simply list that the objective of their program is to provide counselling. In fact, there may be no expectations that the services should effect some permanent change in the behaviour of the client - a behaviour that can be measured objectively.

Self-report assessments of a client's degree of satisfaction with the program when they have terminated treatment are often popularly endorsed by program staff, since most clients will report that they are satisfied with the services they have received. However, this level of satisfaction may not be highly correlated with observable behaviour changes. This type of self-report assessment should be complemented with a follow-up assessment of the client by program staff. Very few programs, however, conduct follow-up assessments. Program staff claim that follow-ups are too intrusive to the family or client and too time consuming to carry out. In actual fact, it may be

that program staff are reluctant to conduct follow-up assessments because they fear that most clients will revert back to their pre-program style of living or behaviour, experiencing the same problems that precipitated their referral to the program. Many program workers will even claim that the program is not expected to have a permanent or long-term effect on the client.

Given these problems, it is understandable why quasi-experimental evaluations of operational programs are seldom properly conducted. The quasi-experimental studies which are attempted are frequently compromised or misdirected and, thus, the results of the study also become questionable. Frequently, programs evaluted by such misguided quasi-experimental methods are found to be ineffective in achieving program objectives. Yet, these very same programs may appear to be very successful when evaluated with 'softer' measures such as the self-report satisfaction questionnaire given to the 'graduates' of a program (Lundman, McFarlane & Scarpitti, 1976). Such was the case with the Reckless & Dinitz study (1972). An experimental and control group of male juveniles identified as likely to become delinquents by their sixth-grade teachers were placed in a special educational program. Subjective assessments of success were highly favourable. However, according to objective standards such as arrest rates, dropping out, school attendance, grades, school achievement and attitudes toward self, there were no differences between the treatment and non-treatment groups.

While many evaluators respond to these incongruencies simply by saying that the subjective assessments and 'softer' measures of effectiveness are completely wrong, it may be that the measurements of objective criteria (e.g., arrest rates) do not reflect the 'true' objectives nor the real impact of the program services on clients. It is possible that a youth has been helped by a program but that this may not be evident in terms of whether or not attendance at school is regular. Or, it may not be evident in terms of complete cessation of delinquent behaviour. But later, as an adult, the program's intervention may be an important factor contributing to the individual's success in acquiring a stable job instead of adopting crime as a means of livelihood.

It is difficult to measure a youth's behaviour and attitudes reliably given the clumsiness and crudeness of our measuring instruments. Extensive natural observation of a youth's behaviour may be the most reliable and valid method for recording a youth's behaviour but this is very time-consuming and intrusive, and it requires professional observer training. Observer ratings and self-report questionnaires are much quicker methods for collecting information, but they are prone to biases and distortions and, most importantly, the specific items and scales may not be uncovering the youth's true behaviour and attitudes. For example, asking a youth whether "he likes himself" may not be a good indicator of his true feelings of self worth. The self-report questionnaire method is dependent, first, on the individual being aware of his true behaviour and attitudes and

secondly, on his degree of truthfulness about these things. The observer rating method is dependent on the observer being able to observe the relevant behaviour and then being able to interpret it properly according to the rating scale provided. In addition, the rating scale may not be able to represent fully the complexities and dimensions of the behaviour being observed.

4. Inadequacy of the Theoretical or Conceptual Model

Despite the seriousness of these technical difficulties and the socio-political difficulties of carrying out experimental or quasi-experimental evaluations of operational programs, a more serious problem concerns the adequacy of the conceptual or theoretical delinquency prevention model (Hackler, 1978). Most operational juvenile prevention programs assume that a child's delinquent or problem behaviour is merely symptomatic of underlying problems or conditions.

Psychological models assume that these underlying antecedent conditions pertain to factors such as poor self-esteem, inadequate social skills, emotionally conflictual parent-child dependency relations, attention seeking needs, etc. Sociological models, instead, examine the effect of aberrant social conditions such as exposure to a criminal or delinquent peer group, poverty conditions, unstable family arrangements, poor schools, lack of education and occupational skills, etc. Preventive techniques or treatment strategies are oriented towards the alleviation or remediation of any of these underlying or antecedent conditions, whether it be individual or family counselling,

educational or occupational training, psychotherapy or recreational experiences.

Each type of preventive delinquency program adopts its own theoretical or conceptual rationale linking specific antecedent conditions to socially problematic behaviour and then to delinquent or criminal behaviour. Specific program activities or services are then provided to the youth or family to treat these underlying conditions.

Evaluation researchers have typically measured effectiveness of a program in terms of delinquency rates such as the number of police contacts or re-arrest rates. However, if the theoretical delinquency model is faulty, that is, the delinquent behaviour is not a result of the underlying conditions proposed, or even worse, antecedent conditions have been correctly diagnosed but treatment of these conditions is in no way related to the diagnosed delinquent behavioural problems, then obviously the evaluation results will also be faulty. For example, while the Cambridge-Sommerville program, which focused on improving the self-esteem of juvenile delinquents quite successfully achieved this objective, it did not have a long-term preventive effect on the youth's future tendency to engage in delinquent acts. There are so many factors influencing a youth's behaviour at the time he or she engages in disruptive or delinquent behaviour, it may be there is no particular service or treatment that can have much of a long term effect inhibiting further acts of delinquency. However, a program, such as a family counselling program, may have an immediate and dramatic (possibly long-term) effect on improving family relations and improving the patterns of family communications.

Using police contacts or arrest rates as the criteria for program effectiveness, therefore, may be faulty because of an inappropriate theoretical rationale underlying the problem. In addition, police arrest rates may be faulty as a criterion for program effectiveness because there may be other intervening variables affecting how many and which youths are identified and apprehended by police or which youths are perceived by social service personnel and schools as troublesome children. The fact that a youth has been identified may predispose significant individuals in contact with him to view his or her behaviour differently; that is, to view all negative behaviour as delinquent behaviour. Therefore, it is possible for a young child's behaviour to remain the same or improve, but parents or teachers or police may focus on the remaining negative or problem behaviour and, therefore, describe him or her as more delinquent.

D. SELECTING AN APPROPRIATE EVALUATION DESIGN - A PRACTICAL GUIDE

1. When is the Quasi-Experimental Evaluation Design Appropriate?

Considering the problems of conducting a methodologically sound and successful quasi-experimental study of operational programs, several researchers (Hackler 1978, Wholey, 1978) now are arguing that it may be more appropriate and useful to conduct extensive process analysis. Hackler (1978) points out that by trying to conduct a quasi-experimental evaluation study in a setting that is resistant or inappropriate for experimental manipulation, compromises are made and bad feelings generated that jeopardize the worth of the study.

While Hackler takes an extremist position that the quasi-experimental and experimental evaluation strategy should be discarded completely and that only "soft" process analysis should be conducted (if at all), this does not allow for the fact that there may be a place for quasi-experimental research, under more selective conditions than the conditions under which they have been previously attempted. One must also be aware that despite the merits of process evaluation and less rigorous outcome evaluation, only the experimental evaluation method can test the validity of the theoretical assumptions underlying the program - that a particular type of service or treatment will result in specific effects or client behavioural changes. It is possible that an experimental study may reject the theoretical hypotheses of the program model while a process evaluation may indicate that the program has been successful in terms of other dimensions of equal or greater value. If this occurs, then it is much more reasonable to conclude that the program objectives need to be changed to reflect reality, rather than to conclude that the program is ineffective and should be discarded.

Suchman (1970) provides a useful model for determining when a quasi-experimental evaluation study is both necessary and feasible, and when a process evaluation would be more appropriate. He observes that a social service program can be classified into four stages of development: (1) an initial research phase, (2) a planning phase, (3) a demonstration project stage and (4) an operational stage. He argues that evaluation can occur at each of these stages of a program's development. He explains:

> "We must evaluate 'research' in terms of its ability to provide the necessary knowledge-base for planning; in turn, 'planning' may be evaluated according to its success or failure in developing a program which can be tried out on a demonstration basis. Similarly, the demonstration program can be evaluated in terms of its utility for the establishment of an operating program, while the operating program becomes evaluated according to its effectiveness in achieving the desired objective." (p. 58)

Suchman believes that the experimental and quasi-experimental 'evaluation' is most important for certain kinds of 'demonstration projects' but not at the planning or 'operational' program stages. Demonstration projects can be classified into (1) pilot programs; (2) model programs; and (3) prototype programs. The pilot program represents a trial-and-error period during which new approaches and new organizational structures or procedures can be tried out on a very flexible and easily revisable basis. Programs, at this stage of development, should be evaluated by means of a process analysis to determine the successfulness or adequacy of each procedure. The pilot project stage requires exploratory research for the purpose of developing a program which can then be evaluated in a more systematic way. The end result of modification and evaluation of pilot projects is a 'model' program.

The experimental or quasi-experimental research study becomes applicable only for evaluating the model program. It is important to demonstrate that the theoretical assumptions underlying the program are correct and that the program objectives can be achieved given certain circumstances. A carefully controlled experiment is in

order. The program input must be highly structured and well-defined. The program services or activities must be specifically defined and delivered in a precise manner to each client or group of clients. Experimental and control groups need to be closely matched.

An appropriate criterion of program effectiveness must be defined and validated, and reliable and valid instruments for measuring behaviour with sufficient precision must be constructed. Pre-program and post-program assessments of client behaviour need to be made. Other extraneous intervening variables that may affect the delivery of the service or the impact of the service must be eliminated or controlled.

The third type of demonstration program - the prototype program - is an attempt to implement the experimental program on a practical and realistic basis given the available resources and the community characteristics. The evaluation design can attempt to approximate the experimental approach and should compare the new prototype program with traditional programs as controls. But since the prototype and traditional programs must be carried out under normal operating conditions, if one is to be able to generalize the findings, vigorous controls over matched experimental and control groups may not be possible. Suchman emphasizes "it is absolutely essential for the prototype program to be evaluated under conditions as similar as possible to the proposed operational program for the results to be applicable to these programs" (p.62). Given this, however, the evaluation research component must be able to determine why the prototype program was a success or failure and to specify what aspects of the program

were more successful than others, and among which population subgroups. A less rigorous quasi-experimental design or comprehensive process evaluation of program input and output would be appropriate for the 'prototype' program.

2. The Appropriateness of a Process Evaluation for 'Operational' Programs

Once a program is operting as an established service in the community, Suchman argues that the focus of interest for an evaluation of the program should be upon the improvement of services rather than upon whether or not a service is worth keeping. Often, a program that has been operating in the community to meet a health, education, or welfare problem will continue to operate regardless of the results of any evaluation study. The political and social pressures to maintain the service regardless of its effectiveness are considerable. Therefore, the primary focus of an evaluation of this type of program should be on "What is happening?" and "How can it be made better?" This type of evaluation design is generally known as a process evaluation (also called formative evaluation).

A process evaluation is concerned with describing all components of a program and with studying the relations between these components. The focus of an evaluation is upon the day-to-day operation of the system as a whole. The administrative, structural, and environmental input, and financial and staff resources can be identified and analyzed in terms of how they facilitate the delivery of services. The program activities must be described and analyzed in terms of

whether they are relevant to the attainment of defined program objectives. Program objectives must be assessed in terms of their relevance to the needs and characteristics of the target population. The delivery of services must be evaluated in terms of time and cost estimates related to the attainment of desired program objectives. According to Suchman, information obtained from such an evaluation process must be used in a feedback mechanism to make decisions about future program directions and revisions.

Suchman's characterization of social service programs and of the role of evaluation in terms of stages of program planning, development and operation is useful to researchers engaged in the task of conducting an "assessment of the evaluability of a program". An assessment of the evaluability of a program is a preliminary step to the design of a post-hoc evaluation of an operational program.

3. Assessment of the 'Evaluability' of a Program

(a) General Issues

It is generally the case that most programs selected for evaluation can be characterized as 'prototype' or 'operational' programs. (Although it is not necessarily true that a 'model' program was first developed and experimentally validated.) For most of these types of programs, it would be inappropriate to conduct a quasi-experimental study: a process evaluation would be more practical and more likely to address the concern of funders and program managers. However, certain conditions are necessary before one begins to even consider the practicality of conducting a process evaluation. Even given these

basic conditions, other information is needed to enable the researcher to design the evaluation study. These preliminary steps are often referred to as a program "review" or an "assessment of the evaluability of operational programs" (Wholey, 1978; Sanderson, 1979; Suchman, 1970).

Suchman (1970) raises some important points concerning the evaluability of programs. Since dissatisfaction and puzzlement lie behind most demands for evaluation, the first step in an evaluation study, according to Suchman, should be to identify as clearly as possible the sources of "dissatisfaction and puzzlement". He goes on to state that an evaluation study should have a clear-cut relationship to some decision-making function. Unless it can be reasonably expected that the results of the evaluation will be utilized to make decisions about the program or its objectives, there is probably little need for the evaluation.

Another preliminary criterion for considering an evaluation study concerns the appropriate and desirable time for evaluation. Suchman argues than an evaluation study should not be undertaken until an activity or program has had enough time to stabilize and prove its possible effectiveness. However, the program should not be evaluated so late that results cannot be applied to the operation of the program. Programs that have been operating for a long time often become so entrenched that change is no longer possible or can only be made with great difficulty and disruption. It might be better not to evaluate programs that are highly entrenched or at least not to introduce

new data collection procedures. However, analysis of existing data maintained by the program may be useful. This level of analysis for established programs is possible assuming that client intake and case management information is maintained. Many programs, however, have poorly defined objectives and maintain no information about their clients or their operations. An evaluation analysis can not be accomplished without an adequate information system operating in the program. It may be best to leave these types of programs completely alone or help to introduce a system for collecting client information at intake. Once there is adequate information available on what is happening to clients referred to the program, then an evaluation might be possible. (The design of an information system necessary for conducting an evaluation study is discussed later in this report.)

One major problem confronting evaluators concerns the lack of agreement among management or policy makers or both about the objectives of the evaluation research. This often leads to an under-utilization of the study. Wholey (1978) recommends that an analysis of the decision-making system be made prior to the initiation of the study. This would require examining which parts of a program are sufficiently defined and stabilized to warrant evaluation, and collaborating with program managers to clarify the program and its goals, and to identify potential side effects not specified in the objectives. It is important for funders, program managers, and researchers to understand the needs of each other and to come to an agreement concerning the research objectives. Lack of understanding on how the collected

information will be used by the decision-makers is a potential source of conflict. Program managers and the policy analysts may require the same data but use the data for quite different purposes. Thus, misunderstandings on how the evaluation results will be used can lead to disappointment and bitterness after the study is concluded.

An analysis of the decision-making and information-using needs of the program managers, funders, policy analysts and researchers will go a long way to forestall or at least to anticipate sources of conflict. Other decision making issues that need to be resolved in advance concern (1) how and to whom will the results be distributed; (2) who owns the data; (3) how to set up mechanisms of feedback to the agency and funder during periodic stages of the evaluation; (4) the setting of a realistic research time and necessary resources; (5) how to maintain the confidentiality of client records; (6) promoting the researcher's awareness of the program's socio-political environment; and (7) how much involvement and time will be required by program staff.

(b) Wholey's system for assessing 'evaluability'

Wholey (1979) offers a procedure for conducting an assessment of the decision-making process and the evaluability of a program. This system involves a sequential series of steps that "bound and refine" the program from two perspectives, that of the user (intended audience for the evaluation) and that of the evaluator. The steps are:

"1) Bounding the problem/program: Determining what federal, state, or local activities and what objectives constitute the program - what is the unit to be analyzed?

2) Collection of program information: Gathering information that defines the program's objectives, activities, and underlying assumptions.

3) Modeling: Development of a model that describes the program and the inter-relationships of activities and objectives from the point of view of the intended user of the evaluation information.

4) Analysis: Determining to what extent the program definition, as represented by the model, is sufficiently unambiguous that evaluation is likely to be useful. This step also includes the identification of potential evaluation studies.

5) Presentation to management/intended user: Feedback of the results of the assessment to representatives of management/intended user and determination of next steps that should be taken." (Wholey, p. 49-50).

Wholey emphasizes that the bounding exercise should not only identify the intended users of the evaluation but also include an examination of which levels of management and policy should be included in defining the program to be evaluated. It may be necessary to design several feedback loops for presenting findings to the different interest groups.

Once the boundaries of the program to be evaluated have been specified, information must be collected on the objectives, type of program activities and any statements describing assumed causal links between activities and objectives. In most programs, there will be a set of objectives related in some hierarchical order or on some time-phased basis. Objectives should be defined as precisely as possible in terms that are measurable. Judgements as to the adequacy or measurability of objectives would not be made at this stage. There are several flow models available for organizing the relationship of inputs to activities, to objectives, and to outputs. Wholey offers a

sample "user survey" for interviewing program managers to obtain the manager's view of the program and its objectives; in particular, the manager's definition of events, how he would measure each event, whether he accepts measures currently in use, and his assumptions linking events:

"1. What are the objectives of the program?

2. What would you consider acceptable evidence of achievement of program objectives?

3. What mechanisms exist (policies, guidelines, staff activities, etc.) to achieve the above objectives?

4. Why will Event A lead to Event B?

5. What does (the funder) expect of the agency/program in terms of performance? Are they consistent from year to year?

6. What do you perceive as the most serious difficulty facing the program in terms of meeting its objectives?

7. What performance information do you need on the job?

8. If you had the above information, what would you do with it?

9. Have you seen the program's present information systems? Is it adequate for your needs?

10. How do you get the information you need to do your job? How satisfied are you with this information?

11. What do you consider to be the most important thing that you must accomplish in the next year? What information do you think you need? How will you get it?

12. What are the most important issues or questions that you believe an evaluation of the program should answer?"(p.56)

The next stage of analysis is to develop a model that graphically represents the important intended relationships among program

activities and objectives as cited in the program documentation and interviews. The "rhetorical program model" (Horst, 1974) is a type of flow model that can represent all the activities and their objectives as defined by managers, policy makers and other intended users, without inserting what might appear to be 'missing' or 'necessary' objectives or activities.

From the evaluator's perspective, two types of analysis should be applied to the rhetorical model: (1) Are the objectives stated in measurable terms? and (2) Are the assumed causal relationships testable? Here "measurable" means that there exists agreement on the part of management/intended users as to what would constitute or signal success. If the program manager is unable to define what he or she wants the program to accomplish and what evidence is needed to determine this, then the objective is eliminated from the "rhetorical model" or classified temporarily as "unmeasurable". This type of examination should be applied to all objectives - process objectives associated with administrative activities as well as program impact objectives. Any assumption that a program service is causally linked to a program objective is testable only if there is pre and post program data on the treatment group and if there is data on a non-treatment comparison or control group. This allows the manager to determine that a pre-to post-program change has occurred and to determine that the observed changes occurred as a direct result of the program services, not some other reason (e.g., maturational change).

Then, a final evaluable model is developed which is a subset of the rhetorical model, retaining measurable objectives and plausible, testable assumptions, and representing that portion of a program which is ready for useful evaluation.

The next stage of the evaluation design defines the information that can be collected about the program as it is represented by the evaluable program model. The most important aspect about this type of evaluability assessment is that the evaluator does not design an evaluation strategy based on the theoretical model and attempt to collect data to determine whether all the theoretical objectives of the program have been attained. Rather, the program managers and policy makers define their own objectives and direct the evaluator to design the evaluation strategy. If program managers and policy makers are able to define only one objective in measurable terms and are able to specify only one part of a causal linkage of activities to objectives, then this is all that should be evaluated. Thus, it is possible, after analysis of the rhetorical program model, that both the evaluable program design and the information collected will be extremely simple - maybe trivial. In itself, this fact may be important feedback to the manager or policy maker.

c) Sanderson's System for Assessing 'evaluability'

Sanderson (1979) offers a similar set of procedures for conducting evaluations of operational programs. She defines three phases through which the evaluation process should proceed: (i) a contract

development phase; (ii) a research phase; and (iii) research termination. (Only the first phase will be discussed here.) Two functions are performed in the contract development phase: (1) an assessment of the feasibility of doing any kind of evaluation of the program and, if evaluation is feasible (2) negotiating the preliminary steps development procedures of a contract to conduct the research.

According to Sanderson, the objective of the contract development phase is to decide whether an evaluation is appropriate at this time, and if so, to produce a formalized written research contract, signed by the program manager, researcher and funder. The task of writing a research contract forces many of the issues described by Wholey into the open. If program managers, funders and policy makers are unable to agree on the objectives of the program in terms of measurable criteria, then it will be impossible to formulate a research contract to be signed by all. The steps of Sanderson's contract development phase are summarized below:

1. The political context of the evaluation must be made explicit.

2. The environmental context in which the program operates must be described.

3. The program objectives and goals must be specified and clarified. Goals are defined as a general statement of purpose for the organization. Objectives must be stated precisely in terms of measurable end results. The objectives guide the agency's activities.

4. The research objectives and goals must be established (This should be considered in light of Wholey's evaluability assessment.)

5. Sources of information which will be available to the researcher must be identified. Data can be collected from documents, observation of actual events or interviewing people.

6. The methods to be used to collect information from these sources should be specified.

7. Anticipated effects of the proposed research methods on agency operation should be presented.

8. The research resources necessary to carry out the research should be specified.

9. The agency and funder should be required to guarantee that research resources are currently available.

10. The project duration must be established for each stage of the evaluation.

11. The cost of the research should be specified for each research component.

12. Funders, program managers and evaluator must be able to agree to all of the above conditions. If not, the evaluation plan will need to be modified or simply not carried out.

This contract development phase will serve to determine the feasibility of conducting any type of evaluation. The political-social environment in which the program operates must be receptive for an evaluation and the results of the evaluation must serve clearly defined purposes for program managers and policy makers to aid program improvement and development. The research contract can specify the parameters of data collection involved in a process analysis. Generally, this involves a description and analysis of the administrative structure, environmental context, resources, clients, administrative procedures, program activities, and possible effects of the program on clients and the community.

E. ESTABLISHING AN INFORMATION SYSTEM FOR PROGRAM MONITORING AND EVALUATION

A major component of the assessment of the evaluability of an operational social service program involves conducting an assessment of the adequacy of the program's information system to collect information relevant for an assessment of program efficiency and effectiveness.

Within any social service program, information is needed for a variety of purposes: (1) for determining optimum treatment or service delivery and monitoring the client's progress, (2) for monitoring the efficiency and effectiveness of program procedures, (3) for measuring the achievement of program and client objectives, (4) for monitoring acquisition and use of program resources (funds, costs, staff time, etc.) and (5) for specific research questions.

These information functions can be classified into four types: (1) client treatment and monitoring, (2) management of resources and monitoring of program procedures, (3) impact evaluation, and (4) research. Different information is generally collected to satisfy each of these functions (although in some areas there is overlap). Likewise, different record-keeping systems are usually employed to collect each type of information. However, most data-collection systems are cumbersome, redundant at times, and often inappropriate for the information needs of management and staff. In addition,

information recording is generally inaccurate, sporadic, and non-standardized. In such cases, information retrieval and utilization is limited.

Most program managers consider information for program evaluation to be a separate (one-shot) collection of data on program impact - specifically, on changes in client behaviour. This information collection function is considered separate from the day-to-day collection of information for administrative purposes. However, it is argued in this report that an evaluation of the successfulness or impact of 'operational' social service programs requires assessment and analysis of all aspects of the program's operation as well as an assessment of the program's effect on client behaviour. A program cannot hope to be efficiently and effectively meeting the needs of the client population if there are problems in the delivery of the service or problems such as poor staff attitudes, inadequate financial resources, insufficient staff, an unsupportive community environment, etc. It is important to know what type of services are being delivered, to whom they are being delivered, how they are being delivered, and what kind of resources are being used to deliver these services.

This type of process monitoring and evaluation, therefore, requires collection of extensive amounts of information. A record-keeping system must be installed to collect information in a comprehensive, efficient, systematic, and accurate manner. In addition, the information must be collected in an organized and standardized format

such that retrieval and use of this information for analysis and later decision-making is economical and effortless. For example, summary statistics on client characteristics should be tabulated automatically in an accumulative count at the time the information on each client is first recorded. Examining client files retrospectively to compile summary statistics is time-consuming and error-prone.

If the existing information system for an operational program is inadequate to meet the needs of management and program evaluation, then a new or revised system must be designed. The first step in the preliminary design phase of an evaluation or management information system (MIS) is to formulate organizational goals and client objectives. Next, one must determine the questions to be addressed by an information system for assessing the achievement of organizational goals and client objectives. In doing this, the information needs of the direct service staff (e.g. counsellors), the administrators and the funders, as they relate to treatment monitoring, management of resources, and evaluation must be considered. The third major step is to determine what data elements need to be collected to respond to these questions. In addition, data capture instruments such as questionnaires or tests must be created to collect this data. (Occasionally, standardized instruments are available commercially.) Finally, report requirements, mechanisms to control data recording errors, and the cost of the system must be considered.

The next phase of a management information system design is to integrate the required new MIS documents into the existing information system documents. New input forms must be designed to record the data. Paying special attention to how the data is to be coded and how the forms are to be laid out can enhance the accuracy, ease and completeness of data recording. Whether the data is to be recorded and statistically analyzed manually or by computer is also an important factor affecting the design of the input documents.

The final phase of information system design involves implementing and validating the system. Forms must be pretested, orientation and training sessions for those using the system must be planned, computer programs (if any) must be developed and tested, decisions about inclusion of current caseloads and historical data must be made, and collection of the relevant data must begin.

In summary, without the design and implementation of a systematic and efficient information system, a monitoring and process evaluation of the program is not possible. Subsequently, it would be impossible to determine whether the program is operating as designed, whether it is effective, and if not optimally effective, how it could be made more effective. Therefore, one of the first steps in designing an evaluation study of an operational social service program is to make an assessment of the adequacy of the information system and if not adequate, to design and implement an effective information system.

FOR EVALUATING A DELINQUENCY PREVENTION, YOUTH AND FAMILY COUNSELLING PROGRAM

As an example, let us consider some preliminary steps in the design of a management information system for evaluating a family counselling program - a program which provides assistance to minor delinquent and problem behaviour youth that have come into contact with police. Information on this program, Langley Youth and Family Services, is available in a B.C. Ministry of Attorney General report - A Process Evaluation of a Juvenile Delinquency Prevention Program: Youth and Family Services. (Rowe, 1981) The information required for program evaluation and the specific data elements needed to be collected will be described here.

The first step in the design of a MIS for program evaluation is to describe all the goals and objectives. The primary goal of the Langley Youth and Family Services program is to resolve family and interpersonal problems that may be contributing or predisposing youths toward juvenile delinquency.

In order to facilitate this program goal there are specific procedural objectives for the program:

(1) provide an immediate response (within 2 days) to a youth or family that has expressed a need or has a crisis;

(2) divert from the juvenile justice system youths under 13 years of age;

(3) persuade both the youth and the family to accept LYFS counselling;

(4) make an assessment whether there are underlying family problems that may be contributing or predisposing the youth toward problem behaviour;

(5) provide short-term family counselling (3 months) for problem behaviour and minor delinquent behaviour youths and provide counselling over longer periods of time only in exceptional circumstances;

(6) refer elsewhere families who require long-term intensive counselling or special services;

(7) Follow structured case management procedures:
  (i) At intake, collect biographical information on the child and family
  (ii) Contact parent by letter and make an appointment
  (iii) Conduct preliminary interview
  (iv) Obtain parent consent form to obtain information from schools and other involved agencies
  (v) Contact referral source informing them referral was received
  (vi) Conduct an assessment of the youth problem behaviour and underlying family problems (if any)
  (vii) Determine treatment strategy or refer elsewhere
  (viii) Establish and conduct counselling
  (ix) Provide sustaining counselling
  (x) Inform referral service that case was closed
  (xi) Conduct a follow-up assessment at specified periods of time.

Some broad impact objectives for the program relevant to the overall program's goal are as follows:

(1) through counselling for approximately three months, inhibit the youth from committing a delinquent offence (a time period can be specified);

(2) improve family communication patterns;

(3) eliminate the youth's negative behaviour in the home and school; and,

(4) increase youth and family self-esteem.

Given the program goal, the procedural objectives and the impact objective for receiving clients, providing assessment and counselling, and addressing client problems, there are a number of specific information questions that need to be addressed by the management information system. These are as follows:

I Client Assessment and Treatment

(1) Who is this client? Why has this person been referred? (Client Characteristics)

(2) What are the client's problems and are there underlying family problems? (Client and Family Assessment - Does it need to be modified? How?)

(3) What kind of action should be taken: no treatment, brief service, refer elsewhere, counselling? (Service Selection)

(4) Is the treatment plan working? When should it be terminated? (Treatment Monitoring, Review Modificaton and Termination)

II Implementation and Achievement of Procedural Objective

(1) Have all case management procedures been followed? e.g. intake information taken, parent letter sent, parent consent form signed, referral letter sent, assessment conducted, counselling provided or client referred elsewhere, file closed, and follow-up assessment conducted. (Case Management Procedures)

(2) Was there an immediate (within 2 days) response to a referral? (Immediate Referral Response Objective)

(3) Are both the youth and the family involved in counselling? (Family Participation Objective)

(4) Are the clients receiving approximately 3 months of counselling? (Three Month Treatment Objective)

(5) Are youths and families who require long-term counselling being referred elsewhere? (Refer Elsewhere Long-term Counselling Objective)

III Achievement of Program and Client Objectives

(1) Are client and family problems being resolved? (Problem Resolutions Objective)
e.g. (a) Is there an improvement in family interaction and communication patterns?

(b) Is there an improvement in the youth's behaviour at school?

(2) Are there recurrences in problem or delinquent behaviour? (Cessation of Problem or Delinquent Behaviour)

In addition to these information questions that are related to the specific goals and objectives of the program, program managers and the funders require information on the cost and use of program resources. This information would enable administrators to conduct a cost-benefit analysis on the program as part of their evaluation of the efficiency and effectiveness of the program.

Table 1 provides a summary of these information questions and of the data elements that need to be collected to address these questions. These questions are relevant to this particular counselling program with these specific goals and objectives. Different questions may be relevant to other types of social service programs.

Formulating the information questions and selecting the data elements to be collected are only first steps in the design of a management information system. Designing documents to record this information in a format that can be easily coded and analyzed (manually or by computer) is a complicated task. Likewise, implementing a data collection system that is cost-efficient and simple enough for all staff to follow systematically for each client without

TABLE 1: Example of Information Questions and Data Elements for a Delinquency Prevention, Family Counselling Program

	Information Questions	Data Elements
I	Client Assessment and Treatment 1. Client Characteristics? 2. Client and Family Assessment? 3. Service Selection? 4. Treatment Monitoring, Review Modification and Termination?	biographical data (sex, age, family members); developmental history of youth; family interaction patterns; school behaviour; emotional-social behaviour of youth; reason for referral; previous treatment received; therapist treatment plans; client behaviour in treatment; client attitudes and responsiveness to treatment.
II	Implementation and Achievement of Procedural Objectives 1. Case Management Procedures Followed? 2. Immediate Response to Referral? 3. Full Family Participation? 4. Three month Treatment Objective? 5. Referred Elsewhere when Long-term Counselling Required? 6. Assessment of underlying causes of youth's behaviour conducted? 7. Follow-up assessments conducted?	date of problem; date referral received; date of client contact; intake information taken (yes, no); assessment conducted (yes, no); type of action taken; date file closed; date file opened (if at all); date follow-up assessment conducted; which family members attend counselling; duration of treatment; number of contacts with client.
III	Achievement of Program and Client Objectives 1. Family Problem Resolution? 2. Cessation of Problem or Delinquent Behaviour?	pre-program and post-program description of youth social behaviour; family communication patterns; number of delinquent offences prior to referral; number of delinquent offences after treatment terminated; response to follow-up assessment.
IV	Use and Cost of Program Resources 1. Cost per client counselling session? 2. Proportion of staff time spent in counselling versus administrative support activities?	program staff salaries; staff time spent on counselling per client per month/year; cost of supplies, equipment; total cost per client; total cost per hour of counselling.

disrupting counselling and program effectiveness is also a difficult task.

G. SUMMARY

In summary, let's consider some of the issues affecting the evaluation of operational social service programs. It is the position of Hackler, Suchman, and this author that the quasi-experimental design is a completely inappropriate strategy for evaluating the 'operational' social service program. Numerous problems have been cited. When such quasi-experimental evaluations have been attempted not only does such an evaluation design not answer many of the questions it is attempting to ask, but the results that are obtained are often misleading and uninformative. The quasi-experimental impact evaluation design is only appropriate when integrally incorporated into the design and implementation of an experimental demonstration project in order to test whether specific objectives are achieved as a result of specific treatment or intervention activities. Full cooperation and understanding of research purposes must be also be present.

No one takes the position, however, that the 'operational program' should not be evaluated. Rather, a non-experimental type of evaluation, but equally valid as the experimental designs, is in order; that is, one that is concerned with assessing the efficiency and effectiveness of the program in terms of identifying the process

of operation, determining the adequacy of the process of operation relative to the theoretical or conceptual expectations for the program, and assessing all possible attitudinal and behavioual effects of the program on clients, community and staff. The effectiveness of a program can be evaluated in terms of the effects of a program on clients and the community, and in terms of the adequacy of the program's process model to maximize client effects and minimize interfering problems. It may not be possible, however, nor may it be relevant to ascertain the causal linkage between program activities or treatment and attainment of program objectives.

In order to conduct this type of process evaluation, Suchman, Wholey and Sanderson provide some useful guidelines for establishing a cooperative atmosphere between administrators, program staff and evaluator researchers, and, for collecting and analyzing the data. Following these procedures constitutes an "assessment of the evaluability of a program". A crucial stage in assessing the evaluability of a program and designing the evaluation strategy is the negotiation of a "research agreement" between program staff, administrators and researchers. This action will effectively identify any problems for research implementation that could obstruct the research process or result in misleading data being collected and reported. The next important action to take in conducting an evaluation of an operational program is to examine the adequacy of the program's information system and, if necessary, to design a comprehensive management information system.

H. IMPLICATION OF EVALUATION ISSUES FOR JUVENILE DELINQUENCY PREVENTION PROGRAMS

It is apparent that the quasi-experimental effectiveness research strategy is an inappropriate strategy for evaluating juvenile delinquency prevention programs that have been operating in the community for a number of years. Rather, a more appropriate strategy would be to analyze the process of operation and the effects of the program on clients and community, without disrupting or modifying the program's operation. A process evaluation research strategy provides the means to document the interrelationships of program services and other staff activities and assess their effect on clients, to determine the influence of environmental factors on program operational efficiency and effectiveness, and to evaluate whether the actual program is operating and offering outcomes in accordance with the conceptual plan. Most importantly, an efficient and effective process evaluation provides information necessary for program development and modification. Experimental and quasi-experimental evaluation designs do not yield this type of information. Conducting an 'assessment of the evaluability of the program' is a first step in determining the nature and extent of a process analysis; for in many cases the goals and objectives of a program are so poorly defined and so few records are kept of the type of clients entering the program that even an analysis of the program structure and process of operation could be premature.

Let us consider juvenile delinquency prevention programs in terms of Wholey's (1979) steps for conducting an assessment of the evaluability of a program. Attempting to determine the goals, objectives, and activities of a juvenile delinquency prevention program is not an easy task when one considers the points of view of the social service system, the justice system and the community. Each of these systems has a different perspective on how a delinquency prevention program should operate, what its goals should be, and even the type of clients that the program should be servicing. The police are concerned with preventing further criminal behaviour among juveniles who have committed previous delinquent acts. Thus, they are concerned with assessing the effectiveness of a program in terms of recidivism rates. Social workers, however, are more concerned with assessing and diagnosing social, emotional, and psychological problems a youth may be having in the school, home, or community. They may believe that these problems can predispose or lead to a youth's involvement in delinquent activities, but they are concerned more with treating the underlying conditions affecting these problems rather than with ensuring that the youth does not commit further delinquent offences. Thus, from the perspective of the social worker, traditional program effectiveness is seen to be a function of the success of the program in addressing the antecedent problem conditions.

Parents may evaluate the effectiveness of the juvenile delinquency preventive program in terms of whether it helps them cope

more effectively with the youth at home. Schools may evaluate the effectiveness of the program in terms of whether it helps them cope with the problem behaviour of a youth at school. If through the program's intervention a youth attends school more regularly and gets into fewer fights with classmates, then the program is considered successful. When store security officers have some place to refer a youth who is presenting a problem in their store, (such as loitering, or is only suspected of engaging in shoplifting) then these people also feel they have been helped. Their workload has been reduced and their helplessness in knowing what to do with a problem youth has been alleviated.

Thus, the criteria for determining the effectiveness of a juvenile delinquency prevention program can vary depending on the perspective of the program users. In addition, even program staff who have the same general perspective will often disagree on program objectives and their relative importance. This makes evaluation difficult.

Most juvenile delinquency prevention programs provide a general strategy for treating or coping with their clients, such as family counselling, recreational activities, or vocational and life skill training. However, there is considerable variation in the specific objectives and activities designated for each client. For example, in one family a youth may be experiencing extreme difficulty in achieving positive recognition from his father and may be 'acting out' or

committing delinquent acts in order to get his attention. In another family, a youth may be reacting in anger to the constant marital conflict in the family. In another family, the youth may experience no problems at home but, at school, frustration with academic demands or peer group pressures may prompt outbursts such as pulling the school fire alarm bell or frequent fighting.

While counselling may be provided to all of these children and their families, the specific plans and activities designed to alleviate the problem, to modify each youth's behaviour, or to resolve each marital confict are quite varied.

To identify all these elements of a program's operation, and to represent them accurately from the point of view of all users of the program is a complicated but highly informative task. It would be impractical and possibly irrelevant to attempt to take only two or three of these objectives from one user's point of view, conduct a quasi-experimental evaluation of the program and assess the effectiveness of the program only in terms of these limited objectives. It is a major, highly relevant, and possibly sufficient evaluative task to identify and represent all the objectives, activities, target population, and environmental inputs of a program from the point of view of all users.

This report has attempted to present some of the issues relevant to an evaluation of 'operational' juvenile delinquency prevention programs. Some practical guidelines have been offered to begin the task

of designing an 'evaluation study' of such programs. These issues and guidelines were highly relevant during the design of the process evaluation of the delinquency prevention program, Langley Youth and Family Services. Information arising from a process evaluation of existing operational programs can be extremely useful to program staff attempting to improve a program. In addition, process evaluations provide critical knowledge for the planning and development of new optimally effective delinquency prevention family counselling programs in other communities.

REFERENCES

BRANTINGHAM, P.J.; FAUST, R.L.
1976 "A Conceptual Model of Crime Prevention", Crime and Delinquency, Vol. 22, #3 July, pp. 284-296.

CAMPBELL, D.T.; STANLEY, J.C.
1966 Experimental and Quasi-Experimental Designs for Research. Chicago, Rand - McNally.

FAMIGLIETTI, J.; ENGLISH, D.H.; JANVIER, R.L.
1980 An Assessment of Evaluation of Family Intervention Programs. Designed to Prevent Juvenile Delinquency. National Center for the Assessment of Delinquent Behaviour and its Prevention, Seattle Washington.

HACKLER, James A.
1974 Why Delinquency Prevention Programs in Canada Should Not be Evaluated. University of Alberta.

1978 Prevention of Youthful Crime: The Great Stumble Forward. Toronto: Methuen.

1978 The Dangers of Political Naivete and Exessive Complexity in Evaluation of Deliquency Prevention Programs. Evaluation and Program Planning. Vol. 1, pp. 273-283.

LEWIS, Ralph; DAVIDSON, William S.
1977 Executive Summary: Model Evaluation Project: Youth Service Bureaus in Michigan. Rochville, Md. NCJRS Micro Program.

LUNDMAN, R.J.; McFARLANE, Paul T.; SCARPITTI, F.R.
1976 Delinquency Prevention: A Description and Assessment of Projects Reported in the Professional Literature. Crime and Delinquency, Vol. 22, # 3.

McCORD, William; McCORD, Joan.
1959 Origins of Crime, New York, Columbia University Press.

NORMANDEAU, Andre; HASENPUSCH, Burkhard.
1980 Prevention Programs and Their Evaluations. Canadian Journal of Criminology, Vol. 22, #3, 307-319.

RECKLESS, W.C.; DINITS, S.
1972 The Prevention of Juvenile Delinquency: An Experiment. Columbus, Ohio: Ohio Stat University Press.

ROMIG, Dennis A.
1978 Justice for Our Children. Lexington Mass: Lexington Books.

SANDERSON, Margo.
1979 Negotiating Strategies for 'Evaluation as a Tool for Management', Vancouver, B. C.

SELKE, William L.
1977 A Systems Model Approach to Evaluation of Social Programs: The Case of Youth Service Bureaus. Dissertation Thesis, Michigan State University.

SUCHMAN, E.A.
1967 Evaluative Research: Principles and Practice in Public Service and Social Action Programs. New York: Sage Foundations.

1970 Action for What? A Critique of Evaluative Research in the Organization, Management and Tactics of Social Research, Richard O'Toole (ed.) Cambridge, Mass: Schenkman Publishing Company.

WHOLEY, Joseph S.
1976 Planning Useful Evaluations. Presented at the Evaluation Workshop, Health Resources Administration, U.S. Dept. of Health Education and Welfare, Jan.

WRIGHT, E.W.; DIXON, M.C.;
1977 Community Prevention and Treatment of Juvenile Delinquency - A Review of Evaluations. Journal of Research in Crime and Delinquency, Vol. 14, pp. 35-67.

END