An exploratory study of the evaluative practices of the member agencies of the United Community Services of greater Windsor.

Lam Wo. Wong

University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/etd

Recommended Citation


This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.
AN EXPLORATORY STUDY OF THE EVALUATIVE PRACTICES OF THE MEMBER AGENCIES OF THE UNITED COMMUNITY SERVICES OF GREATER WINDSOR

by

Wong Lam Wo

A research project presented to the School of Social Work of the University of Windsor in partial fulfillment of the requirements for the degree of Master of Social Work

June, 1974
Windsor, ONTARIO, CANADA
RESEARCH COMMITTEE

Dr. Lola E. Buckley  
Chairman

Prof. Forrest C. Hansen  
Member

Dr. Mary Lou Dietz  
Member

Mr. Michael J. Pfaff  
Member
ABSTRACT

The purpose of this research project was to gain familiarity with concepts and methods of evaluative research in general and the evaluative practices of the member agencies of the United Community Services of Greater Windsor, (herein-after, U.C.S.) in particular. It was hoped that this familiarity would help to (a) develop some ideas for the improvement of service evaluation of the U.C.S.-member agencies; and (b) formulate specific problems of service evaluation for more precise investigation.

An extensive survey of the literature was conducted to gain understanding of various aspects of evaluation research in social services. The literature survey was separated into three areas for discussion: functions and types; models and techniques; problems and constraints.

The entire population, that is, all the twenty-seven member agencies of the U.C.S. at the time of the research, was studied as a sample. Data was obtained from twenty-two agencies. It was collected through a structured interview schedule according to a specific procedure. The data collected was centered on describing the following research focus:

(a) the types of information kept in relation to service evaluation;
(b) the methods used in service evaluation;
(c) the purposes for the use of service evaluation;
(d) the value of service evaluation; and
(e) the kind of help needed for improvement of service evaluation.

Among the major findings were:
(a) All agencies kept some types of evaluative information, mainly the statistics that account for time and money spent.
(b) The more sophisticated and refined methods which were essential to the evaluation of effectiveness, efficiency and benefits were the least common in usage and the lowest in preference.
(c) The greatest limitations and difficulties experienced were those related to the skill and knowledge of evaluation research, as well as those related to financial resources.
(d) The kind of help seen as most useful was consultation and guidance from experts to improve service evaluation.

The implications of the findings were discussed and recommendations were made to the U.C.S. and its member agencies.
ACKNOWLEDGEMENTS

The researcher wishes to express his special thanks to the Windsor U.C.S. and its member agencies for their participation and cooperation which made this research a reality.

He is especially indebted to each of the Research Committee: to the chairperson Dr. Lola E. Buckley who has long been his faculty advisor, for her continuous personal care, support and push until some potentials have been realized; to Professor Forrest C. Hansen for his stimulation and especially his guidance in the use of computer; to Dr. Mary Lou Dietz for her high expectations and encouragement for further development; to Mr. Michael Pfaff for his useful advice on the research settings. The researcher is most grateful for their comments and immediate response to any material presented which made possible the early completion of this thesis. Their sensitivity to the needs of the researcher to rejoin his family as soon as possible is very much appreciated.

Thanks are owing to the following for the informative discussions the researcher had with them: Professor Bernard Kroeker of the University of Windsor; Professor Tony Tripodi of the University of Michigan; Professor Brian Wharf of the MacMaster University; Messers. Gary McCarthy, Ray Mahoney of the U.C.S.; and Mr. John Nugent of the Windsor National Parole Office. Also to Professor Mae Harman who carefully proof read this report, and Mr. Edward Marsh of the Detroit
U.C.S. who rendered great assistance in the pre-testing, the researcher is very thankful.

The researcher also would like to take this opportunity to express his gratitude to the Canadian Commonwealth Scholarship Committee, the faculty of the School of Social Work, and the Public Service Commission of Singapore, for giving him a most valuable experience in life.

Lastly, but most importantly, the researcher wants to thank his wife, Dora and children, Philip and Kathy, for their continued understanding, patience and sacrifice in his absence. It is to them that this report is dedicated.

Wong Lam Wo

Windsor, Ontario
June 21, 1974.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>ii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>iv</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>The Purpose</td>
<td></td>
</tr>
<tr>
<td>II. REVIEW OF THE LITERATURE</td>
<td>10</td>
</tr>
<tr>
<td>Types and Functions of Evaluative Research</td>
<td>11</td>
</tr>
<tr>
<td>Defining Evaluation Research</td>
<td></td>
</tr>
<tr>
<td>Evaluation Research in Historical Perspective</td>
<td></td>
</tr>
<tr>
<td>Need and Functions of Evaluative Research</td>
<td></td>
</tr>
<tr>
<td>Purposes and Reasons for Evaluation Research</td>
<td></td>
</tr>
<tr>
<td>Types of Evaluation Research</td>
<td></td>
</tr>
<tr>
<td>Models and Techniques of Evaluation Research</td>
<td>43</td>
</tr>
<tr>
<td>Evaluative and Non-Evaluative Research</td>
<td></td>
</tr>
<tr>
<td>Models of Evaluation Research</td>
<td></td>
</tr>
<tr>
<td>Techniques of Evaluation Research</td>
<td></td>
</tr>
<tr>
<td>Problems and Constraints on Evaluation</td>
<td>63</td>
</tr>
<tr>
<td>Research</td>
<td></td>
</tr>
<tr>
<td>Methodological Problems</td>
<td></td>
</tr>
<tr>
<td>Socio-Political and Organizational Constraints</td>
<td></td>
</tr>
<tr>
<td>Summary</td>
<td>88</td>
</tr>
</tbody>
</table>
III. RESEARCH DESIGN AND METHODOLOGY

- The Focus
- Definitions
- Population and Sample
- The Setting
- Methods of Data Collection
- The Interview Schedule
- Pre-Test
- Methods of Data Analysis
- Utilization
- Limitations of the Study
- Summary

IV. ANALYSIS OF DATA

- The Characteristics of the Sample
- Types of Evaluative Information
- The Methods in Service Evaluation
- The Purposes for the Use of
  Service Evaluation
- The Value of Service Evaluation
- The Kind of Help Needed
- Summary

V. CONCLUSION AND RECOMMENDATIONS

- The Major Findings
- Implications
- Recommendations

APPENDICES

BIBLIOGRAPHY
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Length of Agency Existence</td>
<td>114</td>
</tr>
<tr>
<td>2</td>
<td>Agency Length of U.C.S. Membership</td>
<td>115</td>
</tr>
<tr>
<td>3</td>
<td>Number of Full-Time Paid Staff</td>
<td>116</td>
</tr>
<tr>
<td>4</td>
<td>Number of Full-Time Paid Professional Staff</td>
<td>117</td>
</tr>
<tr>
<td>5</td>
<td>Total Number of Part-Time Paid Staff</td>
<td>118</td>
</tr>
<tr>
<td>6</td>
<td>Number of Part-Time Paid Professional Staff</td>
<td>118</td>
</tr>
<tr>
<td>7</td>
<td>Annual Operating Budget</td>
<td>119</td>
</tr>
<tr>
<td>8</td>
<td>Types of Services</td>
<td>120</td>
</tr>
<tr>
<td>9</td>
<td>Types of Services by Annual Operating Budget</td>
<td>122</td>
</tr>
<tr>
<td>10</td>
<td>Distribution by Types of Evaluative Information Used by Agency</td>
<td>125</td>
</tr>
<tr>
<td>11</td>
<td>Distribution by the Methods the Sample Used</td>
<td>129</td>
</tr>
<tr>
<td>12</td>
<td>Distribution by the Methods the Sample Preferred</td>
<td>133</td>
</tr>
<tr>
<td>13</td>
<td>Distribution by the Respondents' Ratings on the Five Degrees of Adequacy of Evaluative Methods Used</td>
<td>137</td>
</tr>
<tr>
<td>14</td>
<td>Distribution by Respondents' Preference for Sources of Evaluators</td>
<td>139</td>
</tr>
<tr>
<td>15</td>
<td>Distribution by Difficulties and Limitations the Sample Experienced in Service Evaluation</td>
<td>141</td>
</tr>
<tr>
<td>16</td>
<td>Distribution by the Respondents' Ratings on the Reasons for Evaluation</td>
<td>146</td>
</tr>
<tr>
<td>17</td>
<td>Distribution by the Use of Evaluation Results</td>
<td>149</td>
</tr>
<tr>
<td>18</td>
<td>Distribution by Respondents' Ratings on Degrees of Usefulness of Service Evaluation</td>
<td>151</td>
</tr>
<tr>
<td>Number</td>
<td>Title</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>----------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>19</td>
<td>Distribution by Respondents' Ratings on Various Degrees of Importance of Service Evaluation</td>
<td>152</td>
</tr>
<tr>
<td>20</td>
<td>Comparison of the Mean Values of Ratings on Importance, Usefulness and Adequacy of Evaluations</td>
<td>152</td>
</tr>
<tr>
<td>21</td>
<td>Distribution by Respondents' Ratings on Direction of Change</td>
<td>153</td>
</tr>
<tr>
<td>22</td>
<td>Distribution by Kinds of Help the Sample Needed for Improvement of Service Evaluation</td>
<td>155</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

In the Western Hemisphere there has been a long history of public support for social service programs such as public education, health care, rehabilitation, housing and the like. Over the last decade or two many new programs have been developed in the United States and Canada for victims of social ills such as high school dropouts, juvenile delinquents, drug addicts and the mentally ill. All these services are intended to improve human condition and alleviate its attendant ills. As new problems become visible, the response has traditionally been to set up programs to cope with the problems. Despite all these attempts, "dissatisfaction is vehemently expressed towards almost all aspects of our traditional policies, organizations and approaches for dealing with social problems."\(^1\) How well do the host of programs, new and old, succeed in achieving the goals for which they were established? As Mashkin points out the major evaluations of social programs have yielded negative or "no clear" findings, and these findings have come to be interpreted as saying, the results

anticipated from social policies are not being achieved. For instance, school lunches neither improve the health of children nor their learning capacity. Francis Caro concludes that:

In virtually all areas the increased public and private expenditure of the past decade simply have not appreciably improved the social order, or at least new problems have emerged that are distressing as any that have been ameliorative.³

Beginning with the Johnson administration in the United States, policy analysis and evaluation staffs were installed in a number of federal civilian agencies to carry out planning and budgeting efforts, as an attempt to improve the situation. "Programs must be evaluated to identify those that must be redirected, reduced, or eliminated because they do not justify the taxes required to pay them," declared President Nixon in his 1974 Budget Message to the Congress.⁴

In the Canadian scene, Carter and Wharf, in a booklet published by the Canadian Council on Social Development, remark that "the lack of provision for evaluating the results of social programs is a serious liability for those responsible for program planning, policy decisions and

---

funding. Wallin of Vancouver, strongly advocates that in the social services both industry and government must take stock of their record to date, must be more attentive to the feedback from those for whom the services have supposedly been designed, and must vastly expand the search for more effective programs.

The public and social services are subject to constant scrutiny and criticism from political leaders and the citizenry who increasingly ask why a program deserves public support and why it merits priority. The legislators and those responsible for budgeting public funds have to know what impact that service will have on client and community and must evaluate it in terms of cost. Unfortunately, as observed by Newman and Wilsnack, measurement of achievement is missing from social service repertoire. Measurement is the basis for accountability and accountability is the evidence of responsibility. Evaluation research that measures accomplishments is clearly then a striving towards accountability and responsibility.


Highlighting the importance of evaluation research, Caro (1971) argues that neither the rhetoric of politicians nor the pleas of do-gooders is adequate to guide program development. In the same way, he says, the theories of academicians and the exaggerated statements of efficacy by practitioners do not constitute "an adequate basis for the support and expansion of various human service activities." 

It is evaluation research that can have a major impact on social problems... the adequate assessment of existing and innovative programs can be a vital force in directing social change and improving the lives and the environments of community members. 

However, despite the importance and current popularity of evaluation research, there are very few evaluation projects "which have the elegance of design and clarity of execution which would achieve widespread admiration among social researchers." Guba (1969) is of the opinion that the world of evaluation itself requires reshaping. As quoted by Campbell, Etzioni (1968), Hyman & Wright (1967)

---

and Schwartz (1961) all think that most ameliorative programs end up with no interpretable evaluation.\textsuperscript{12} Weiss believes that much evaluation research falls short of expectations because of organizational constraints on the researcher's ability to apply what he knows.\textsuperscript{13} Mashkin (1973) comments "the methodology of evaluation is still inadequate to serve as an overall policy guide and urges that evaluation should be used with caution."\textsuperscript{14}

Recognition of the importance of the evaluation research has grown rapidly. A good deal of confidence in and expectations for evaluation research have been expressed. Yet, some see that in the actual conduct of the evaluation research there are many difficulties, problems and constraints. In reality, what is evaluation research? What kinds of functions does or can it fulfill? What is needed in order to carry it out? What are some of the models and techniques being used today? What are some of the problems that attend its being carried out? These are some common questions about evaluation research in social services.

\textsuperscript{12}Donald Campbell, "Reforms as Experiments," in Readings in Evaluation Research, p.233.

\textsuperscript{13}Carol Weiss, "Between the Cup and the Lip," Évaluation, Vol. 1, No. 2 (1973), p.49.

\textsuperscript{14}Mashkin, "Evaluations," p.35.
The Purpose

This research project is concerned with this very important and much talked about topic of evaluation research in social services. The purpose of this research project is to gain familiarity with concepts and methods of evaluative research in general and the evaluative practices of the member agencies of the United Community Services of Greater Windsor (hereinafter, U.C.S.) in particular. It is hoped that this familiarity would lead to:

(a) discovery or development of some ideas for conceptualization of some approaches designed to improve service evaluation of the U.C.S. member agencies;

(b) formulation of specific problems of service evaluation for more precise investigation.

The stimulus to explore this area came from challenges experienced during work in Singapore and several observations formulated during a field practice experienced at the U.C.S. while a graduate student of the University of Windsor School of Social Work. The researcher was involved in developing training programs as senior instructor of the National Youth Leadership Training Institute, Singapore. What effects actually resulted from the program and how the effects can be determined are two of the many questions that perplexed him very much at that time. He faced similar questions in attempting to find ways and means of
obtaining more reliable information about the service effectiveness of the community centres in Singapore.

At the U.C.S. the researcher noted that both the reports of the Program and Budget Review Committee\textsuperscript{15} and of the Social Planning Review Committee\textsuperscript{16} recognized the need for and recommended the improvement of the evaluation of services or programs of member agencies. In fact the former suggests that the U.C.S. "Should consider the use of instruments for measuring effectiveness of any given service or program."\textsuperscript{17} However, given these recommendations, what should be done to begin with? What should be the strategy or approach for the U.C.S. to help improve service evaluation of member agencies?

The researcher's second observation was that the questions on service evaluation in both FORM A and FORM B of the U.C.S. Agencies Program and Budget Presentation were very general and unspecific.\textsuperscript{18} It was the opinion of the researcher that a better knowledge of agency evaluative practices would result by developing questions which would


\textsuperscript{17}United Community Services, "Report of the Program & Budget Review Committee," p.17.

yield information more truly reflecting service delivery and need situations towards which the services were addressed.

The researcher further noted that in the agency budget presentations some agencies expressed a strong desire to attain a "better" evaluation, as well as more objective and specific information on service effectiveness. The researcher felt that perhaps the funding organization, U.C.S., could take the lead in this respect by conducting research into agency evaluation to provide some data base from which some ideas for improving service evaluation could be developed.

The next observation arose from the researcher's reading. He found that there was a lack of empirical research on evaluative practices of social agencies, though there was much written about evaluation.

The researcher was further motivated by the fact that service evaluation has increasingly become a concern for the social work profession. Besides, the knowledge resulting from the research project would be useful to him in the kind of job he anticipates to take on his return to his home country, Singapore.

It was these observations and challenges which led the researcher to formulate this particular research project.

The literature was explored to gain understanding of the various concepts of evaluation research which helped formulate the study of the evaluative practices of the member agencies of the U.C.S. An interview schedule was
designed to collect data on various aspects of the evaluative practices of these agencies. The researcher believes that the findings obtained from this research should be useful to the U.C.S. in promoting/service evaluation, to any social planning groups, to the agencies themselves, and to the researcher. It was hoped that, based on the findings, some types of measures or strategies would be recommended to the U.C.S. and its member agencies for improving their service evaluations.

The research focus, data collection, samples and limitations of the study are further discussed in greater detail in Chapter III, Research Design and Methodology, whereas particulars of the survey of the literature are discussed in Chapter II. Chapter IV is devoted to analysis of data followed by recommendations and suggestions in Chapter V. In Chapter VI the project is summarized.
CHAPTER II

REVIEW OF THE LITERATURE

In the literature review, the researcher has surveyed some of the standard literature, journals and books in the social service field including public health, social work, social planning, public administration, education and corrections, and in the field of evaluation research itself.

The term evaluation research in social services as used in this paper, was defined as using the scientific method in the evaluation of social programs or services offered by social agencies. It does not focus on the evaluation of intervention skill or practice skill of the line social worker or those handling the clients in the delivery of services, although the skill part is not totally excluded from the consideration of overall program performance and effectiveness.

The literature survey has been separated into three areas: functions and types of evaluation research, models and techniques of evaluation research and constraints on evaluation research. This categorization reflects the researcher's interest in the relationship between these various dimensions of investigation and understanding of evaluation research. A description of these areas follows.
Types and Functions of Evaluation Research

The functions of evaluation research, purposes and the reasons for it, and the various types as well as levels and approaches to evaluation research are discussed in this part. In addition, the various definitions of evaluation research are examined, because variations in definitions are greatly due to the manners in which the writers see the purposes, the utility, and the focus of evaluation. Further, the types and functions of evaluation are explored in terms of historical perspective.

Defining Evaluation Research

Evaluation research like any analytic term, can be defined in many essentially arbitrary ways. Each of the ways which have gained common acceptance has certain utilities. The variations in definition are often due to the ways the writers see the purposes, utility and focus of the evaluation. For example, Bard distinguishes evaluation from monitoring. To him evaluation is the assessment of the extent of the program objective achievement after the program has been completed whereas monitoring is the assessment of program objective achievement while the programs are in operation. Riecken (1953) sees evaluation as "the measurement of desirable and undesirable consequences" of an action or sets

---

of operations or a program.\textsuperscript{20} Similarly, Greenberg and Mattison would restrict evaluation to the "follow-up" results.\textsuperscript{21} However, Klineberg defines evaluation as "a process which enables the administrator to describe the effects of his program, and thereby to make progressive adjustments in order to reach his goals more effectively."\textsuperscript{22} Tripodi et al also regard evaluation as a monitoring device for the feedback of program information to the appropriate persons at various stages of program development.\textsuperscript{23} Thus social program evaluation is defined as

...the systematic accumulation of facts for providing information about the achievement of program requisites and goals relative to efforts, effectiveness, and efficiency within any stage of program development. The facts of evaluation may be obtained through a variety of relatively systematic techniques, and they are incorporated into some designated system of values for making decisions about social programs.\textsuperscript{24}

As in the case of education, evaluation is first used for judging a product, for example student achievement, and

\textsuperscript{20} Henry Riecken, "Memorandum on Program Evaluation," in Evaluating Action Programs, p.86.


\textsuperscript{24} Ibid, p.12.
later comes to have utility also for judging a process, for example the means of instruction. The above definition of Tripodi et al (1971) also introduces concepts like (a) various foci of evaluation; efforts, effectiveness and efficiency; (b) data collection techniques; and (c) system and values for decision-making. Implicit in Klineberg's definition is that evaluation research extends beyond simply determining success or failure toward knowing why success or failure occurred and what can be done about it. This idea was later made very clear by Suchman (1969) when he commented that evaluation "also encompasses research on conditions affecting success or failure." In addition, Riecken's definition introduces the idea that social program may have negative consequences.

The aspects of analyzing, summarizing and reporting information in evaluation, in addition to collecting data, are stressed by Jenkins (1961) and Alkin (1969). Jenkins (1961) defines evaluation operationally as:

The process of acquiring, analyzing and using information for making decisions associated with planning, programming, implementing and recycling program components and activities. Its purposes are to

---

26 Edward Suchman, "Evaluating Educational Programs," in Readings in Evaluation Research, p.44.
facilitate continual improvement of an educational program and to assess the effectiveness of the program at its conclusion.27

Also explicitly clear in this definition is that evaluation is both a monitoring device feeding back information while the program is ongoing as well as a device assessing program effectiveness after the conclusion of the program.

To Alkin

evaluation is the process of ascertaining the decision areas of concern, selecting appropriate information, and collecting and analyzing information in order to report summary data useful to decision-makers in selecting among alternatives.28

From this definition, five areas of evaluation are identified: systems assessments, program planning, program implementation, program improvement and program certification. The first two and the last are somewhat similar respectively, to "context," "input" and "product" evaluations described by Stufflebeam.29

Similar to the various classifications by Tripodi et al (1970), Alkin (1969), and Stufflebeam (1969), are


Suchman's four categories of evaluation, namely, effort, effect, process and efficiency;\(^3^0\) and the four basic classes of evaluation criteria of the Planning-Programming-Budgeting System, namely, volume, performance, effectiveness and benefit.\(^3^1\) All these will be discussed later in a separate section.

The judgmental dimension of evaluation is stressed by Guttentag to whom "evaluation research...always involves a judgment of the worth of a program or project."\(^3^2\) It is also emphasized by Scriven (1967) who sees that evaluation must include "judgment of goals (which) are subjective value judgments not open to rational argument."\(^3^3\) Glass, in discussing the growth of evaluation methodology has stressed that evaluation is an attempt to assess worth or social utility.\(^3^4\)

The scientific aspect of evaluation is emphasized by Hyman and Wright when they specify the use of "methods


that yield evidence that is objective, systematic and comprehensive.\textsuperscript{35} It is also stressed by Suchman that "the need today is for more scientific evaluative research."\textsuperscript{36} Williams uses the term research to delineate "all studies using scientific methodologies."\textsuperscript{37}

However, Hayes discusses the art of evaluation other than the scientific part.\textsuperscript{38} Sherwood broadens the definition of evaluation to include strategies and skills other than those of social research.\textsuperscript{39}

All these variations in definition reveal that there is a mixture of conceptual and operational definitions and that evaluation is not a simple unitary concept. Despite the varieties, it seems clear that the following elements basic to research are common to many definitions:

(a) a pre-determined area for evaluation such as efforts, effectiveness;


\textsuperscript{38}Samuel Hayes, Evaluating Development Projects (Belgium: UNESCO, 1967).

(b) "scientific" method of investigation including data collection procedures, techniques, and instruments;

(c) analysis and interpretation of findings.

Reflected in the above discussion on definition is the fact that not many authors explicitly distinguish evaluation from evaluation research or evaluative research. These terms are used loosely and sometimes interchangeably. However, Suchman (1967) uses the term evaluation to denote "the general process of judging the worthwhileness of some activity regardless of the method employed," and evaluative research to mean "the specific use of the scientific method for the purpose of making an evaluation." Scriven (1967) uses the term estimation to refer to "over-crude evaluative procedures" as opposed to evaluation which embodies more scientific method or objective procedures. Weiss views evaluation research as "a specific method of evaluation-(in which) the tools of research are pressed into service to make the judging process more accurate and objective." What constitutes a scientific research or a scientific method? This question will be examined in the methodology section.

---

40 Suchman, Evaluative Research, p. 31.
For the purpose of this paper, the conceptual distinction made by Weiss (1972), Suchman (1967) and Scriven (1967) between scientific and less scientific ways of evaluating is adopted; that is evaluation which utilizes the tools and methodology which are normally accepted as scientific. No substantive distinction is made between "evaluative research" and "evaluation research" although the former means the stress is on research, that is, the type of research which has the purpose of evaluating something, whereas the latter means that the stress is on evaluation, that is, the evaluation of something by using research (scientific) methods.

Evaluation Research in Historical Perspective

As pointed out by Suchman (1967), as early as in the 17th century there were "scientific" evaluations of effectiveness in public health services. Morbidity and mortality data was used to measure the state of health as well as service effectiveness. Health surveys for the collection of data were popular through to the beginning of this century. Comparative ratings of communities by using a series of rating sheets were then developed in addition to mere statistical indices. The development of Community

43The information in this paragraph was condensed from Suchman, Evaluative Research, pp.13-9.
Appraisal Forms, Evaluation Schedules, Grading Standards, and Guides to a Community Self-Survey was at that time one of intense public service activity in the fields of health, education and welfare. Gradually, there was real demand for critical self-evaluation of the activities of local officials in order to increase standardization and to provide models for record-keeping. These self-rating appraisal forms, if judged by present standards of research, have many shortcomings, which eventually disqualify these attempts as being labelled "scientific" research. The major weaknesses in the area of method of procedure are: (a) samples are of biased or unknown representativeness; (b) there is the absence of "experimental" design; and (c) the measurements are of unknown accuracy and reliability.

According to Caro (1971), anticipation of formal social experimentation and evaluation research can be traced back to the writings of early social scientists in the 19th century. As he further pointed out, before World War II, there were some evaluative studies utilizing experimental method of the effects on participants in public health programs. During and immediately after World War II, a number of impressive evaluative research studies on group dynamics were contributed by social psychologists. In the

---

fifties, the number and the scope of evaluative studies were enlarged tremendously. For instance, inter-racial housing projects, volunteer work camps, delinquency prevention projects became subjects of evaluative studies.45

As indicated in the introduction, in the 1960's there was a great interest in evaluation research by the citizenry, the client public, the politicians and the social scientists. Evaluation research studies were conducted on a great number of federal and state sponsored projects and programs such as delinquent prevention projects, antipoverty programs and economic opportunity programs. Evaluation requirement in education in the United States, was written into an Act (1965),46 and hence evaluation has become an increasingly common legislative requirement. A committee of the National Academy of Sciences of the United States in 1969 recommended built-in evaluation and the allotment of five to ten per cent of program funds for evaluation.47

Undoubtedly, at present, the potential contribution of evaluation research is recognized in a great number of fields. However, as Caro (1971) puts it, the fundamental

concepts and methods of evaluation research have been widely
diffused. As reflected in the earlier discussion, there
is the wide range of variation in what is labelled "evaluation"
and there is the lack of any clear-cut definition of either
the objectives or procedures of evaluation research. The
possibilities and the extent of actual use of scientific
method in the evaluation of social services are still in
debate and discussion especially in face of political and
administrative considerations. Moreover, evaluation research
studies are often not made use of; are often left unpublished
or published in widely scattered journals. Besides, most
new social programs do not include a plan for evaluation.
It is fair therefore to say that the potential productive
contribution of evaluation research to social reforms, to
programs of directed social change, has yet to be realized.

Need and Functions of
Evaluation Research

Presently, all social institutions or sub-systems
whether social development, religious, economic or political,
have been increasingly asked or required to provide "proof"
of their legitimacy and effectiveness in order to justify
society's continued support. The funding sources, profes-
sional groups, clientele, political leadership and a more

---

sophisticated general public are no longer satisfied with demonstrations of the needs to which programs are addressed; they also want evidence of the contributions that these programs make in meeting these needs. Besides, questions such as those relating to the management of funds and the relative costs and efficiency of alternative programs are constantly being raised. In addition, there are increasing pressures for "proof" of effectiveness or ineffectiveness of various levels and methods of intervention. "The social welfare field and its associated professions must move from a fixed service delivery stance to one characterized by experimentation, model development and evaluation."\(^49\)

As indicated by Suchman (1967), the current desire to judge the worthwhileness of social programs and services is but one aspect of modern society's belief that many of its social problems can be met most effectively "through planned action based upon existing knowledge, including the design of even better solutions in step with advancing knowledge."\(^50\) This commitment to planned social change is very clear on the national and international scene. It is to be found in current approaches to the political, economic and social problems of the advanced societies and in attempts to


\(^{50}\)Suchman, Evaluative Research, p.2.
change the structure and functioning of the developing countries.

Underlying the current increasing demand for evaluation of social services and programs are changes in the nature of social problems, in the structure and functioning of the various fields of service, and in the needs and expectations of the public. Fundamental changes taking place in each of these areas constitute the basic need for evaluation research.

To the researcher, the continuous advances in human knowledge and technology result in continuous changes in social and physical environments such as living conditions, communication systems and social organization. This process of continuous change naturally gives rise to new problems, needs and values which in turn stimulate further search for new knowledge and technology. As new problems and needs emerge, new objectives and programs will have to be formulated and developed. Evaluation research is essential in this process to identify new needs and problems, modify services and approaches, and even to predict and plan for change. As a result of the changing social problems, Suchman (1967) observed that the dividing line between prevention and treatment is becoming increasingly difficult to determine. Social institutions which, rather than

\[51\text{Ibid, p.3.}\]
individuals, bear the major responsibility in attacking social problems and planning for change, have to modify and adjust their organization and operations, as their functions and programs change. This unchanging need for change in program and organization provides pressure for professional growth - better training, better knowledge of effects, higher degree of specialization and precision. Consequently, there is need and increased demand for evaluation research which provides the necessary feedback and proof of success or otherwise, thus ensuring continued refinement and development of knowledge and skills. Furthermore, as Suchman (1967) points out, a better educated and more sophisticated public is less willing to accept things on faith alone. They ask for scientific evidence of accomplishments. Coupled with this is an increased public expectation of bigger and better services. Thus the social problem, the service agencies, the professionals and the public reinforce each other to produce growing demand for evaluation research.

The above explanation suggests that there is an intrinsic relationship between evaluation research and social planning including program planning and development, and professional growth.

As observed by Johnson (1970), the most characteristic function of evaluation is to establish merit, and to

\[52\text{Ibid, p.4.}\]
assure that the information obtained is to serve as basis for action of some kind.\textsuperscript{53}

To Tripodi \textit{et al} (1971), evaluation research feeds back information on program efforts, effectiveness and efficiency at each of the three "sequential and interrelated stages" of program development: program initiation, program contact and program implementation.\textsuperscript{54} Changes or modification of programs thus depend on, or are effected, to a great extent by this feedback, taking for granted, of course, the program operators are receptive and value evaluation results. Suchman (1969) looks beyond the program itself to include external factors. He strongly maintains that evaluation study should do more than judging "pass" or "fail" on a program; it should attempt to find out why a program was or was not effective. The answer to this question "why" requires an analysis of such factors as (1) the attributes of the program that make it more or less successful; (2) the population exposed to the program; (3) the situations such as auspices, locale, competing programs and public opinion; (4) the different kinds of effects produced, such as attitudinal, cognitive, or behavioral including any negative side-effects.\textsuperscript{55} Thus, evaluation also encompasses

\begin{itemize}
\item \textsuperscript{53}Johnson, "Purpose of Evaluation," p.4.
\item \textsuperscript{54}Tripodi, Fellin and Epstein, \textit{Social Program Evaluation}, Chapter 1.
\item \textsuperscript{55}Suchman, "Evaluating Educational Programs," p.44.
\end{itemize}
research on conditions affecting success or failure. The major function of evaluation research is then to aid the administrators or program operators to plan and adjust their activities in an attempt to increase the probability of achieving the desired action or service goals. To indicate the role of evaluation, Suchman (1970) identifies a sequence of developments which relates to four major aspects of programmatic activity: Research - Planning - Demonstration - Operation. He visualizes a constant interaction among them with success at one stage moving the process forward to the next, but with failure at any stage leading to a recycling to some earlier stage.56

This process model as Suchman (1970) points out, has been applied by Guba and Clark to educational programs in terms of the following series of steps: research leads to two stages of development, invention and design, which is followed by two stages of diffusion, dissemination and demonstration, which in turn leads to three steps of adoption, trial, installation and institutionalization.57

Similarly, Caro (1971) envisions a planning-action-evaluation cycle which "may be repeated indefinitely until objectives are realized or until problems and objectives are redefined."58 Evaluation research then is one aspect


57 Ibid, p.57.

of a process of planned change, for its results can be used to increase the likelihood of realization of long-term goals and for redefining problems and objectives. But, as pointed out, unless its findings are implemented, it does not bring about change. 59

Seeing from another angle, Carter (1971) in her discussion of accountability holds the view that evaluation is potentially one of the change agents in shaping services from within the public welfare organizations. She makes implicit the evaluation function as means of helping the public welfare agency to face change and to formulate clear objectives for a public social service system available to all people. She indicates that the willingness of public welfare agencies to have their programs of direct service evaluated could well provide the key to whether they can be the instrument for social change in a reformed welfare system. 60

Kahn (1969) recognizes the contributions of evaluation research to social policy saying that it "may produce major needed review of policy and programs." 61 He (1973)


further considers that the effectiveness evaluation provides social indicators essential to describe or measure the state of affairs of different aspects of society, such as health, education, and corrections.\textsuperscript{62} This function of evaluation research as providing indicators of the state of affairs is further supported by Carlisle (1972).\textsuperscript{63}

Brooks' descriptions of the following four functions of evaluation research in action program are rather comprehensive:

(a) the accounting function - to inform the finding agent as to the value being received for dollars spent;
(b) the feedback function - to refine and improve the program being evaluated, through a continuous feedback of its results to the planning process (planning-action-evaluation-planning---etc.);
(c) the dissemination function - to make available to other interested communities the results of the program being evaluated;
(d) the theory-building function - to clarify, validate, disprove, modify or otherwise affect the body of theory from which the hypothesis underlying the program were derived.\textsuperscript{64}

To state it in the simplest way, the function of evaluation research is to contribute to rational decision-making.


making. With objective information on the outcomes of programs, actions or techniques, wise decisions can be made on budget allocations, policy, program planning, program adjustment and techniques improvement.

Purposes and Reasons for Evaluative Research

The purpose of the evaluation study determines the kind and amount of information the study might produce, the type and level of evaluation, and has in turn some bearing on the methods of measurement. Therefore, it is necessary to be quite clear at the very outset about the purpose of any evaluation study.

For example, in the study to determine effectiveness of social casework, the stated aim is to receive sound and validated information for the enlightenment and improvement of the profession. This calls for an evaluation of casework as a practice. If a voluntary agency asks whether the social services available are adequate to meet the needs of children in the community, the stated aim then is to secure a basis for deciding whether to continue present services, to expand them, or to inaugurate a new specialized service. It calls for evaluation of the quality and quantity of services for children, against the need for such services. Similarly, assessing inputs and effectiveness of the same program needs different techniques and methods of measurements.
The different purposes, stated or explicit, for conducting evaluation study can be partially derived from discussion on the functions. The evaluator or program director may want to analyse the features of the program, weigh effects of alternative methods or strategies. The legislators or funding agency may want overall assessment of program success, or of cost-benefit, or measurement of the effects of a program against the goals, et cetera. However, in practice, evaluation is sometimes undertaken for less noble motives or latent reasons. Program decision-makers may turn to evaluation to delay a decision; to justify and legitimate a decision already made; to extricate themselves from controversy about future directions by "passing the buck;" to vindicate the program in the eyes of its constituents, its founders, or the public; to satisfy conditions of a government or foundation grant through the ritual of evaluation. Suchman (1970) provides a brief catalogue of these "eyewash-whitewash" kinds of motivations. In an "eye-wash" evaluation, an attempt is made to justify a weak program by selecting for evaluation only those aspects that look good on the surface. A "white-wash" attempts to cover up program failure by avoiding any objective appraisal.

65 Weiss, Evaluation Research, Chapter 2.
66 Suchman, "Action for What?"
Knutson (1961) has categorized the explicit and overt reasons for the evaluation of a program of program activity into two groups, the organizational oriented and the personally oriented. From the point of view of the organization, a program evaluation may be undertaken to:

1. demonstrate to others that the program is worthwhile;
2. determine whether or not a program is moving in the right direction;
3. determine whether the needs for which the program is designed are being satisfied;
4. justify past or projected expenditures;
5. determine the costs of a program in terms of money or human effort;
6. obtain evidence that may be helpful in demonstrating to others what is already believed to be true regarding the effectiveness of a program;
7. gain support for program expansion;
8. compare different types of programs in terms of their relative effect;
9. compare different program methods or approaches in terms of effect;
10. satisfy someone who has demanded evidence of effect.67

The personal reasons underlying evaluation studies may originate with any, some, or all of the key people in the program, that is, the chief administrator, project director, or evaluator himself. Any of these persons may have many personal reasons for wanting to evaluate. For example, he may:

- see evaluation as "the thing to do" if one wants to belong;
- wish to make an evaluative study as a means of bringing favorable attention, better budgets, and better staff to his unit;

-see it as a means of gaining status and acceptance of peers and superiors,
-see it as a way of making his job easier and more interesting,
-see it as a step toward promotion,
-simply have a vague but urgent need to know if he is progressing. 68

It is easy to see that a wide range of personal as well as group reasons may underlie the desire to seek evidence of program success or failure.

Herzog claims that "organizational evaluations are governed by three main factors: survival, maintenance and knowledge." 69 That organizations are dedicated to survival and maintenance is largely responsible for the "success cult" in evaluation studies. As indicated by Sherwood (1967) the success cult is geared to find ways of convincing others of a program's accomplishments, and relies on the results of evaluation to "pass" the program. When this is the case, important goals are obscured. 70

The success cult or need for generating enthusiasm leads Elinson (1967) to give the following sarcastic advice:

When one wishes to show that a program has been successful, the evaluation should include the following:

1. A control group should not be used.
2. If a control group is used, it should be selected

68 Ibid, p.44.
purposively, rather than by random methods, and matched on "relevant variables".
3. If a prospective design cannot be used, the matching for the control groups should be done as criteria of success rather than effect variables.
4. The evaluation should be done preferably by those in charge of the program.
5. The results should not be published in the scientific literature, but should be issued as a report to one's self or the program director. 71

It is impossible to identify all the reasons involved in the decision to evaluate any particular activity, yet it is important to identify the reasons with as much precision as possible, because many aspects of the evaluation hinge upon the individual and group motivations for undertaking evaluation. As Knutson (1961) points out the reasons for evaluation:

(a) may govern the kind of program selected to be evaluated;
(b) will tend to govern the kinds of assumptions the evaluator is willing to accept or reject, thus helping to identify the action level for evaluation;
(c) help to govern the intensity of investigation to be undertaken and the level of critical analysis to be completed. 72

The evaluator is strongly urged by Weiss (1972) to spend enough time investigating who wants to know what, and why, to find out whether his study will serve a genuine purpose or whether it is likely to be ignored or used for political ends. 73

---

72 Knutson, "Evaluation for What?" pp.64-6.
Most of the decisions that need to be made in planning and conducting an evaluation research depend very much on the reasons or purposes for it. Just like in any research, major decisions regarding the form, instruments, methods, and sample are made around the hypothesis or statement of problem and level of proof desired. In spite of its importance, as Carter and Wharf (1973) observe, spelling out the "why" of evaluation is a part of the design or planning that is frequently neglected.⁷⁴

Types of Evaluation Research

The purposes and reasons for evaluation have great bearing on the approach to, and level of evaluation which in turn influence the considerations for choices of model and techniques.

Hutchison (1960) has drawn attention to the distinction between two types of evaluation. He distinguishes "evaluation of intermediate objectives," that is, the technique, from "evaluation of ultimate objectives" that is, the benefits derived from the procedure. As explained, the evaluation of technique does not require the identification of cause-effect relationship, but merely the definition of technique and the observation, recording and

⁷⁴ Carter and Wharf, Evaluating Development Programs, p.15.
cataloguing of events of which it consists.\textsuperscript{75} This distinction is similar to Scriven's "intrinsic" evaluation versus "pay-off" evaluation; the former involves an appraisal of the instrument itself (such as in education: the content, goals, grading procedures, teachers' attitude), whereas the latter involves an examination of the effects the teaching instrument has on the pupil.\textsuperscript{76}

Among formal approaches to evaluation, Caro (1971) makes a distinction between those emphasizing inputs and those emphasizing outputs. Examples of input approach are: (a) educational programs are evaluated on the basis of such factors as teacher qualifications and ratios of library books to students; and (b) criteria like plumbing facilities and sleeping arrangements are used to determine the adequacy of housing. Program accounting which focuses on maintenance and quantitative analysis of records of project activities is one procedure to study input or effort. Such a procedure may provide a sound basis for screening programs on the basis of ability to establish contacts with clients and the cost of program-client contacts. Those emphasizing output or effects use scientific methods (i.e. experimental design) for collecting data concerning the degree to which

\textsuperscript{75}George B. Hutchison, "Principles in the Evaluation of Community Mental Health Programs," in Program Evaluation in the Health Fields, p. 52.

\textsuperscript{76}Scriven, "Methodology of Evaluation," p. 128.
some specified activity achieves some desired effect. Scriven (1967) suggests "mediated" evaluation as a way of combining input and output variables so that the process through which goals are pursued can be studied. Scriven (1967) also introduces the terms "formative" and "summative" to distinguish evaluation research which is designed to improve a program while it is still fluid from evaluation research designed to appraise a product after it is well established.

The concepts of these two approaches, formative and summative, have been further refined and elaborated on by many authors, such as Planagan (1969), Provos (1969), Johnson (1970), Lumsdaine (1970), Carter and Wharf (1973). In commenting on the use of these two terms by Johnson, Lumsdaine (1970) clarifies:

The term summative evaluation has two implications. These often go hand-in-hand, but not always. First, summative is distinguished from formative in the sense that summative is a terminal evaluation of a finished product. Second summative is an overall evaluation that relates to those purposes of evaluation that involve making an overall decision: Shall we buy the program or not? Shall we go on with it or not? It is summative in that it has an overall decisive function one way or another.

The notion of formative evaluation also has two aspects: (1) use of evaluative data in the course

of developing the program rather than only at the termination of development; and (2) a fine-grain set of decisions relating to numerous specific aspects of a program, rather than just a single overall figure of merit.

The formative/summative distinction should perhaps be regarded as a sort of continuum.\textsuperscript{79}

The summative evaluation is particularly useful for providing a basis for making decisions about alternative programs. The formative evaluation appears most suited for application to development programs and is being used increasingly in areas such as community action.\textsuperscript{80}

Cain and Hollister (1969) have the similar concept in approach to evaluation; they recognize two broad types of evaluation, process evaluation and outcome evaluation. But, there is a different emphasis in their conceptualization of these two types of evaluation. To them, process evaluation deals with integrity of its financial transactions, accounting system, record keeping and other managerial functions; it addresses the question: is the program being run honestly and efficiently? Their outcome evaluation is

\textsuperscript{79} Arthur Lumsdaine's comments that accompany George Johnson's article, "Purpose of evaluation," pp.18-9.

\textsuperscript{80} Carter and Wharf, \textit{Evaluating Development Programs}, p.34.
more appropriately described as "cost-benefit analysis" which has, to them, to be differentiated from "cost-effectiveness" study.\textsuperscript{81}

According to implications for conduct of evaluation research, Suchman (1970) classified evaluation studies into the following types:

1. System-oriented evaluations
   (a) Institutional or social sub-systems (i.e. health, education, or welfare systems).
   (b) Complex organizations (i.e. hospitals, public schools, welfare agencies).
2. Program-oriented evaluations
   (a) Simple organizations (i.e. clinics, classrooms, employment offices).
   (b) Simple programs (i.e. immunization program, reading improvement, enrollment campaign).\textsuperscript{82}

Evans (1969) introduces the three-fold scheme of evaluation of the Office of Economic Opportunity. It is described that all evaluations can be categorized into this scheme:

Type 1. The assessment of overall program impact and effectiveness where the emphasis is on determining the extent to which programs are successful in achieving basic objectives.

Type 11. The evaluation of the relative effectiveness of different program strategies and variables where the emphasis is on determining which alternative techniques for carrying out a program are most productive (for example,


\textsuperscript{82}Suchman, "Action for What?" p.76.
determining whether one type of curriculum is more effective than another for teaching Head Start children).

Type III. The evaluation of individual projects through site visits and other monitoring activities where the emphasis is on assessing managerial and operational efficiency.83

Evaluation can be performed at different levels of achieving the program's objectives. As indicated earlier, Caro (1971), Scriven (1967), Johnson (1970), Cain and Hollister (1969), et cetera, conceive two major approaches or levels at which evaluation is desired: process and outcome. James (1962) has suggested to classify "a variety of levels" into four: effort, performance, adequacy of performance, and efficiency.84 In the Planning-Programming-Budgeting System, the "four basic classes of evaluation criteria" are volume, performance, effectiveness, and benefit.85 Tripodi et al (1971) suggest three dimensions of evaluation: efforts, effectiveness and efficiency.86

There are similarities and differences in the conceptualization by different authors of these terms. The researcher tends to classify the levels in achieving the program's objective into the following: efforts, effect-


86Tripodi, Fellin and Epstein, Social Program Evaluation, Chapter 1.
iveness, efficiency and benefit.

Program Efforts refers to the type and quantity of program activities, in other words, quantity of goods and services. Its evaluation means an assessment of the amounts and kinds of program activities considered necessary for the accomplishment of program goals, including staff time, activity, and allocation and use of resources such as funds, space, equipment, outside consultation, public relations, *et cetera*. It is difficult to agree to Carter and Wharf’s observation that results of the effort evaluation are “often highly subjective;” and that this kind of assessment “appears to be extremely limited and generally lacking in utility.”

The results can be objectively arrived at, and the description of results can certainly be quantified. Even the intensity of work of a person can also be measured though it is more difficult. Information about program effort may be extremely useful to administrators. For example, an analysis of staff activity may reveal that too much effort has been spent on, instructional duties for example, to the neglect of other aspects; or that in spite of a high proportion of staff time and effort spent on instructional activities, the effectiveness (from effectiveness evaluation) may not be there. The findings on effort then may have implications for intervention techniques, quality or

---

training of staff, et cetera. Thus, evaluation of program effort is important for the achievement of program goals, but evidence of program activity or effort is not sufficient to determine whether or not these goals have been reached.

Program effectiveness is concerned with achievement in terms of contribution to the program objectives. Levine (1967) sees this type of evaluation as answering questions such as: is the program achieving its desired objectives? And to what extent are these goals being realized? But as Rossi (1972) observes, program goals are usually vague. Therefore to make evaluation of effectiveness possible, the following two conditions are of central importance: (a) identification and operational definition of the objectives of the program in a way that is measurable, and (b) the selection of the appropriate measures (sometimes each objective requires a different measure of effectiveness).

Program efficiency evaluation refers to the assessment of quality of work performed and the relative costs of achieving the objectives: Efficiency focuses on the relationship between effort and effectiveness; it is the

---


ratio of effectiveness to cost, and is concerned with the relative costs for achieving objectives. Evaluation of efficiency often involves comparison of two or more different techniques, strategies or programs, with respect to the relative costs; comparison with results in previous years, or with results achieved by others in similar situations.

This type of evaluation is essentially addressed to the question: Can the same end be achieved at lower cost and how? Its results may have implications for techniques and strategies, use of staff and resources, program adjustments, designs and alternatives.

Program benefit evaluation refers to measures of program accomplishment in monetary terms. Being in monetary terms direct benefits accruing from the fulfillment of program objectives "can be added with the indirect or 'spillover' benefits to obtain an expression of total benefits." Rothenberg as cited by Rossi (1972), suggests that it should be possible to rely heavily on monetary values. For example, the benefits accruing from subsidized higher education may be converted into increased lifetime income returns to individuals, and into terms of the productivity of the labor force resulting from a higher level of

---

90 Tripodi, Fellin and Epstein, Social Program Evaluation, p.49.

91 Province of Ontario, PPBS, p.27.

education. Through comparing total benefits (direct and indirect benefits) with total costs, a cost-benefit ratio can be established. This reduction to monetary terms is considered by the Planning-Programming-Budgeting System as one of the principal advantages of benefit measures over effectiveness measures.\textsuperscript{93} To Rothenberg and Levine, as cited by Rossi (1972), the results of benefit evaluation also provide a very useful basis for rational decision making.\textsuperscript{94} Benefit evaluation has been common to the military and the industry, and is being adopted slowly in social programs.

Models and Techniques of Evaluation Research

Included in this part are the differentiation between evaluative and non-evaluative research, various models of evaluation research, and some of the techniques used in evaluation research.

Evaluative and Non-Evaluative Research

To the researcher, social research whether evaluative or non-evaluative is a scientific undertaking. It is concerned with both scientific method which is a logical

\textsuperscript{93}Province of Ontario, PPBS, pp.27-8.

\textsuperscript{94}Rossi, "Testing Success and Failure in Social Action," pp.23-5.
and systematized application of the fundamentals of science to the general questions of study, and scientific techniques which provide precise tools, specific procedures and technical rather than philosophical means for collecting and ordering the data. Experimental design, sampling, statistical rules common in social research, are all derived from the fundamentals of science and are thus considered scientific. As asserted by Ogburn, without the scientific criteria, many research projects would become "merely effort and good intentions." Although, as Ogburn further points out, social science research cannot be as exact and "ideal" as physical science research, it still demands a high degree of objectivity, accuracy of approximation that will satisfy scientific requirements. It is very different from general scholarly intellectual work. It has to follow definite procedures and must meet certain stipulated conditions.

All these scientific principles and logic are aimed at getting information that is valid, reliable and precise. Reliability varies with precision. There are many sources of invalidity and unreliability in the instruments, techniques and designs used. For example, Campbell and Stanley identify 12 sources of invalidity in experimental

and quasi-experimental research designs. Further, the primary source of data collection is done through human sense perceptions, and psychologists have shown that the accuracy and reliability of human perception can be notoriously difficult to evaluate. However, it should be pointed out here that the awareness of knowledge of these possible sources of invalidity and reliability which are essential to know, should not create incomprehended fears or confidence in research. And, the reader should maintain this perspective for the discussion on methodology of evaluation research.

To reiterate, it is the degree of accuracy desired, the level of proof aimed at, the state of existing knowledge of the research topic, the reliability of the instrument and tool, the rigidity of the procedure, which all combine to determine the amount of concern one can have with the degree of "science" in one's design. But, as has been emphasized elsewhere, "science offers no possibilities of absolute proof."  

Selltiz et al (1951) using research purposes and

---


probable level of proof desired as criteria, classify research studies into three categories: formative or exploratory, descriptive, and studies testing causal hypothesis. Kahn (1960) using the state of available knowledge and the type of design as criteria identifies four types of research studies: random observation, exploratory or formulative, diagnostic and descriptive, and experimental. Tripodi et al (1969) present a classification system which is based on the major purposes of research and the empirical methods employed. They classify research studies into three major types: the experimental, with two sub-types, the quantitative-descriptive with four sub-types, and the exploratory with three sub-types.

Hyman (1955) distinguishes three types of research studies: (a) the theoretical or experimental, (b) the evaluative or programmatic, and (c) the diagnostic. The theoretical study emphasizes the testing of specific hypo-

---


thesis relevant to some larger body of theory, while the evaluative study is designed to test the practical value of some action program; the diagnostic study, similar to the exploratory type conceived of by Kahn, Selltiz et al and Tripodi et al, explores some unknown, novel problem. A similar distinction is made by Zetterberg, (1963), who differentiates between diagnosis, as leading to descriptive studies, and explanation, which requires verificational studies the purpose of which is to test hypotheses.102

With the rise in interest in social intervention programs and action programs in recent years, terms like "action research," and "pure" and "applied" research have been in much discussion. For example, the Zuckerman Committee (1961) has made a distinction between "pure basic" and "objective basic" research.103 Chern, in an attempt to provide a "more useful categorization of research in the social sciences," suggests four types of research - pure basic, basic objective, operational, and action.104 This categorization is based on the functions, and the nature of the topic for research. Also, this categorization,


103 As cited by Albert Chern in his article, "Social Research and Its Diffusion," in Readings in Evaluation Research, p.64.

104 Chern, "Social Research and Its Diffusion," p.64.
especially the operational research and action research, has a close relationship with evaluation research of social programs, and therefore, the distinctions made by Chern are summarized by the researcher as follows:¹⁰⁵

1. **Pure Basic Research** is aimed at illuminating a theoretical problem.

2. **Basic Objective Research** is directed towards a problem arising from theoretical application but is not aimed at providing a solution to a practical problem.

3. **Operational Research** is oriented towards tackling an ongoing problem within an organization, but it does not include experimental action. Rather, it identifies the goals of the organization, establishes criteria of goal attainment and measurement of performance.

4. **Action Research** is usually inclusive of some operational research as part of its strategy, but with an emphasis on a systems approach. It emphasizes planned change or some social improvement. Therefore, it is oriented towards practical problems, and must rely on empirical science.

   The significant difference between evaluative or non-evaluative research "is one of purpose and not of method."¹⁰⁶ In other words, the basic methods, principles, logic in terms of procedure, design, measurement and

---

¹⁰⁵ *Ibid.,* pp.64-5.

analysis that apply to all other types of research apply to evaluation research. "What distinguishes evaluation research is not method or subject matter, but intent - the purpose for which it is done."\textsuperscript{107} Weiss (1972) however, outlines some of the differences and similarities between evaluation research and other types of social research. The similarities are mainly related to methods, procedure, data collection instruments and techniques, and research design. The differences which are, to the researcher, worthy of note, are in the following dimensions: (a) use for decision-making; (b) program-derived questions; (c) judgemental quality; (d) action setting; (e) role conflicts; (f) publication; and (g) allegiance.\textsuperscript{108}

Models

Closely related to the types of evaluation research and scientific methods of research are various evaluative models and techniques. The various models and techniques have been grouped in different ways by different authors reviewed for this study.

As to models, which is defined by Wallin as "integrated conceptualizations of purposes and means,"\textsuperscript{109} Etzioni (1960), Sherwood (1964), Levinson (1966), Schulberg and

\textsuperscript{107} Weiss, Evaluation Research, p.6.
\textsuperscript{108} Ibid, pp.6-8.
Baker (1969), Carter and Wharf (1973), all identify two basic evaluation models: goal or goal-attainment and system. Houston (1969) adds an impact-effectiveness model, which is similar to Carter's (1971) impact evaluation. Wallin (1972) synthesizes Chern's operational research (1969) and Carter's operations research (1971) into an operational (process-oriented) model. To these four models, the researcher of this project would like to add the cost-benefit model which is consistent with his program benefit evaluation proposal. Use of cost-benefit analysis in social programs has been strongly advocated by Levine (1966), Cain and Hollister (1969), and Planning-Programming-Budgeting System (1970), Rossi (1969) and Rothenberg as cited by Rossi (1969).

These models have been adequately discussed in references cited here, and the discussion of various types and approaches to evaluation research in the earlier section also gives a glimpse of these models. The researcher therefore, does not propose to discuss them. Instead, he summarizes the common characteristics and purposes of these models in Chart I as follows.

Of these five the goal model and the system model have a higher frequency of usage in evaluation.
**CHART I**

**MODELS OF EVALUATION RESEARCH**

<table>
<thead>
<tr>
<th>Models</th>
<th>Purposes</th>
<th>Characteristics</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Goal-Attainment</td>
<td>- to evaluate to what extent a specific goal is achieved</td>
<td>- concerned with goal setting measures of goals</td>
<td>Experiments: classical, quasi or approximations to experiments.</td>
</tr>
<tr>
<td>Model</td>
<td></td>
<td>- measuring program stimuli (the independent variable), the changes (the intermediate variables) and the outcome (the dependent variable).</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- particularly useful for evaluating change</td>
<td></td>
</tr>
<tr>
<td>Systems Model</td>
<td>- to evaluate the degree to which organizational goals are reached under a set condition.</td>
<td>- being criticized for being too narrow in its evaluation methodology.</td>
<td>Experiments and monitoring techniques</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- alternative programs and goals and their relationship are considered</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- sub-goals are considered as part of a hierarchy of goals and the failure of one can be compensated by the success of another</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- goal attainment model used for evaluating a specific</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Models</th>
<th>Purposes</th>
<th>Characteristics</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Impact Model</td>
<td>- to determine the impact of intervention, action or a program activity.</td>
<td>- goal achievement is often used as part of the system model.</td>
<td>- Experiments</td>
</tr>
<tr>
<td></td>
<td>- to determine what is going to be done to change things; was the intervention linked to the change; and what causes the changed behaviour or condition.</td>
<td>- evaluation of the multiple goals becomes complex; some sub-goals may be overwhelmingly realized, but the main goal may not be achieved.</td>
<td></td>
</tr>
<tr>
<td>Operational Model</td>
<td>- to assess performance against criteria of goal attainment</td>
<td>- concerned techniques and methodologies of intervention.</td>
<td>- Common social research methods</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- particularly useful to evaluate professional performance.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- a good approach for dogmatically justified programs.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- process-oriented; on-going.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- feedback the findings to the program administrators</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Continued)</td>
<td></td>
</tr>
<tr>
<td>Models</td>
<td>Purposes</td>
<td>Characteristics</td>
<td>Methods</td>
</tr>
<tr>
<td>---------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------------</td>
<td>----------------------------------------------</td>
</tr>
</tbody>
</table>
| Cost-Benefit Model  | - to find out congruence or discrepancy between goal and achievement.    | - for immediate program modifications  
|                     | - provides a useful framework for rational decision-making.              | - permanent part of the system; built in an organization.                        | - cost-benefit analysis (it is possible to avoid using the comprehensive empirical research project). |
|                     |                                                                          | - expressed in monetary terms  
|                     |                                                                          | - provides cost-benefit ratios  
|                     |                                                                          | - not popular in social programs, which some argue, have their main justifications not in monetary terms but in terms of human and social values.  
|                     |                                                                          | - conducted primarily on assessment of employment training programs.            |                                              |

*This chart was developed by the researcher.*
The goal model directs its evaluation efforts at measuring how well a specific organizational goal is achieved. The systems model contends that such an approach is unproductive and even misleading, since an organization constantly functions at a variety of levels. Even though directing part of its means directly to the goal activities, an organization must simultaneously devote segments of its resources to such other vital functions as maintenance and recruitment.¹¹⁰ Etzioni (1960), the foremost proponent of the systems approach, argues that the starting point for a study of program effectiveness should not be the specific goal or objective but rather a working model of social unit which is capable of achieving a goal.¹¹¹ Using the systems approach, "the evaluation of the multiple goals becomes more complex but in many cases more real...and by using the systems model approach, some degree of realistic comprehensive assessment can be achieved."

However, provocative as the notion is, the systems model sets such demanding requirements for the evaluator (knowing more about the organization than the organization knows itself) that it is difficult to imagine its practical application....For the time


being, most evaluators will probably stick with the goal model...and give as much attention to the organizational and community systems that affect the program as the situation seems to warrant. 112

Techniques

As the approaches, types and models only provide the conceptual framework for doing an evaluation research, the evaluator must eventually select the appropriate techniques which can be used to provide information pertinent to the questions of evaluation. However, it is impossible here to describe and discuss each of the available techniques, their procedures, reliability, special skills required (like sampling, data analysis, cost accounting). What the researcher proposes to do is to present a summary of common uses and purposes of some selected techniques. It is the belief of the researcher that acquaintance with the methods and uses of various evaluation techniques should be of help to those who, though not experts, are interested in evaluation techniques. Some knowledge of various evaluation techniques should be useful to those who wish to seek consultation regarding the kind of evaluation strategy that is most appropriate for their program and to estimate the expense involved.

The summary to be presented is based on the scheme

of Tripodi et al (1971) who group the techniques into three categories.\textsuperscript{113}

1. Monitoring techniques, including procedures used for the direct review of program operations: accountability audit, administrative audit, and time-motion studies.

2. Social research techniques, referring to procedures which exclude cost considerations and which are used for developing, modifying and expanding knowledge about the program which can be communicated and verified by independent investigators: experiment, survey and case study.

3. Cost analytic techniques, referring to procedures used to appraise the relative value of a program in relation to program costs: cost-accounting, cost-benefit analysis, cost-outcome analysis, and operational research, which blends experimental and cost-analytic methods.

In general, they all have proven useful for certain types of evaluation. The brief description of these techniques and their uses as condensed by Carter and Wharf (1973)\textsuperscript{114} was adapted by the researcher in Chart 2 that follows. The title of the chart, Techniques of Evaluation Research, was given by the researcher.

Despite the variety of techniques and models, there are really no distinct tools or procedures which are uniquely assigned to evaluative research. The research processes of evaluation are similar for social research in general - defining the problem, or asking the right questions, developing a design which will utilize our knowledge base,

\textsuperscript{113} Tripodi, Fellin and Epstein, Social Program Evaluation, pp.61-111.

\textsuperscript{114} Carter and Wharf, Evaluating Development Programs, pp.26-8.
### CHART 2

**TECHNIQUES OF EVALUATION RESEARCH**

<table>
<thead>
<tr>
<th>Techniques</th>
<th>Methods</th>
<th>Purpose</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>MONITORING TECHNIQUES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accountability</td>
<td>- to review consistency, dependability and accuracy of records re program expenditures, allocations of resources and processing of program beneficiaries</td>
<td>- to establish program accountability</td>
<td>(1) General accounting</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(2) Social accounting</td>
</tr>
<tr>
<td>2. Administrative audit</td>
<td>- to evaluate the suitability of program policies and practices</td>
<td>- to provide data for planning to improve effectiveness and efficiency in administrative practices</td>
<td>(1) Simple or focused</td>
</tr>
<tr>
<td></td>
<td>- to evaluate adherence to staff practices to designated divisions of responsibility and function</td>
<td></td>
<td>(2) Complex or comprehensive</td>
</tr>
<tr>
<td></td>
<td>- to evaluate organizational patterns of work</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Time-Motion studies</td>
<td>- to specify the amounts of time devoted by staff to program activities</td>
<td>- to provide data for planning to improve effectiveness and efficiency of staff use</td>
<td>(1) Observation</td>
</tr>
<tr>
<td></td>
<td>- to locate uses of staff time which were not anticipated</td>
<td></td>
<td>(2) Self-reports</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(3) Sample time-motion study</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>using samples of both staff</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>and time</td>
</tr>
<tr>
<td>Techniques</td>
<td>Methods</td>
<td>Purpose</td>
<td>Types</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>SOCIAL RESEARCH TECHNIQUES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Experiments</td>
<td>-to recommend re-allocations of staff time to those activities which might be more directly related to the potential achievement of program goals.</td>
<td>-to provide evidence on whether or not program efforts are related causally to the accomplishment of program goals</td>
<td>numerous (a. Selltiz et al., 6 types; b. Campbell &amp; Stanley, 16 types; c. Chapin 3 types)</td>
</tr>
<tr>
<td></td>
<td>-classic experimental design, quasi-experiments or approximations to experiments</td>
<td></td>
<td>-flexible</td>
</tr>
<tr>
<td>2. Surveys</td>
<td>-setting of survey objectives -designation of a target population -selection of a representative sample from that population -collection of data -accuracy checks</td>
<td>-descriptive function to obtain accurate facts and statements of opinion representative of the target population -explanatory function to derive</td>
<td></td>
</tr>
<tr>
<td>Techniques</td>
<td>Methods</td>
<td>Purpose</td>
<td>Types</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>3. Case Studies</td>
<td>-analysis</td>
<td>possible causal connections or to describe relationship among variables</td>
<td>participant observation, informa interviews, methods of group analysis, content analysis of written documents, etc.</td>
</tr>
<tr>
<td></td>
<td>-accumulation of as much information as possible -development and use of conceptual scheme for generating ideas</td>
<td>-to provide detailed description of social programs as they develop for evaluation of program efforts -particularly useful for developing programs where there is difficulty in specifying objectives and in selecting programmatic means to accomplish these objectives -to pinpoint potential problems in program operation</td>
<td></td>
</tr>
<tr>
<td>Techniques</td>
<td>Methods</td>
<td>Purpose</td>
<td>Type</td>
</tr>
<tr>
<td>-------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
<td>---------------------------</td>
</tr>
<tr>
<td>COST ANALYTIC TECHNIQUES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Cost accounting</td>
<td>production of unit cost figures on a basis for analyzing, budgeting and allocating resources</td>
<td>to relate program costs to program outputs</td>
<td>(1) Cost accounting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to improve program budgeting</td>
<td>(2) Program budgeting</td>
</tr>
<tr>
<td></td>
<td></td>
<td>to provide information for determining program service priorities as a function of cost</td>
<td></td>
</tr>
<tr>
<td>2. Cost-benefit analysis</td>
<td>program inputs are related to program outputs or staff actions; program outputs are then related to the results of those actions</td>
<td>to evaluate the relative effectiveness of alternative programs, strategies, etc., in terms of cost</td>
<td></td>
</tr>
<tr>
<td></td>
<td>translation of criteria of goal achievement into monetary units</td>
<td>to ascertain the relationship of costs to benefits (specified goals)</td>
<td></td>
</tr>
<tr>
<td>3. Cost-outcome analysis</td>
<td>to relate program costs to the results of program activities</td>
<td>to gauge the relative efficiency of the costs of alternative program inputs</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Techniques</th>
<th>Methods</th>
<th>Purpose</th>
<th>Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Operations research</td>
<td>- combines scientific experimentation, mathematics, statistics, and computer technology - deliberate</td>
<td>for the accomplishment of specified objectives - to provide data on alternative ways of conducting and co-ordinating program activities within an organization</td>
<td></td>
</tr>
</tbody>
</table>

*(Adapted from Carter and Wharf, 1973, pp. 26-8)*
using survey, sampling, accounting principles, observation techniques, statistical procedures, analysis and the like. To give an example, Hawkridge, senior research scientist of the American Institutes for Research, after reviewing "the evaluations of many hundreds of programs, most of them federally funded," has come up with seven "commonly used phrases of evaluation" which are:

1. Setting up objectives
2. Selecting objectives
3. Choosing instruments and procedures
4. Selecting samples
5. Establishing measurements and observation schedules
6. Drawing conclusions and recommendations
7. Choosing analysis techniques. 115

They are not much different from those on social research proposed by any other social scientist, like Miller (1964), Selltiz et al (1951), Tripodi et al (1969), Kahn (1960), Ackoff (1953), Phillips (1967), Isaac and Michael (1971).

There is no one best method for evaluation as there is no one best method for research. Rather it becomes a matter of assessing the evaluation situation before determining if a particular model or technique or a combination of models or techniques best fits the task at hand.

However, it has to be recognized that there are some constraints and problems special to evaluation research which will be examined in the next part.

Problems and Constraints on Evaluation Research

The common problems and constraints of evaluation research that are being examined include those related to methodological and social-political and organizational context.

The Methodological Problems

Social services or programs of all sorts are expected to result in some sort of changes in the client, either his behaviour or attitude or state of well being. Therefore, measurement of the changes in the client is basic to service evaluation, despite the various models, approaches or levels of evaluation research. A measurement of client changes is necessary in the study of service effectiveness, efficiency and benefit; it is also essential in the goal-attainment model, system model, impact model or operation model.

The study of a client's changes requires a high degree of accuracy and reliability, as well as inferences about causality. In other words, the study requires evidences that can prove that there are changes in the client after receiving services and that the changes are due to the programs or services. It is rather commonly acknowledged that the technique or research design which represents the highest degree of proof and at the same time permits inference about causality is the experimental, or quasi-experimental design. In other words, an experimental or
quasi-experimental design is basic to evaluation research. For instance, Suchman (1967) highly commends the research model presented by Greenberg and Mattison for evaluation in terms of a graphic flow chart involving the definition of a target population, the drawing of a representative sample for study, the allocation of this sample at random into experimental and control groups, the administration of the program to one group and a placebo to the other, and the comparison of resulting differences between the two groups. This model actually represents the ideal experimental design.

Variations to the ideal experimental design are many. Selltiz et al (1951) identify six types of experimental designs. Campbell and Stanley (1966) give a very comprehensive classification of experimental design into three pre-experimental designs, three true experimental designs and ten quasi-experimental designs. Suchman (1967) though recognizing that "there is no one best way to design evaluative studies," identifies five variations to ideal experimental design in evaluative research; they are: (a) the one-shot case study; (b) the one group, pre-test, post-test design; (c) the static group comparison; (d) pre-test,
post-test control group design; and (e) the Solomon four-group design. ¹²⁰

As Suchman points out, the ideal evaluation study would follow the classical experimental model, but, as in the case for most areas of research, evaluative research projects, by and large, utilize some variation or adaption of this model. ¹²¹

Setting up equivalent experimental and control groups

In the use of experimental method to evaluative research, there are a few common problems which have to be recognized. The first concerns the setting up of equivalent experimental and control groups. Unlike the laboratory situation, there are many problems in obtaining such groups in service evaluation. For instance, in the case of community programs, it is difficult to find an equivalent community. Also, there is the ethical problem of withholding service from a group that needs it. As Blenkner points out, "No casework agency is so dedicated to science as to permit it to make a random sort of its applicants, offering help to one half, while merely following up the other half to see what happens to them." ¹²² The recent

¹²⁰Ibid., pp. 93-96.
¹²¹Ibid., p. 102.
cancellation of an experimental research program sponsored by the provincial intoxication recovery center in Alberta by the Provincial Health Minister, Neil Crawford, is a good indication of the ethical problem involved in setting up experimental and control groups. The Globe and Mail reports:

"Under the program, which has been operating for about two weeks, every 10th person who applied for admission to the centre and qualified was turned away. Such people were told that they did not qualify but that they could get $10 for allowing themselves to be interviewed in a week at the offices of the Alberta Alcohol and Drug Abuse Commission, which runs the centre.

Every ninth person who applied for admission and received it also was interviewed. Information from the two groups interviewed was then compared to determine how effective the centre is in treatment.

"I [the Minister] consider what happened to have been an error in judgment. I think it would be useful to get the data (which the program was trying to collect) but it can't be done that way. We'll just have to find another way."123

Truly, as Smith (1955) observes, "Practical difficulties of gaining access to genuinely comparable groups have made the use of this ideal [experimental] design a rarity."124

Formulation, isolation and manipulation of the independent variables

The second major methodological problem in evaluative research concerns the formulation, isolation and


manipulation of the variables which are intended to produce the desired effects. This second major problem involves two main sources of difficulty. One is to identify and specify the program dimensions, the independent variables in an evaluation.

As Hyman and Wright (1967) point out, a program is "a most deceptive term." There is always the discrepancy between program as plan and program as reality. A program as plan is simply a statement on paper of what the planners in an agency hoped to do that may never be fully translated into action by the field staff. Taking the word for the deed, caution Hyman and Wright (1967) an evaluator may try to observe the effects of a non-existent stimulus or treatment. A program as reality, to them, includes three elements: staff, site (including facilities) and treatment or activity. The treatment or activity element in most programs is "anything but a unitary variable;" it demands analysis and our powers of conceptualization. Further, a program is carried out over a certain period of time consisting of various stages of development. A program may be in cyclical operation. In this case, the very first cycle of a program may not yet function at a maximum

---

127 Ibid, p.197.
efficiency. In overlapping cycles, there may be input or influence which should also be taken into consideration. The personal styles and personality differences among the staff and the effects of sites are part and parcel of the independent variables. Therefore, specifying the program variables for evaluative purpose means a specification of the main elements of a program as reality. In other words, the evaluator must know how a program is carried out as well as what the program encompasses.

Weiss (1972) gives quite a comprehensive account of the elements with which the program variables may be associated: purpose, principles, methods, staffing, persons served; length of service, location, size of program, auspices and management.\(^{128}\) She further elaborates that characteristics of participants are also associated with input variables. To her, such participant measures can include: age, sex, socioeconomic status, race, length of residence in community, attitudes towards the program, motivations for participation, aspirations, expectations from the program, attitudes of other family members towards the program, and degree of support from family, friends, fellow workers, supervisors and so on for the intended goals of the program.\(^{129}\)

---

\(^{128}\) Weiss, Evaluation Research, p. 46.

\(^{129}\) Ibid, pp. 46-7.
A note of caution is necessary here. There are many variables that are interesting to study. But most evaluations have limited resources, and it is far more productive, as Weiss suggests, to focus on a few relevant variables than to go on a wide-ranging fishing expedition.\footnote{130}

The other main source of difficulty is related to isolation and manipulation of the independent variables, that is, the control of communication channels. For the evaluator, it is necessary to isolate insofar as possible, the service or treatment method that is being evaluated from all other services or influences. Without such isolation, it is impossible to tell whether change is due to the new treatment or other sources. To allow such isolation, some procedural changes in the program are often made. Very often the subjects and sometimes even the program staff are not told the real reasons for the changes. The staff who know the real reasons are encouraged not to talk about them. Consequently, resentment against the research may be built up among persons who are kept in ignorance of its nature and goals. As Mann (1965) observes, there is no simple solution for this conflict between research requirements and social reality.\footnote{131} Despite the necessity to

\footnote{130}{Ibid., p. 47}

control communication, he concludes that it is almost impossible for anyone to limit communication channels artificially. The next logical step then is to identify the intervening variables, the communication or input that may contribute to the client's changes. These intervening variables have to be separated from the independent variables; their effect from that of the independent variables. To be able to do so requires good knowledge of the technical details of intervening variable analysis which are presented in several textbooks and articles on research methodology.\textsuperscript{133} Weiss, (1972) classifies the intervening variables into two categories: program-operation variables, and bridging variables. The former has to do with the implementation of the program, whereas the latter has to do with the attainment of certain intermediate sub-goals. As to the sources of intervening variables, she observes that intervening variables are usually constructed out of the theoretical assumptions of the program; but they can also be developed empirically during the course of analysis.\textsuperscript{134}

\textsuperscript{132} Ibid, p.179.
\textsuperscript{133} Suchman, \textit{Evaluative Research}; footnote No.39, p.114.
\textsuperscript{134} Weiss, \textit{Evaluation Research}, pp.48-51.
Definition and measurement of the dependent variables

The third major methodological problem resolves around defining and measuring the dependent variable, the effect of the service or treatment. Many programs or services have unspecific, vaguely defined goals. To Rossi (1972) vague program goals is one of the main problems in evaluation.

A social welfare program which does not have clearly specified goals cannot be evaluated without specifying some measurable goals... When the policy makers do not specify goals clearly and still require evaluation, the evaluators are put into the uncomfortable position of deciding what were the goals the policy makers had (or perhaps should have had) in mind.\footnote{Rossi, "Testing for Success and Failure in Social Action," p.18.}

When the evaluator pursues the question, "What is the program trying to accomplish?", many program people give some sort of global statements like "improve education," "enhance quality of life," "reduce crime," "strengthen democratic processes." Thus he begins the long, often painful, process of getting people to state goals in terms that are clear, specific and measurable. The goal of any program for evaluation must be able to be translated into operational terms and observable events. For instance, the reduction in swearing is used as an indicator of the goal of "improving the behaviour of the youth." However, most programs find it extremely difficult to formulate goals in
these terms; and usually there is little consensus on goal indicators among the program people, and between the evaluator and the program people. Sometimes the disagreement centres around the adequacy of the indicators to represent the intended broad goals. Sometimes the evaluator has the trouble of helping the program people to differentiate long-term and short-term goals. Sometimes, he is faced with the dilemma of deciding which of the multiple goals of a program to evaluate. Furthermore, not all goals are of equal importance, and many can be translated into a variety of alternative indicators. How does one select from all possible indicators, those to be used for any particular purposes?

Value judgement

A very crucial question relating to defining and measuring the dependable variable is, "what do we mean by a successful result?" In other words, what should be the valid or appropriate criteria for measuring the result? Ballard and Mudd raise the issue "Whose judgement of improvement then represents the most appropriate measure of the effectiveness of treatment?"\textsuperscript{136} They distinguish at least four evaluating agents in relation to counselling programs whose definition of effectiveness might differ: the client, the persons closely related to the clients, the

society in general and the counselor. These four groups are likely to give different sets of criteria for measuring "success." The establishment of the criteria for measuring result thus involves value and choice.

Writing about value and subjectivity in evaluative research, Guttentag (1973) comments, "The dissatisfaction enters because evaluation research, in contrast to classical research, always involves a judgement of the worth of a program or project." Guttentag believes that the inherent subjectivity of decision-making renders classical experimental design not very useful. She, instead, presents a new model called "decision-theoretic approaches" for program evaluation. This model quantifies the decision-makers' values. "In decision-theoretic-approaches, the value judgements and preferences of the decision-maker are intrinsically bound to the data gathering process." To Guttentag (1973), it is only after the value dimensions have been clearly specified that the evaluator has a clear outline of what he must investigate. In this way, the research will be able to test the hypotheses of the decision-makers rather than those of the evaluator.

There is always the value dimension in defining and establishing the criteria for measurement. As Hemphill (1969) puts it:

---

138 Ibid., p. 62.
Evalitative studies differ from [non-evaluative] research in the manner in which value questions are involved — especially value questions that under-gird choices about what information or knowledge is sought.  

Despite the elements of value judgements, and in measurement criteria is crucial for evalitative research. "No evaluation can be made in the absence of some standard by which to judge success or failure."  

Value judgement is embodied in not only the establishment of criteria for measuring dependable variables but also in the following considerations: How valid are the objectives of the program? Are the outcomes of the program commensurate with the cost? What are the alternatives against which we compared or should have compared this program? And so forth.

Unintended effects.

Evaluation aims to provide objective, systematic and comprehensive evidence on the degree to which a program achieves its intended objectives plus the degree to which it produces unanticipated or unintended effects. As Wallin points out, social science literature is rich in examples of programs that produced totally unexpected side effects.  

---


140 Suchman, Evaluative Research, p.111.

141 Wallin, "Evaluating Effectiveness," p.32.
Reliability and validity

The last but not the least kind of difficulties in measuring dependable variables is the reliability and validity of the measurement itself. The reliability of a measure refers to the degree to which this measure can be depended upon to secure consistent results upon repeated applications. There are different types of reliability, those related to the criteria, to the instrument, to the design and to the object. Constancy or consistency is the main consideration in the analysis of reliability. Suchman (1967) in his analysis of reliability, identifies five major sources of inconsistency in evaluative research: subject, observer, situational, instrument and processing. Reliability is the necessary condition for validity which refers to the degree to which any measure or procedure succeeds in doing what it purports to do. The occurrence of bias is the focus of discussion on validity. The factors affecting validity in evaluative research have been identified by Suchman (1967) as: (a) propositional validity (the use of wrong theoretical assumptions); (b) instrument validity (the use of irrelevant operational indices); (c) sampling validity (not representative enough); (d) observer or evaluator validity (consistent bias); (e) subject validity (modes of expression, motivation); (f) administration validity (use of staff, conditions); and (g) analysis validity.

\footnote{Ibid., pp.118-9.}

\footnote{Ibid., pp.122-3.}
The above-said problems pertaining to reliability and validity are indeed common to social research, the discussion of which is easily available. They are therefore not discussed further in this paper.

Socio-Political and Organizational Constraints

The process of evaluation involves getting information within a context of social relationships. The key persons who are involved or have vested interest in evaluation include the program planners, administrators and staff, the evaluator and his staff, and the potential consumers who have no direct fiduciary or operating responsibilities regarding the program. In the planning stage of the evaluation, the key persons involved are likely the evaluator and those closest to the decision-making power, that is, the program sponsors and administrators. In the conduct of an evaluation, there is interaction primarily among the evaluation staff and program staff. Among these key interest groups, there exist several common but crucial problems that are likely to effect the quality as well as utilization of the evaluation.

Conflicting expectations regarding evaluation

The first and foremost of all is that the key persons may have conflicting perceptions of evaluation purposes. Some program administrators and staff may see evaluation as a ritual to secure funding; others may view it as a potential
guide to program modifications. Evaluators, on the other hand, are concerned with the production of a piece of work which is respectable to his professional colleagues. Program sponsors may be most interested in program expenditures and efficiency. Differing expectations about the purposes of evaluation may result in several unfortunate consequences. When administrators or program staff expect formative evaluation to aid on-going program developments, evaluators may design summative studies to render judgement on the program after its conclusion. The administrators seeing little value in having summative studies may exert control over funds, staff and decisions affecting research, such as criteria measures, size of sample and sources of data. Further, the administrators and the staff may expect the evaluation to legitimate the program. Most do not seriously entertain the notion that evaluation results should show their program in an unfavourable light. The fear of negative findings may then prompt some to interfere with evaluation.

Even when evaluation is undertaken for bona fide purposes people can have widely differing expectations. Top policy makers need the kind of information that will help them address the broad issues: Should the program be continued, dropped or modified? Should more money be allotted to this program or to others? They want information on the overall effectiveness of the program. Direct-
service staff have practical day-to-day concerns about techniques. Should they spend more time with clients or in case conference? Should they accept more middle class people or more lower class people? They often challenge evaluation to come up with something practical on topics such as these. Weiss (1972) identifies seven possible consumers of the evaluation: (a) a funding organization (government, private, foundation); (b) a national agency (government, private); (c) a local agency; (d) the directors of the specific project; (e) direct-service staff; (f) clients of the program; and (g) scholars in the disciplines and professions.

Different groups of consumers may have discrepant notions of what the evaluation should be. No single evaluation can serve all consumers in the same way. Therefore, as Tripodi et al. (1971) point out, a crucial element in obtaining useful evaluations is the extent to which the key persons involved in evaluations are able to make appropriate accommodations to each other. Tripodi et al. (1971) suggest that useful evaluations are more likely to occur when the following two conditions are met:

(1) There is a clarification of the purposes of the evaluation among key persons involved.

---

(2) There is an agreed-upon commitment, contractual or understood, regarding the uses and possible consequences of the evaluation.¹⁴⁷

It is clear that the first task confronting the evaluator is for him to find out the answers to the questions: Who initiated the idea of having an evaluation and for what purposes? Are there groups in the organization who question or object to the evaluation? For whom is the evaluation conducted? What are the different expectations of the staff, administrators and funders or both? The challenge that follows is for the evaluator to solicit the involvement of the various interest groups in clarifying objectives and criteria for evaluating the program. There may not be consensus even after long periods of discussion and negotiation. Sometimes, the evaluator might well ponder whether he wishes to get involved in the situation or whether he can find more productive uses for his talents elsewhere.

Role Differences

"Practitioners have to believe in what they are doing; evaluators have to doubt."¹⁴⁸ Practitioners or program staff are committed to a project; they have invested enormous amounts of time, energy and professional expertise. Evaluators are committed to the search for evidence to

¹⁴⁸ Weiss, "Between the Cup and the Lip," p.52.
prove the success or failure of the project; they are basically critical of the effectiveness of the practitioner. For the practitioner, an attitude of confidence in his own method is a prerequisite to good practice. In the light of this attitude, professional defensiveness against having his practice brought under question is understandable. When several practitioners are involved, the problem becomes even more delicate and complex.

The practitioner, as in the case of many social work settings, focuses upon a series of individual cases, while the researcher focuses upon a general problem. This occasionally gives rise to the attitude that the researchers are not interested in the individuals and the practitioners are not interested in the general problems. As Rodman and Kolodny (1964) comment:

In their less charitable moments, researchers complain that practitioners 'can't see the forest for the trees' while practitioners, in turn, wonder whether researchers 'can see the human beings behind the Statistics'.

Another problem that may arise from role differences concerns the credit that is assigned for the publication of research reports. To the researcher, publication is an essential part or indication of his work, and he expects to get primary if not sole, credit for publication. The practitioner who has co-operated with the researcher,

\[149\] Hyman Rodman and Ralph Kolodny, "Organizational Strains in the Research-Practitioner Relationship," Readings in Evaluation Research, p.118.

\[150\] Ibid, p.123
typically feels that he has contributed a great deal to the research project and expects to get substantial, if not equal, credit for its publications. As a solution to this problem, Rodman and Kolodny suggest the use of co-authorship, footnoting, and secondary authorship ("with the assistance of" or "with the collaboration of").

Programs do not always keep to their original course; they are responding to a host of factors: budget cutting, budget expansion, changes in administration and in top officials, public appraisal, or initiation of rival agencies and programs. Whereas the evaluator wants to study the effects of a stable and specifiable stimulus, program administrators have much less interest in the integrity of the study than in assuring that the program makes the best possible adaptations to conditions.

There is also basic difference in stance between evaluative research and the public relations of an agency. Because evaluative research forces attention to what is actually done rather than what is intended, evaluative research may expose the realistic operations of the agency in contrast to its public appearance. An organization that has existed for some time will ordinarily acquire many secondary functions beyond those for which it has been established. To make a public acknowledgement of these

---

activities may well be upsetting, and resistance to evaluative research may arise from this basis.

Organizational Strains

Evaluation usually places a strain on the ordinary administrative and operative procedure. It may require changes in the routine such as an interruption in the flow of cases into the agency as well as work habits. Adjustments on the functioning of the agency and its staff may thus be necessary. One area in which change may be asked of the practitioner is in terms of fuller and more frequent recording. The other is that the practitioner may be asked to attend research interviews or meetings. Sometimes, the practitioner prefers to save time from these activities required by evaluation so that he can devote more time to doing what he considers his basic job—clinical work.

The staffing of the evaluation is another source of strain. The evaluation staff who are from outside of the agency may face entry problems. They may have difficulties in getting cooperation from the agency staff; they need some time to learn about the program objectives and operations. However, the evaluation conducted by the "insiders" may be suspect to some, as not being completely objective, and as providing biased findings.

The organization location of evaluation may give rise to conflict. To whom should the evaluator report? At what level of authority should the evaluator be? If the evaluator
reports directly to the funding body, the agency administrators and staff may see the evaluator as "inspector" checking on them, and become wary of divulging information that might reflect poorly on their performance. The evaluator will not have the easy, informal contact with program staff. If the evaluator is responsible to program directors, he is under all kinds of pressure not to come up with findings that disparage the effectiveness of the whole program. If he does, the directors may stall the report or may color it with comments like, "the evaluation was too crude to measure the important changes that took place." That is why Williams and Evans (1969) recommend, "Major evaluations of programs should be performed by an office and staff removed from the operating program." 152 Rodman and Kolodny (1964) suggest the use of research consultants as one way of providing external structural supports for the researcher who occupies a marginal position within an agency. 153

Differences in value systems, personality differences and lack of clarity about boundaries of responsibilities and procedures in reality at the initial stage of evaluation, are also potential sources of conflict.


The Political process in the functions of evaluation.

By systematic and objective methods, evaluation research examines the effects of policies and programs on their targets, the goals they are meant to achieve. The assumption is that by providing the "facts," decision-makers are assisted in making wise choices among future courses of action. However, how valid is this assumption? To what extent can careful and unbiased data on the results of the programs really improve decision-making? How much of the decision-making on policy and programs are political, and how much of it is affected by evaluation? As Weiss (1973) observes, evaluation is a rational enterprise that takes place in a political context. Political considerations intrude in evaluation in terms of its functions in three major ways.  

First, social programs are creatures of legislative politics and bureaucratic politics. The policies and programs are proposed, defined, debated, enacted and funded through political processes, and in implementation they remain subject to pressure - both supportive and hostile. Because of the political processes of persuasion and negotiation that are required to get a program enacted,  

---

154 This and much of the following of this section are adapted from the article by Carol Weiss, "Where Politics and Evaluation Research Meet," *Evaluation*, Vol. 1, No. 3, 1973, pp. 37-45.
inflated promises are made in the guise of program goals. For instance, public housing will not just provide decent living space, it will improve health, enhance marital stability, reduce crime and lead to improved school performance. Sometimes program goals are diffuse and incompatible because holders of diverse values and different interests have to be won over; in the process, a host of realistic and unrealistic goal commitments are made. Given the consequent diffuseness of program goals, there tends to be little agreement, even within the program, on which of the goals should effort be spent and which are mere window-dressing. People at different levels in the system perceive and interpret goals in different ways. "Unless the evaluator is astute enough to direct his research towards the authentic goals, he winds up evaluating the program against meaningless criteria."155 It is not uncommon for evaluation reports to meet remarks like, "But that's not what we were trying to do." Further, as said earlier, programs, because of a host of factors they have to respond to, can shift in activities and in overall strategy and even in the objectives they seek to attain.

The second element of politics in the evaluation context is that the evaluative reports have to enter the political arena and to "compete for attention with other

155Ibid., p.39
factors that carry weight in the political process.\textsuperscript{156} The decision-making system in a democratic setting has its own values and its own rules. The decision on the evaluative outcomes may involve trade-offs with interest groups, professional guilds and salient publics. A considerable amount of ineffectiveness may be tolerated if a program fits well with prevailing values, if it satisfies voters or funding sources or if it pays off political debts. What evaluation research can do then is to clarify what the political trade-offs involve. Evaluations are sometimes readily disregarded because they address only official goals. A program's effectiveness in meeting political goals should also be assessed. "Political-benefit analysis might help to resolve questions about political benefits and foregone opportunities."\textsuperscript{157}

The third element is the stance of evaluation itself. By the very nature of evaluation, when a social scientist agrees to evaluate a program, he gives legitimacy to: (a) the rightness of the social diagnosis, of the program goals, program strategy and program activities, and (b) that the program has a realistic chance of reaching the goals. Further, there is the implicit acceptance of the political decisions in the selection of some programs to undergo evaluation.

\textsuperscript{156}Ibid, p. 37.
\textsuperscript{157}Ibid, p. 40.
while others go unexamined, and in the structuring of the evaluation research such as funding, staffing and structural position. But the basic proclivity of evaluation research should be to improve the way the society copes with social problems and to subject program assumptions to scrutiny, in addition to locating discrepancies between intent and actual outcomes.

Evaluation conclusions are the identification of discrepancies between goals and outcomes, and the usual recommendations will call for modifications in program operation. The assumptions here are: (a) that reforms in existing policies and programs will serve to improve government or agency performance without drastic reorganization, and (b) that decision-makers taking the evaluative evidence will respond by improving programming. However, from the past evidences and experiences there are no strong indications that evaluation research of one program at a time can draw useful implications for action or that piece-meal modifications will improve effectiveness. As for the other assumption on which the utility of evaluation research is based, there is not much positive evidence either. The politics of survival and the politics of higher policy-making accord evaluative evidence relatively little weight in the decision making. The evaluation researchers are then confronted with the political stance which is real and painful.
Because evaluation researchers tend to be liberal, reformist, humanitarian, and advocates of the underdog, it is exceedingly uncomfortable to have evaluation findings used to justify an end to spending on domestic social programs. On the other hand, it is extremely difficult for evaluators to advocate continuation of programs that they have found had no apparent results.158

Summary

Service evaluation is a form of accountability which has been increasingly demanded by funding bodies, legislators and the public, especially after recent disclosures that the major social programs do not achieve the anticipated results. As a result, many administrators, planners and policy-makers have become greatly interested in evaluation research, and have been trying to build in evaluative devices to monitor feedback for service modifications, and for changes in service objectives. As social science research advances, evaluation has, in turn, been expected to yield more accurate, reliable, specific and objective information on program results.

Different expectations of evaluation have led to various definitions and conceptualization of functions and types of evaluation research. Different types or levels of evaluation require different methods and skills to carry out. Hence, various models and techniques of evaluation research have been developed.

158 Ibid, p.44
At this stage of development, evaluation research has yet to fulfil its functions to the fullest extent; it is still limited by some technical difficulties and various practical constraints.

In this chapter the researcher has explored in great detail, evaluation research, an increasingly popular subject in social service. The following dimensions of evaluation research have been looked at closely: functions and types; models and techniques; problems and constraints.

It is hoped that after reading through this exploration, a better understanding of the nature, functions and problems of evaluation research can be developed.

To the researcher, the concepts relating to these dimensions derived from the literature search have provided some kind of conceptual clarity on the study of the evaluative practices of the U.C.S. member agencies. They are especially useful in the formulation of the research focus; in the construction of the interview schedule; and in the development of ideas for improving agency service evaluation.
Leaf 90 omitted in page numbering.
CHAPTER III

RESEARCH DESIGN AND METHODOLOGY

As discussed earlier in this report, there are various systems of classification of research design.\textsuperscript{159} Research design varies with the purposes of the research, the types of questions it seeks to answer, the levels of certainty or degree of accuracy desired. These three interrelated independent variables are further associated with the state of available knowledge on the problems being investigated.\textsuperscript{160} Therefore, the choice of a study design is complicated by many factors.

In this research project, there has not been any previous research conducted on the problem to be researched, that is, the evaluative practices of the U.C.S. member agencies. In other words, the state of available knowledge on the research topic does not allow for a more refined type of research design, the experimental design, or even the quantitative-descriptive type. There needs to be more knowledge based on some preliminary research before the

\textsuperscript{159} Please refer to p.39 & 40 of this report.

\textsuperscript{160} For discussion on this topic, please refer to Claire Selltiz et al, Research Methods, pp.50-1; Alfred Kahn, "The Design of Research," pp.49-50.
nature or certain dimensions of the agency evaluation could be determined or formulated for further study. The purpose of the study, as stated, is to gain familiarity with various dimensions of the evaluative practices. Therefore, the design which is most appropriate for the research is the exploratory type as conceived by Kahn (1960), Selltiz et al (1951), and Tripodi et al (1969).

While sometimes regarded as rather unscientific, the exploratory research fulfills the purpose of being "relevant to broader issues than those posed in the experiment,"\(^{161}\) by exploring the dimensions of the problem with which the research is attempting to deal. Its main function is to develop hypotheses or formulate problems for more precise investigation.

The research study could be further sub-typed as a combined exploratory-descriptive as described by Tripodi et al (1969). This sub-type "is intended to serve as a transition between exploratory and quantitative-descriptive studies."\(^{162}\) One of the distinguishing characteristics of this sub-type is that the phenomena being studied are described both from a quantitatively as well as a qualitatively position.\(^{163}\)

\(^{161}\)Selltiz et al, Research Methods, p.53.

\(^{162}\)Tripodi et al, Social Research, pp.25-6, 49-51.

\(^{163}\)Ibid, pp.25, 49.
The nature of an exploratory study demands a certain flexibility in terms of research procedures. As pointed out by Selltiz et al. (1951), methods must be used flexibly in exploratory study; "frequent changes in the research procedure are necessary"\textsuperscript{164} in order to provide for the gathering of data meaningful to the research purpose. To Tripodi et al., "sampling procedures are flexible, and little concern is usually given to systematic representativeness."\textsuperscript{165}

In light of the flexible nature of exploratory research, evaluators should trust that the researcher used scientific judgement discreetly and with as little bias as possible.

**The Focus**

This research project was an attempt to find out the why, what and how of the service evaluation of the member agencies of the U.C.S. Restated more precisely, the research project was designed to yield empirical data on the following aspects of the evaluative practices of the U.C.S. member agencies:

- the purposes for the use of service evaluation;
- the types of information kept in service evaluation;
- the methods used in service evaluation;

\textsuperscript{164}Selltiz et al., Research Methods, p. 53
\textsuperscript{165}Tripodi et al., Social Research, p. 49.
- the value of service evaluation; and
- the kind of help needed for improvement of service evaluation.

The data was collected through the use of an interview schedule (Appendix A). The respondent in the interview was the executive director of each agency. Therefore, to be accurate, the information on the above aspects actually represented only the respondent's reporting of these aspects, not necessarily his perception of or the "objective" facts about these aspects. The chief executive was considered to be the most appropriate person to talk about the agency evaluative practices. It was stressed to the respondent that the evaluative practices of the individual agency would not be identified or disclosed; and that knowing the true level of functioning was essential to developing any useful plan for improvement. Such a strategy should be helpful in getting truthful reporting. As an additional measure to ensure that the data as a whole would reasonably reflect the evaluative practices, the total population would be studied rather than a sample.

The researcher therefore, considered it reasonable to assume that the data thus obtained would be adequate for: (a) the agencies themselves to look at their evaluative practices in relation to those of others; (b) the conceptualization of some approaches to improve the agency evaluation. Further, the findings should be of interest to both
academicians and practitioners who are interested in agency service evaluation.

**Definitions**

Included in this section are definitions and operational definitions of some terms. So as to allow for conceptual clarity, definitions indicate exactly what the term meant to the researcher. Operational definition refers to a description of any indexes to be used as indicators of the concept, such as I.Q. score to indicate intelligence. In other words, the researcher has to "devise some operation that will produce data he is satisfied to accept as an indicator of his concept."\(^{166}\)

The purposes for the use of service evaluation, in the statement of the research focus was operationally defined as the respondent's reporting of:

- (a) the main reasons for having service evaluation;
- (b) the utilization of the evaluation results.

The methods used in service evaluation would be indicated by the respondent's reporting of the following:

- (a) ways (methods) used in evaluating service effectiveness;
- (b) methods preferred for evaluating services;
- (c) specification of the objectives in measurable terms;

---

\(^{166}\) Selltiz et al, Research Methods, p.43.
(d) adequacy of the methods used;
(e) preference for agents to conduct evaluation;
(f) difficulties or limitations experienced.

The value of service evaluation was expressed in terms of the respondent's reporting of the following:
(a) usefulness of existing service evaluation;
(b) importance of service evaluation in general;
(c) direction of change for the existing service evaluation.

The term evaluation research has been previously defined on page 10 of this report. The term evaluation was defined as "the general process of judging the worthwhileness of some activity regardless of the method employed." Evaluative practices was therefore whatever was done in the area of evaluation by agencies. No substantive distinction was made between evaluative research and evaluation research, though in the former, the stress is on research, whereas in the latter on evaluation. They were used interchangeably in this report.

Improvement in service evaluation was conceived as movement towards (a) wider area of evaluation, that is, covering more areas or aspects, and (b) attainment of more objective, specific or accurate information, that is, data obtained by more reliable, scientific method of investigation.

167 Suchman, Evaluative Research, p. 31.
Member agencies of the U.C.S. was defined as those agencies which had U.C.S. membership status on the date the interview schedule was mailed. For reference, the member agencies of the U.C.S. are listed in Appendix B.

Population and Sample

The population in this research project was all the member agencies of the U.C.S. The reasons for not defining the population as member agencies of all U.C.S. or United Funds in Canada was that the researcher was limited by time and resources to sample from such a population and to study such a wide-spread sample. Further, he had not studied the representativeness of the Windsor U.C.S. member agencies in terms of evaluation as a sample for their counterparts throughout Canada. However, he believed the findings may have some implications for other settings.

The researcher decided to study the total population, twenty-seven agencies, as a sample so that the findings could give a more accurate picture of the overall evaluative practices of the member agencies of the U.C.S. Windsor. The implications of such an arrangement for the research methods were that there would be no sampling procedures and no statistical testing on the significance of the findings for generalization to the population. Since the entire population was studied, the term sample was not used in its strict statistical sense.
The Setting

Windsor is Canada's southern-most city. It is linked to Detroit of the United States by a bridge and a tunnel. Its population is about two hundred thousand.

Windsor is one of Canada's big industrial cities and is the fifth largest manufacturing centre of Canada. 168 It is well-known for its liquor and automobile industries.

There are strong labour unions in Windsor. Both the labour and the big corporations actively support the U.C.S. in terms of contributions and participation in campaigning for funds.

The United Community Service of Greater Windsor was formed in 1962 when the then Community Welfare Council merged with the United Fund. The objective of the U.C.S. is "the provision of voluntary community services to meet existing and emerging needs in the geographic community it serves." 169 At present, the U.C.S. is composed of two major parts: Social Planning and United Fund Campaign with each having its own board of directors. The executive director of the U.C.S. is a professional social worker. He, together with the professional staff of the Social Planning Division,

---


provide consultation to the member agencies of the U.C.S. Currently, all the professional staff of the Social Planning Division are professional social workers.

Each year before June, the U.C.S. member agencies submit their program and budget presentations for funding. The U.C.S. does specify some conditions for allotting funds. All its member agencies have been urged or advised to evaluate their services. Besides, they are in the second year of the three-year program to establish functional budgeting.

The characteristics of the U.C.S. member agencies were discussed in Chapter IV, Data Analysis.

**Methods of Data Collection**

At the beginning the researcher toyed with the idea of conducting the survey by either a questionnaire or interviews with an interview guide. Questionnaires might give rise to different interpretations or miscomprehension of some technical terms to be used; they tend to result in a low return rate. Interviews, usually of brief nature, would give the researcher too much judgement of the evaluative practices. These considerations finally led to the development of a standardized interview schedule which would have been termed questionnaire without the interview. The questions and the technical terms necessary for the investigation were

170 Sometimes the term questionnaire is used to include standardized interview schedule.
made as plain and as simple as possible.

The first draft of the interview schedule was reviewed by professors, fellow students and a group of twelve practitioners consisting of those working in public bodies in Windsor, and staff members of both the Windsor U.C.S. and the Detroit U.C.S. Seven of the twelve participating indicated they would not need an interview to complete the interview schedule. In the pre-test six of eight gave the same indication. Of the six, five indicated they preferred not to have the interview. As a result of these processes, the researcher decided to use the following procedures in collecting data.

First, the interview schedule would be mailed out with a cover letter to the executive director of each agency with a duplicate copy of the letter sent to the president of the agency. In addition to explaining the rationale and purpose of the research project, the letter states that the researcher would phone soon for an appointment for an interview.

In the next stage, between four to seven days after mailing out the interview schedule, the researcher would phone each for an interview. In the conversation, the researcher would indicate the option of their completing the interview schedule without an interview. It would be indicated at the same time that should they choose not to have an interview, a further phone call would be necessary
to discuss any questions that might arise in their going through the interview schedule.

In the third stage, either an interview or a telephone conversation took place for completion of the interview schedule. In the latter case, the interview schedule could be delivered, picked up or sent by mail.

To the researcher, the procedures should have the following advantages. They would

(a) give the respondent the choice of following whichever way he felt more comfortable with, thus reducing resentment or resistance;
(b) ensure a higher return than by questionnaire;
(c) avoid possible undue alarm or frustration about some technical terms;
(d) in some cases, save time for both the researcher and the respondent.

The researcher did not believe that the option would result in discrepancy in reporting or that the reporting without interview would be less truthful.

The researcher feared that if the research project was seen as one sponsored by the U.C.S. there might be some undue influences on responses. He, therefore, decided to use the stationery of the School of Social Work and to state in the letter that the project was carried out under the guidance of a Thesis Committee of the University of Windsor School of Social Work (Appendix C).
The Interview Schedule

Though the term questionnaire is sometimes used to refer to standardized interview schedules as well as forms to be filled in by the respondents,171 in this research project the term interview schedule was preferred because the majority of those who were asked to comment on the first draft of the interview schedule indicated their preference for the term interview schedule to the term questionnaire. However, the considerations necessary for the construction of a questionnaire are relevant and essential to the construction of the interview schedule.

In constructing the interview schedule, the researcher followed the six steps suggested by Kornhauser and Sheatsley: deciding what information should be sought; deciding what type of questionnaire should be used; writing a first draft; re-examining and revising questions; pre-testing; editing the questionnaire and specifying procedures for its use.172 He further took into consideration their suggestions of the three main areas of considerations for questionnaire construction: decisions about question content; decisions about question wording; and decisions about form of response to the question.173


172 Ibid, pp. 547-552.

173 Ibid, pp. 552-574.
The researcher was very conscious of the use of some technical terms in the interview schedule, such as "staff's quality of work," "efficiency of service delivery," "effects of the services," "the benefits the services have brought to the clients," and "the effects of intervening variables." He could not find other terms which expressed the same meaning. As a remedy, written explanations were given in the interview schedule, and advanced preparations were made on the explanation of these terms in interviews and telephone conversations. However, in the pre-test, only two respondents sought clarification on terms.

As to question five on "types of services," the researcher first used the U.C.S. classification which groups services into six types, namely, counselling, character-building, health, social and recreational, correctional, and services to the handicapped. Feedback indicated that these groupings were not satisfactory for use in the interview schedule. The researcher also found the classification of the U.C.S. of Metropolitan Detroit unsatisfactory. The classification of 99 types of services suggested by the Community Funds and Councils of Canada for voluntary organizations was indeed not practical for the interview.

174 It was told to the researcher that such classification was arbitrary and for internal use of the U.C.S.

175 Community Funds and Councils of Canada, Functional Budgeting for Canadian Voluntary Organizations, (November, 1972), pp. 7-9. Place of publication was not given.
schedule. Finally, after a majority of the sample of eight fellow students and three agency directors had indicated their preference among the three types previously mentioned, the researcher decided to use Rapoport's grouping of social services.176

The researcher also considered the advantages and limitations of both open and closed questions.177 In the first draft five questions were open-ended. After feedback, the researcher decided to standardize the format of the interview schedule, by categorizing the possible responses, and by adding two other columns, "other" and "comments" for any other responses.

In constructing the interview schedule, the researcher also took into consideration the length of the questions, the order of the questions, the total number of questions, and the number of questions relating to each subject as suggested by Duvenger.178 The questions in the interview schedule were centered on the research focus with emphasis given to finding out the method of evaluative practices. In addition, a few questions were devoted to

177 Selltiz et al, Research Methods, pp.256-263.
yield data on the characteristics of the sample.

The five-point scale, from "a great deal" to "not at all" used in the interview schedule was adopted from Fleishman's Leadership Opinion Questionnaire.\textsuperscript{179} It was used to permit the expression of different degrees of agreement-disagreement. In all the three stages of development of the interview schedule, namely, the first draft for comments, the second draft for pre-test and the final form for use, there were only two or three respondents suggesting to change the phrase "fairly much" in the scale. The researcher had thought of using "very much" to replace "fairly much," "extremely" to replace "a great deal" and "somewhat" or "so so" to replace "to some extent." He was, however, not satisfied with the would-be replacements.

The scale used belongs to the Likert-type which is an ordinal scale. In other words, the scale makes possible the ranking or ordering of five degrees of agreement-disagreement, but it does not measure the amount of change or distance between the different degrees. The distance between the two degrees of agreement-disagreement, for example, "a great deal" and "fairly much," or that between "fairly much" and "to some extent" are not measured or indicated on the scale.\textsuperscript{180}

\textsuperscript{179}Fleishman, as cited by Miller, \textit{Handbook of Research Design}, p.229.

\textsuperscript{180}For discussion on the advantages and limitations of Likert-type scale, please refer to Selltiz \textit{et al}, \textit{Research Methods}, pp.366-70.
As a whole, in constructing the questionnaire, the researcher did not follow a mathematical model, nor did he standardize the number of items for each question and the form of responses, though model and standardization usually give better efficiency and precision to data analysis. The main reason was that he believed that the use of any one model or rigid regulation of responses might have caused a sacrifice of some responses (data) essential to the research purpose which was essentially a survey of some aspects of the evaluative practices.

Pre-Test

The questionnaire was pre-tested with a sample of eight voluntary agencies, five in Detroit and three in Windsor. Like the U.C.S. member agencies, each of the agencies in the pre-test had a board of directors consisting of volunteers. The respondent was the executive director.

In the Windsor area the researcher made the initial telephone contact to some agencies suggested by the U.C.S. staff. In Detroit, the initial contact was made by Mr. Edward Marsh, of the Agency Relations Division of the U.C.S. of Metropolitan Detroit. The interview schedule with a cover letter was then either sent by mail or delivered by hand to those who had agreed to participate in the pre-testing. In the letter the researcher solicited comments on wording, order, gaps, et cetera of the interview
schedule; in addition, it stated the general purpose and for whom the interview schedule was intended.

Three to five days after the delivery or mailing of the interview schedule the researcher made a telephone call to each respondent for an appointment to go through the interview schedule. However, at this juncture the option of completing the interview schedule without interview was indicated. Two gave appointments for interviews; six requested another telephone call to ascertain the situation or to allow them to seek clarification on questions arising from the interview schedule. Of the six, only two sought clarification on one or two terms. All six completed the interview schedule without interview.

Consequential to the pretest, improvements were made, mostly on the layout and strategy of the interview schedule.

**Methods of Data Analysis**

Simple statistical procedures were used to tabulate and analyze the data. Since there were no sampling procedures, there were no statistical tests on the significance of the findings. Only descriptive statistics were used to summarize and cross-tabulate the data obtained.

The data analysis consisted of two levels of operations. The first was the tabulation of frequency distributions of the responses which provided a general
picture of the aspects of evaluative practices investigated. The frequency distributions reflected some general trends or patterns and led to some conclusions. The second level of analysis was concerned with seeking variable relationships which is one of the important functions of the exploratory study. There were ninety seven variables, and many of their associations offered possibilities for investigation. Limited by time, the researcher, however, studied only those about which he had some hunches or those in which the U.C.S. had indicated an interest. The variables were cross-tabulated, using mainly contingency tables and chi square.

The computer was used in data analysis. The interview schedules returned were carefully checked and coded; the information was key-punched on the computer cards and verified accordingly. The Statistical Package for Social Sciences\textsuperscript{181} was used for the computer run.

The analysis of data was divided according to the five areas of the research focus. A sixth section was devoted to a description of the sample.

\textbf{Utilization}

A copy of the report would be given to the U.C.S. which may seek permission from the University of Windsor to

reproduce parts or all of it for distribution to any agencies desiring it. As agreed upon, all agencies participating in the project may have access to the findings of this research.

According to the Graduate Studies policies for all thesis, two copies of this report would be placed for reference in the library of the University of Windsor and one in the School of Social Work of the University of Windsor.

Limitations of the Study

There were a number of limitations of the study. The first was related to the research design. The exploratory study allows for flexibility in not only the research procedures but also the width and depth of the research. Ideally, the interview schedule could have been expanded to cover other aspects of the evaluative practices. Similarly, more variable relationships could have been studied in the analysis of data. 182

The second was related to the reporting of the respondents. The chief executive was selected as the respondent. In reporting, he might take into considerations the public image of the agency. Here, the discussion earlier

182 For further discussion on limitations of the exploratory study, please refer to pp. 92-3 and p. 108 of this report.
on the disadvantages of using "inside" evaluator might fit in. In order to achieve uniformity, the researcher had to accept whatever was reported despite instances in which he was not in agreement with the ways the respondent interpreted and reported the practices. However, to take the researcher's interpretation based on such a short interaction would be "unacceptable" also since this would have injected another kind of bias.

The third limitation was the connection with the U.C.S. The researcher had his field placement at the U.C.S. and his field supervisor, the executive director of the U.C.S., was also a member of the Thesis Committee. During the period of the collection of data, the U.C.S. agencies were about to make their submissions for funding. Since the agencies had been encouraged by the U.C.S. to evaluate their services, the connection and the timing may have had undue influences upon the reporting.

Fourthly, the interview schedule consisted of some technical terms which might not be fully comprehended by some respondents, and they may have been hesitant to ask for clarification despite attempts to provide this.

Fifthly, the survey of the literature could be viewed as a limitation in that it was well beyond that which was required for the Masters' thesis. Actually, the literature survey reflected more the researcher's interest at that time in exploring as many aspects of the topic as
possible, as such an extensive search was permitted by the exploratory type of study.

Sixthly, as explained earlier, the generalizations of the findings were not certain.

In addition, the researcher would consider it a limitation to the dissemination of the data that the full report of the findings could not be given to each of the participating agencies.

Summary

Discussed in this chapter were the decisions on various aspects of the research project, from the choice of a design, through the transformation of research questions, the tool and procedures of data collection, the methods of data analysis, the plan for utilization of the report, to the limitations of the study. In the discussion, the reasons for, principles of, and circumstances leading to making the decisions were given.

The research project was carried out according to the design described. The data obtained accordingly was presented and analyzed in the following chapter.
CHAPTER IV

ANALYSIS OF DATA

The data was collected through the use of an interview schedule according to procedures as described under "method" and "pre-test." Sixteen interviews or 59 per cent were conducted in person and the balance by telephone. A total of 23 returns or 85 per cent was obtained from the sample of 27. However, only 22 were used for data analysis, as 1 respondent gave information about the practices of the main office instead of the practices of the local branch. The interviews ranged from 20 to 90 minutes, with most being within 35 to 50 minutes.

Of the 4 from whom the data could not be obtained within the time limit, 1 reported to have sent the interview schedule for completion to their regional representative stationed outside of Windsor. The researcher did not make the subsequent contact. Of the 3 remaining, either there was no full-time paid staff or the executive director was too busy to complete the interview schedule.

It took 6 weeks to complete all the necessary interviews and obtain the 23 returns. There was a total of 68 telephone calls with an average of 2.5 calls per agency. When first contacted by phone, 2 respondents reported that
they received only the letter not the interview schedule. Two others asked for an additional copy of the interview schedule; a third asked for ten copies for distribution to his staff and board of directors.

As a whole the respondents were very co-operative. Some were very enthusiastic about the project and asked that the researcher give them a copy of the study result. There were a few cases in which the researcher met resistance. However, there was a common phenomena. Many respondents felt uncomfortable because their evaluation was not inclusive of all the aspects described in the interview schedule. The researcher had to explain that the interview schedule was meant to cover many aspects as well as those related to various types of services; and that the researcher did not expect each to have covered, tried, or experienced all those activities described in the interview schedule. Generally, this explanation helped to put them at ease.

As to the interview schedule, some described it as "very good." A few respondents sought clarifications on terms such as "quality of work," "efficiency of service delivery," "effects of the services," "the benefits," "definition of need," "having much value judgement," "the effects of intervening variables," and "complications caused by various stages of program development."

The processes in the collection of data described
above could serve as a background of the analysis of data. In order to answer the research questions, the analysis was divided into the following six sections for discussion: the characteristics of the sample, the types of information kept in relation to service evaluation, the methods in service evaluation, the purpose of service evaluation, the value of service evaluation, and the help needed for improvement of evaluation practice.

The Characteristics of the Sample

The term sample henceforth refers to the 22 agencies on which the data was obtained for study.

Questions 1 to 5 yielded data that described the sample in terms of length of existence, length of U.C.S. membership, staff complement, annual operating budget, and types of services.

Table 1 shows the distribution of the sample by the length of their existence.

<table>
<thead>
<tr>
<th>TABLE 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>LENGTH OF AGENCY EXISTENCE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Length in Years</th>
<th>Number of Agencies</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 4</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>5 - 9</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>10 - 14</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>15 - 19</td>
<td>2</td>
<td>9.1</td>
</tr>
<tr>
<td>20 and above</td>
<td>13</td>
<td>59.1</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>100.</td>
</tr>
</tbody>
</table>
The majority (59 per cent) of the sample had been in existence for more than 20 years, while only 18 per cent were "young", that is in existence less than 10 years. The table was collapsed into 2 categories for cross-tabulation. The first consists of 41 per cent or 9 agencies in existence less than 20 years, while the second consists of 59 per cent or 13 agencies in existence 20 years or above.

Table 2 indicates the distribution by the agency length of U.C.S. membership.

<table>
<thead>
<tr>
<th>Length in Years</th>
<th>Number of Agencies</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 2</td>
<td>4</td>
<td>18.2</td>
</tr>
<tr>
<td>3 - 4</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>5 - 6</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>7 - 8</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>9 - 10</td>
<td>2</td>
<td>9.1</td>
</tr>
<tr>
<td>11 and above</td>
<td>12</td>
<td>54.5</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>100.0</td>
</tr>
</tbody>
</table>

As in the case of agency life, 55 per cent fell in one category which made interpretation of cross-tabulation less ideal. The U.C.S. has been in existence for about 12 years. Before that, it was Community Welfare Council. From the distribution it can be assumed that the majority of the sample were chartered members of the U.C.S., that is, they were members at the time the U.C.S. was incorporated.
The staff complement of the sample is shown in tables 3 to 5. The distribution of the sample by the number of full-time paid staff was presented in Table 3.

### TABLE 3
**NUMBER OF FULL-TIME PAID STAFF**

<table>
<thead>
<tr>
<th>Number of Full-Time Paid Staff</th>
<th>Number of Agencies</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 - 5</td>
<td>11</td>
<td>50.0</td>
</tr>
<tr>
<td>6 - 10</td>
<td>6</td>
<td>27.3</td>
</tr>
<tr>
<td>11 - 15</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>16 - 20</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>over 20</td>
<td>4</td>
<td>18.2</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22</strong></td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

Seventy-seven per cent of the sample employed no more than 10 persons, while only 18 per cent employed over 20 full-time paid staff. The distribution would indicate a high trend toward small agencies.

Of the total number of full-time paid professional staff, there is an indication again of small rather than large operations.

Each agency had its own definition of professional staff. The researcher found that professional staff was generally defined as one who (a) had a professional degree; (b) had gone through some sort of specialized training; or (c) had acquired some sort of expertise in his field of
work through experience. The term professional staff, was used rather loosely among the sample without a uniform definition.

**TABLE 4**

**NUMBER OF FULL-TIME PAID PROFESSIONAL STAFF**

<table>
<thead>
<tr>
<th>Number of Full-Time Paid, Professional Staff</th>
<th>Number of Agencies</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 3</td>
<td>13</td>
<td>59.2</td>
</tr>
<tr>
<td>4 - 7</td>
<td>5</td>
<td>22.7</td>
</tr>
<tr>
<td>8 - 11</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>12 - 15</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>16 - 19</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>10 and above</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>Total</td>
<td>22</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 5 gives the distribution by the total number of part-time paid staff.

The distribution pattern was consistent with those previously discussed. One-half of the agencies hired a very small number of part-time staff.

How many of the sample hired part-time professional staff?

Fifty-five per cent did not have any part-time professional staff and none hired more than 4, as indicated in Table 6.
### TABLE 5
**TOTAL NUMBER OF PART-TIME PAID STAFF**

<table>
<thead>
<tr>
<th>Number of Part-time Paid Staff</th>
<th>Number of Agencies</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 - 2</td>
<td>11</td>
<td>50.0</td>
</tr>
<tr>
<td>3 - 5</td>
<td>2</td>
<td>9.1</td>
</tr>
<tr>
<td>6 - 8</td>
<td>5</td>
<td>22.7</td>
</tr>
<tr>
<td>9 - 11</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>12 - 14</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>15 - 17</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>18 - 20</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>21 and above</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22</strong></td>
<td><strong>100.</strong></td>
</tr>
</tbody>
</table>

About 82 per cent of the sample had no full-time non-paid staff; 59 per cent had no part-time non-paid staff. It could be concluded then that the majority of the sample had no non-paid staff or volunteer staff.

### TABLE 6
**NUMBER OF PART-TIME PAID PROFESSIONAL STAFF**

<table>
<thead>
<tr>
<th>Number of Part-time Paid Professional Staff</th>
<th>Number of Agencies</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>12</td>
<td>54.5</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>22.7</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>9.1</td>
</tr>
<tr>
<td>3</td>
<td>2</td>
<td>9.1</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>22</strong></td>
<td><strong>100.</strong></td>
</tr>
</tbody>
</table>
However, one agency reported having 200 part-time non-paid staff while another had as many as 525. Again, the term volunteer staff was defined differently by different agencies. A volunteer could mean anyone having worked at one time or another with the agency; he might not necessary be a regular worker. From observation, agencies offering services for primary prevention had a greater number of volunteer staff. University students on placement were classified as part-time non-paid staff.

The size of the operational budget was another indicator of the characteristics of the sample. Table 7 shows the distribution by the annual operating budget.

### TABLE 7

**ANNUAL OPERATING BUDGET**

<table>
<thead>
<tr>
<th>Thousands of Dollars</th>
<th>Number of Agencies</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 20</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>20 - 50</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>50 - 80</td>
<td>4</td>
<td>18.2</td>
</tr>
<tr>
<td>80 - 100</td>
<td>4</td>
<td>18.2</td>
</tr>
<tr>
<td>110 - 140</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>140 - 170</td>
<td>2</td>
<td>9.1</td>
</tr>
<tr>
<td>170 - 200</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>200 - 300</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>300 - 400</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>over 400</td>
<td>1</td>
<td>4.5</td>
</tr>
</tbody>
</table>

**Total** 22 100.
The annual operating budget of more than one-half of the sample did not exceed $110,000, although 19 per cent had budgets exceeding $300,000.

For cross-tabulation, the table was collapsed into 3 categories. The first consists of 8 agencies each with budgets within $80,000; the second 7 agencies with budgets between $80,000 and $140,000; and the third 7 agencies with budgets over $140,000.

The distribution by types of services is indicated in the following table.

**TABLE 8**

**TYPES OF SERVICES**

<table>
<thead>
<tr>
<th>Types of Services</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary Prevention</td>
<td>7</td>
<td>31.8</td>
</tr>
<tr>
<td>Secondary Prevention</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>Tertiary Prevention</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>Primary and Secondary</td>
<td>6</td>
<td>27.3</td>
</tr>
<tr>
<td>Primary and Tertiary</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>Secondary and Tertiary</td>
<td>2</td>
<td>9.1</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>9.1</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>22</td>
<td>100.</td>
</tr>
</tbody>
</table>

Nearly one-third of the sample offered services for primary prevention. There were 73 per cent of the sample

---

183 For explanations of the terms, primary, secondary and tertiary prevention, please refer to question 5 of the interview schedule (Appendix A); and Lydia Rapoport, "The Concept of Prevention in Social Work," *Social Work*, Vol. 6, No. 1, pp.3-12.
offering services for primary prevention or secondary prevention, or both. There was only a small number of agencies offering services which were for tertiary prevention either as a primary or secondary function.

In addition, and at the request of the U.C.S., the sample was divided into two types: professional and non-professional. The professional agency was defined as one which was classified by the U.C.S. as professional. The U.C.S. defined professional agency as one in which the key positions were occupied by those who had a professional degree from university. The professional degree was loosely defined as a degree in nursing, psychology or social work. Ten out of the sample of twenty-two were classified as professional agencies.

There was a statistically significant\textsuperscript{184} relationship between the types of agencies and the annual operating budget. As illustrated in the following contingency table, 60 per cent of the professional agencies had budgets between $80,000 and $140,000, while only 8 per cent of the non-professional agencies had such budgets. Chi square was used to indicate the significance of the relationship, and contingency coefficient (c) was used to describe the degree of relationship.

\textsuperscript{184}Statistical significance was defined as that the chi square value obtained on the relationship was larger than the critical value of chi square required for significance at 5 per cent level.
### TABLE 9

**TYPES OF AGENCIES BY ANNUAL OPERATING BUDGET**

<table>
<thead>
<tr>
<th>Types of Agencies</th>
<th>Budget Below $80,000</th>
<th>Between $80,000 and $140,000</th>
<th>Above $140,000</th>
<th>Raw Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional</td>
<td>2</td>
<td>6</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>20.0</td>
<td>60.0</td>
<td>20.0</td>
<td>45.5</td>
</tr>
<tr>
<td></td>
<td>25.0</td>
<td>85.0</td>
<td>28.6</td>
<td></td>
</tr>
<tr>
<td>Non-Professional</td>
<td>6</td>
<td>1</td>
<td>5</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>50.0</td>
<td>1</td>
<td>41.7</td>
<td>54.5</td>
</tr>
<tr>
<td></td>
<td>75.0</td>
<td>14.3</td>
<td>71.4</td>
<td></td>
</tr>
<tr>
<td><strong>Column Total</strong></td>
<td>8</td>
<td>7</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>36.4</td>
<td>31.8</td>
<td>31.8</td>
<td></td>
</tr>
</tbody>
</table>

Chi square = 6.73095  2 df  \( P < .05 \); c = 0.48

**Summary**

The majority of the sample had been in existence for over 20 years, and were members of the U.C.S. at the time of its incorporation. Among the sample there were only a few "young" agencies or "new" members.

The majority of the sample had a relatively small number of staff and annual operating budget. Most had no more than 10 full-time paid staff; no more than 3 full-time paid professional staff; no more than 5 part-time paid staff; and had no non-paid staff either full-time or part-time. About 60 per cent of the sample operated with an annual budget between $20,000 and $140,000. However, among the
sample, there were 4 which were relatively large in terms of staff complement and operating budget. The annual operating budgets of each of these big 4 exceeded $300,000; each had over 20 full-time paid staff, and over 18 part-time paid staff.

The services offered by the sample were mainly for primary prevention or secondary prevention, or both. Only a few offered services for tertiary prevention as a primary or secondary function.

Of the sample, ten were classified by the U.C.S. as professional agencies as distinguished from the non-professional, consisting of twelve.

There was a statistically significant relationship between the types of agencies and the annual operating budget.

Types of Evaluative Information

Evaluative information as used here was an abbreviated expression for the types of information kept in relation to service evaluation. What were the types of information kept by the sample in relation to service evaluation? The answer to this question should shed some light on what the sample attempted to evaluate.

The ten categories described in question six of the interview schedule were the types of information basic to the four types of evaluative research previously discussed.
effort, efficiency, effectiveness and benefit. Table 9 shows the distribution of the sample by the types of evaluative information.

There were some interesting phenomena in the distribution. Of the total responses, 63 per cent were affirmative response,\textsuperscript{185} whereas the negative response was only 16 per cent. Everyone of the sample kept information on "staff use of time" and "number of clients served." But there were very few who kept information on "benefits" and "efficiency" of services.

If keeping evaluative information was taken as doing evaluation, then most of the sample did not do benefit evaluation and efficiency evaluation.\textsuperscript{186} What they did was essentially effort evaluation, that is, accounting for the time and money spent. Slightly more than one-half attempted to do effectiveness evaluation, as indicated by the percentage of the affirmative response on the following categories which were essential to effectiveness evaluation: "staff's quality of work," "the effects of services on clients," "client's reactions" and "the need situations."

\textsuperscript{185}In this data analysis, the affirmative response was referred to the "yes" response; and the negative response the "no" response.

\textsuperscript{186}For explanation of the various types of evaluation, please refer to chapter two of this report.
### TABLE 10

**Distribution by Types of Evaluative Information Used by Agency**

<table>
<thead>
<tr>
<th>Types of Evaluative Information</th>
<th>Yes</th>
<th>No</th>
<th>Not Sure</th>
<th>Not Applicable</th>
<th>No Answer</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Staff use of time</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>(2) The cost of different services</td>
<td>18</td>
<td>3</td>
<td></td>
<td></td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>(3) Number of clients served</td>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>(4) Characteristics of clients served</td>
<td>17</td>
<td>4</td>
<td></td>
<td></td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>(5) Staff's quality of work</td>
<td>14</td>
<td>2</td>
<td></td>
<td></td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>(6) The efficiency of service delivery (i.e., quality of services over relative cost)</td>
<td>3</td>
<td>6</td>
<td></td>
<td></td>
<td>4</td>
<td>22</td>
</tr>
<tr>
<td>(7) The effects of the services on the clients</td>
<td>11</td>
<td>4</td>
<td></td>
<td></td>
<td>1</td>
<td>22</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Types of Evaluative Information</th>
<th>Yes</th>
<th>No</th>
<th>Not Sure</th>
<th>Not Applicable</th>
<th>No Answer</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(8) Client's reactions to the services (e.g. client's satisfaction, attitudes and suggestions)</td>
<td>11 (50.0)</td>
<td>6 (27.3)</td>
<td>3 (13.6)</td>
<td></td>
<td>2 (9.1)</td>
<td>22</td>
</tr>
<tr>
<td>(9) The benefits (usually expressed in monetary terms) the services have brought to the clients</td>
<td>6 (27.3)</td>
<td>8 (36.4)</td>
<td>3 (13.6)</td>
<td></td>
<td>2 (9.1)</td>
<td>22</td>
</tr>
<tr>
<td>(10) The need situations towards which the services are directed</td>
<td>14 (63.6)</td>
<td>3 (13.6)</td>
<td>4 (18.2)</td>
<td></td>
<td>1 (4.5)</td>
<td>22</td>
</tr>
<tr>
<td><strong>Column Total</strong></td>
<td><strong>138</strong></td>
<td><strong>36</strong></td>
<td><strong>23</strong></td>
<td><strong>10</strong></td>
<td><strong>13</strong></td>
<td><strong>220</strong></td>
</tr>
</tbody>
</table>

*(Figures in bracket represents per cent of row total).*
The responding agencies or the sample had been for two years on the functional budgeting program in which they were assisted to carry out financial budgeting and accounting according to procedures set out by the Community Funds and Councils of Canada.\(^{187}\) This fact should well explain that each of them kept some types of statistics.

To study variable relationship, the various types of evaluative information were cross-tabulated by (a) agency length of existence, (b) agency annual operating budget, and (c) types of agencies.

The cross-tabulation was to study the relationship between each of the eight\(^{188}\) types of evaluative information and each of the three variables. Only the affirmative and negative responses of each of the eight types of evaluative information were included in the cross-tabulation.

It was found that none of the relationships cross-tabulated was statistically significant.

Summary

Each agency in the sample kept some sort of evaluative information on their services. Most of them did not try to evaluate service efficiency and benefits. About one-half of them attempted to evaluate service effectiveness.

\(^{187}\) Community Funds and Councils of Canada, Functional Budgeting.

\(^{188}\) The remaining 2 types with 100 per cent affirmative response were excluded.
Practically all of them did, in one way or the other evaluate their effort.

There were no statistically significant relationships between each of the various types of evaluative information and (a) annual operating budget, (b) types of agencies, or (c) agency length of existence.

The Methods in Service Evaluation

Questions seven to twelve in the interview schedule were intended to yield information on the six aspects of the methods of evaluative practices as previously defined. Each of these six aspects was discussed separately under a sub-heading.

Methods Used

Provided in question seven were the various methods for evaluating service effectiveness. Categories one to three of question seven were the types of statistics basic to different types of evaluation. Categories five and six were the methods used during treatment; and seven and eight after treatment. The experimental methods in category four could be used during or after treatment or both. The methods to evaluate the need situations were described in the ninth and tenth categories. The respondents were asked to indicate which of these methods they had used. Table 11 indicated the distribution by methods used.
<table>
<thead>
<tr>
<th>Methods</th>
<th>Yes</th>
<th>No</th>
<th>Not Sure</th>
<th>Not Applicable</th>
<th>No Answer</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Statistics on number of clients served</td>
<td>22</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>(2) Statistics on characteristics of clients served</td>
<td>18</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td>22</td>
</tr>
<tr>
<td>(3) Record of time spent in delivery of service</td>
<td>17</td>
<td>2</td>
<td></td>
<td>2</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td>(4) Experimental or quasi-experimental designs (e.g., use of control group, 'before-after' testing)</td>
<td>2</td>
<td>13</td>
<td></td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>(5) Staff observation of client changes</td>
<td>15</td>
<td>2</td>
<td></td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>(6) Systematic collection of client responses by questionnaire, interview schedule, or a scale during treatment period</td>
<td>5</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Methods</th>
<th>Yes</th>
<th>No</th>
<th>Not Sure</th>
<th>Not Applicable</th>
<th>No Answer</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(7) Selective or casual collection of client reports (i.e. letters of appreciation, taking the &quot;successful&quot; cases)</td>
<td>11</td>
<td>6</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>(50.0)</td>
<td>(27.3)</td>
<td>(4.5)</td>
<td>(13.6)</td>
<td>(4.5)</td>
<td></td>
</tr>
<tr>
<td>(8) Systematic follow up of clients after termination of services</td>
<td>9</td>
<td>6</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>(40.9)</td>
<td>(27.3)</td>
<td>(9.1)</td>
<td>(13.6)</td>
<td>(9.1)</td>
<td></td>
</tr>
<tr>
<td>(9) Periodic systematic survey and definition of need</td>
<td>9</td>
<td>8</td>
<td>2</td>
<td>3</td>
<td></td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>(40.9)</td>
<td>(36.4)</td>
<td>(9.1)</td>
<td>(13.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10) Make some sort of observation or &quot;guess&quot; on need situation periodically</td>
<td>10</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>22</td>
</tr>
<tr>
<td></td>
<td>(45.5)</td>
<td>(31.8)</td>
<td>(4.5)</td>
<td>(9.1)</td>
<td>(9.1)</td>
<td></td>
</tr>
<tr>
<td>Column Total</td>
<td>118</td>
<td>60</td>
<td>12</td>
<td>19</td>
<td>11</td>
<td>220</td>
</tr>
<tr>
<td></td>
<td>(53.6)</td>
<td>(27.3)</td>
<td>(5.5)</td>
<td>(8.6)</td>
<td>(5.0)</td>
<td></td>
</tr>
</tbody>
</table>
Shown in the table was that a vast majority of the sample kept the first three types of statistics: the number of clients, the characteristics of clients and the time spent. About 68 per cent carried out observation on client changes; one-half of the sample collected casual reports of success. The percentages of frequency for the rest of the categories were less than 50. All the 3 methods with the adjective "systematic," had low percentage of affirmative response. Experimental methods and the "systematic collection of clients' responses," had the lowest percentage of affirmative response and the highest percentage of negative response; that is, it was positive that these two types of methods were the least common in usage.

Although 68 per cent of the sample carried out observation of client changes, systematically to collect clients' reactions was uncommon --- about 23 per cent of the total sample.

Using the same method as described in the last section, each of the various methods used was cross-tabulated with (a) annual operating budget, (b) agency length of existence, and (c) types of agencies. There was only one significant relationship found. As indicated in the contingency table of Appendix D, none of the agencies which had budget below $80,000 used the method, "systematic follow up of clients," which was used by all agencies which had budget exceeding $140,000. The indication was that
the larger the budget the more likely it was that the agency used this method.

In another level of cross-tabulation in which all types of responses were included, it was detected that the non-professional agencies tended to reply "no," "not sure," "not applicable" or make no answer. In other words, there was a higher percentage of the professional than the non-professional agencies which used each method, especially methods numbered 3, 4, 5, 7 and 8 where the frequency percentage of the affirmative response was much higher.

Methods Preferred

Question eight consisted of the same categories as in question seven. It was intended to find out the methods which the sample preferred to use instead of those used for evaluating their service effectiveness. The distribution by methods preferred was shown in Table 12.

The methods preferred by more than one-half of the sample were the first three types of statistics which also were those with highest frequency in usage. The distributions on methods used and on methods preferred supported the interpretation on the distribution on the types evaluative information that most of the sample were still at their level of evaluating effort.

Of the total, there was only 29 per cent of affirmative response in methods preferred, whereas in methods used the percentage was about 54 per cent. There
TABLE 12
DISTRIBUTION BY THE METHODS THE SAMPLE PREFERRED

<table>
<thead>
<tr>
<th>Methods</th>
<th>Yes (%)</th>
<th>No (%)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Statistics on number of clients served</td>
<td>15 (68.2)</td>
<td>7 (31.8)</td>
<td>22</td>
</tr>
<tr>
<td>(2) Statistics on characteristics of clients served</td>
<td>13 (59.1)</td>
<td>9 (40.9)</td>
<td>22</td>
</tr>
<tr>
<td>(3) Record of time spent in delivery of service</td>
<td>13 (59.1)</td>
<td>9 (40.9)</td>
<td>22</td>
</tr>
<tr>
<td>(4) Experimental or quasi-experimental designs (e.g., use of control group, 'before-after' testing)</td>
<td>3 (13.6)</td>
<td>19 (86.4)</td>
<td>22</td>
</tr>
<tr>
<td>(5) Staff observation of client changes</td>
<td>8 (36.4)</td>
<td>14 (63.6)</td>
<td>22</td>
</tr>
<tr>
<td>(6) Systematic collection of client responses by questionnaire, interview schedule, or a scale during treatment period</td>
<td>5 (22.7)</td>
<td>17 (77.3)</td>
<td>22</td>
</tr>
<tr>
<td>(7) Selective or casual collection of client reports (i.e. letters of appreciation, taking the &quot;successful&quot; cases)</td>
<td>5 (22.7)</td>
<td>17 (77.3)</td>
<td>22</td>
</tr>
<tr>
<td>(8) Systematic follow up of clients after termination of services</td>
<td>9 (40.9)</td>
<td>13 (59.1)</td>
<td>22</td>
</tr>
<tr>
<td>(9) Periodic systematic survey and definition of need</td>
<td>9 (40.9)</td>
<td>13 (59.1)</td>
<td>22</td>
</tr>
<tr>
<td>(10) Make some sort of observation or &quot;guess&quot; on need situation periodically</td>
<td>5 (22.7)</td>
<td>17 (77.3)</td>
<td>22</td>
</tr>
<tr>
<td>Column Total</td>
<td>85 (38.6)</td>
<td>135 (61.4)</td>
<td>220</td>
</tr>
</tbody>
</table>
were a few possible explanations. One was that some did not prefer evaluation in the first place. Another reason could be that they did not prefer any method at all, because each took up their time and effort. The third reason could be that after using, they discovered that the methods were not good enough for their purposes, yet they could not think of better ones as there were no methods suggested under "other" and "comments."

As reflected in Tables 11 and 12, the identical numbers of agencies used and preferred the methods numbered 6, 8 and 9, the more refined and systematic ways. Only about one-half the agencies which used the methods numbered 5, 7 and 10 preferred them. Probably those which did not show preference thought these more refined methods were not good or objective enough.

There were no significant relationships between each of the methods preferred and (a) annual operating budget, (b) types of agencies, and (c) agency length of existence.

In another level of cross-tabulation in which the test of significance of the association between variables was not applied, it was found that there was a higher percentage of the professional than the non-professional in their preference for almost everyone of the listed methods, especially the more refined and systematic methods. The difference in terms of percentage between the professional and non-professional agencies seemed to testify that some
of the professional agencies would prefer to continue progressing on the more systematic and sophisticated methods. Although the same number of agencies used and preferred the more refined and systematic methods, the cross-tabulation indicated that there was a higher number of professional agencies which preferred than used three of the four relatively more refined methods. These three methods were: experimental methods, systematic collection of client responses, and periodic systematic survey and definition of need. It was reasonable to assume then that those who had used them preferred them. Logically they must have found these methods fruitful, effective or necessary. Another indication was that more professional agencies would like to try out the more sophisticated methods, except the method of systematic follow up of clients.

Specification of Objectives

Vague goals were one of the major hurdles in evaluation of social services. As pointed out, the objectives of many social agencies were global and unspecific.\textsuperscript{189} Specifications of the agency objectives in measurable terms was the first and foremost step in the evaluation of effectiveness, efficiency and benefit. Therefore a question was set to find out the level of the evaluative practices of the sample.

\textsuperscript{189}Rossi, "Testing for Success and Failure in Social Action," p.18.
How many of the sample did specify their agency objectives in measurable terms? Almost one-half (10) out of the sample of 22 reported they did so, 6 reported they did not do so, 2 were not sure, and 4 thought that specifying objectives in measurable terms was not applicable to their agencies. In cross-tabulation, 50 per cent of the professional as compared to about 42 per cent of the non-professional indicated that they did specify objectives in measurable terms. Though the difference was not statistically significant, the findings were consistent with those in the foregoing sections.

However, how successful did they feel they were in specifying objectives in measurable terms? Of the 50 per cent or 11 who responded, \textsuperscript{190} one felt "a great deal," five felt "fairly much," four "to some extent." One was not sure. The mean value \textsuperscript{191} was 3.7. In other words, as a whole they felt "fairly much" successful. Those who had done so, found it successful. This subjective feeling of success should give a boost to those who had not tried out.

\textsuperscript{190}Eleven instead of 10 responded to the second question.

\textsuperscript{191}In computing mean value for rank ordering, "not sure," "not applicable," and no answers were not assigned any value. Each degree of the five-point scale was arbitrarily assigned a value ranging from 1 to 5 corresponding to from "not at all" to "a good deal." The researcher was aware that in assigning the values, not only the scale was transformed into an interval scale but also the distance between each degree on the scale was arbitrarily determined. Further, there was no testing of the significance of the difference between the mean values for rank ordering.
Adequacy of the Methods Used

How adequate did the respondents consider the methods used for evaluating their services? The distribution on the five degrees of adequacy was shown in the table below.

<table>
<thead>
<tr>
<th>Degrees of Adequacy</th>
<th>Number</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A Great Deal</td>
<td>5</td>
<td>22.7</td>
</tr>
<tr>
<td>Fairly Much</td>
<td>10</td>
<td>45.5</td>
</tr>
<tr>
<td>To Some Degree</td>
<td>6</td>
<td>27.3</td>
</tr>
<tr>
<td>Comparatively Little</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>22</td>
<td><strong>100.0</strong></td>
</tr>
</tbody>
</table>

The mean value was 3.9, indicating that as a whole they felt "fairly much" adequate.

There was no statistical justification to say that there was any relationship between the ratings of adequacy and (a) operating budget, (b) agency length of existence, and (c) types of services.

In spite of the above, the mean value of the ratings of the professional agencies was 4.2 as compared to 3.6 of the non-professional agencies. It could mean that the professional agencies were more satisfied with the methods used and that the non-professional agencies were less satisfied and sought other methods.
Preference for Evaluators

As discussed in the literature survey, the source of the evaluator is an important consideration in evaluation research. Therefore, it was important to find whom the sample preferred to conduct their service evaluation. Listed in question eleven were five “sources” of evaluators. The respondents were requested to rate each of them in terms of the degree of their preference. Table 14 shows the distribution by their preference for evaluators.

The least preferred were those where the evaluation of their services did not involve the agency personnel. In fact, they highly preferred to conduct their evaluation by themselves. Eighteen per cent of the respondents did not indicate any preference; they made no answer. Could it mean that they basically did not prefer evaluation at all?

The professional agencies gave a higher rating on the first four sources of evaluators than did the non-professional. On the fifth, the rating of the non-professional was slightly higher than that of the professional. One implication was that the professional agencies had more faith in the first four types of evaluators than the non-professional agencies, which had more faith in the team effort with their involvement. Possibly, these agencies recognized the inability of evaluation internally, and felt that the funding source itself was not acceptable.

There was a statistically significant relationship
### TABLE 14

**DISTRIBUTION BY RESPONDENTS' PREFERENCE FOR SOURCES OF EVALUATORS**

<table>
<thead>
<tr>
<th>Sources of Evaluators</th>
<th>A Great Deal</th>
<th>Fairly Much</th>
<th>To Some Degree</th>
<th>Comparatively Little</th>
<th>Not At All</th>
<th>Not Sure</th>
<th>No Answer</th>
<th>Row Total N=22</th>
<th>Mean Value</th>
<th>Rank</th>
<th>Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Your agency itself</td>
<td>10</td>
<td>4</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>22</td>
<td></td>
<td>4.3</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Some experts outside the agency</td>
<td>2</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>6</td>
<td></td>
<td>22</td>
<td></td>
<td>2.7</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>The funding source(s)</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>22</td>
<td>3.1</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>The agency itself with outside experts as advisors</td>
<td>7</td>
<td>5</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>22</td>
<td>3.8</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>A joint team of the agency staff, outside experts and the funding source(s)</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>22</td>
<td></td>
<td>4.1</td>
<td>2</td>
<td></td>
</tr>
</tbody>
</table>
between the preference for "some experts outside the agency" and the annual operating budget. As shown in the contingency table of Appendix E those agencies which had an annual operating budget between $80,000 and $140,000 tended to rate either "a great deal" or "not at all" in terms of degrees of preference for "some experts outside the agency." Agencies having an annual operating budget above or below this stated amount tended to rate "to some extent" or "comparatively little."

Difficulties and Limitations

What were the common difficulties and limitations experienced by the sample in service evaluation? Some knowledge of the limitations and difficulties experienced by the sample was essential in trying to help improve their service evaluation. Provided in the list were fifteen limitations and difficulties for the respondents to rate. The higher the value number represents the greater the difficulty.

In the frequency distribution in Table 15, the five highest terms of mean value, that is the five most difficult, were: (a) sparing staff and time; (b) finding proper scientific ways; (c) lacking financial resources; (d) specifying service goals; and (e) meeting complications caused by various stages of program development. The (a) and (c) related to funding whereas the other three related to the techniques of evaluation. It was clear then, that
### TABLE 15

DISTRIBUTION BY DIFFICULTIES AND LIMITATIONS THE SAMPLE EXPERIENCED IN SERVICE EVALUATION

<table>
<thead>
<tr>
<th>Difficulties and Limitations</th>
<th>Great</th>
<th>Fairly</th>
<th>Some</th>
<th>To Deal</th>
<th>Much</th>
<th>Degree</th>
<th>Little</th>
<th>All</th>
<th>Sure</th>
<th>cable</th>
<th>Answer</th>
<th>Mean</th>
<th>Value</th>
<th>Rank Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Specifying Service goals in measurable terms</td>
<td>2</td>
<td>4</td>
<td>7</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td></td>
<td>3.2</td>
</tr>
<tr>
<td>(2) Finding proper scientific ways of yielding reliable data</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>1</td>
<td>3.6</td>
</tr>
<tr>
<td>(3) Determining what information is required</td>
<td>2</td>
<td>5</td>
<td>9</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>8</td>
<td>8</td>
<td>2.6</td>
</tr>
<tr>
<td>(4) Getting feedback from clients</td>
<td>1</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>3</td>
<td>2.9</td>
</tr>
<tr>
<td>(5) Meeting resistance from the clients</td>
<td>2</td>
<td>4</td>
<td>8</td>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>4</td>
<td>2.2</td>
</tr>
<tr>
<td>(6) Sparing the staff and time</td>
<td>8</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
<td>5</td>
<td>3.9</td>
</tr>
<tr>
<td>(7) Preventing hindrance to the smooth functioning of the services</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>1</td>
<td>4</td>
<td>2.5</td>
<td>10</td>
<td></td>
<td>(continued)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Difficulties and Limitations</td>
<td>A Great Deal</td>
<td>Fairly Much</td>
<td>To Some Degree</td>
<td>Little</td>
<td>Comparatively At All</td>
<td>Not At All</td>
<td>Sure Cable</td>
<td>No Application</td>
<td>No Answer</td>
<td>Mean</td>
<td>Rank Order</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>----------------</td>
<td>--------</td>
<td>---------------------</td>
<td>------------</td>
<td>-----------</td>
<td>---------------</td>
<td>-----------</td>
<td>------</td>
<td>------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) Lacking financial resources</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>3.4</td>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) Lacking understanding and enthusiasm among the staff</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>1</td>
<td>5</td>
<td>1.6</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10) Meeting resistance from the staff</td>
<td>5</td>
<td>11</td>
<td>1</td>
<td>5</td>
<td>1.3</td>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11) Lacking encouragement from the board of directors/management</td>
<td>1</td>
<td>4</td>
<td>12</td>
<td>5</td>
<td>1.4</td>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12) Having much value judgement or subjectivity in evaluation</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>7</td>
<td>2.7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(13) Having nobody or nowhere to turn to for consultation</td>
<td>1</td>
<td>4</td>
<td>5</td>
<td>7</td>
<td>1</td>
<td>4</td>
<td>1.9</td>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(14) Having to assess the effects of the intervening variables</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>6</td>
<td>2.7</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(15) Meeting complications caused by various stages of program service development</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>6</td>
<td>3.1</td>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
the greatest problems for them were in fact money and knowledge and skill of evaluation research.

In a similar way, it could be concluded that staff understanding and resistance, encouragement from the board, and availability of persons for consultation were the least difficult problems or limitations for them.

The percentage of the total number of "no answers" was as high as 23. It could mean that some of the sample (a) had not experienced the listed difficulties or limitations, or (b) were tired of checking the long list of 15 categories.

Summary

The experimental methods and the other three comparatively more systematic and refined methods were the least common in usage, and also the least preferred. These more refined methods were essential to the evaluation of effectiveness, efficiency and benefits.

The methods more preferred and used most commonly were the accounting methods, that is, keeping statistics on the number of clients served, characteristics of clients, and time spent in delivering service. In other words, most of the sample were still at the level of evaluating effort and preferred this for the time being.

On the whole, some of the sample used the methods listed but did not prefer them. A higher percentage of the professional agencies than the non-professional used and
preferred the methods listed, especially the more refined ones. Among the professional agencies, there was a higher number who preferred than used the more refined methods.

Ten or 45 per cent of the sample reported they did specify their objectives in measurable terms. A higher percentage of the professional than the non-professional agencies did specify their objectives. Those who did specify their objectives generally felt "fairly much" successful in their endeavour.

How adequate did the respondents consider the methods used for evaluating the services? As a whole, they considered the methods used "fairly much" adequate for their purposes.

As to preference for evaluators, the sample highly preferred to conduct their evaluation by themselves or in conjunction with outside experts and the funding source. They preferred least the evaluation of the services without their agency taking part in it, such as by some experts outside the agency or the funding source.

The greatest limitations or difficulties experienced by the sample were those related to the knowledge and skill of evaluation research, and those related to financial resources. They had least difficulties with their staff understanding resistance, encouragement from the board of directors and availability of experts for consultation.
The Purposes for the Use of Service Evaluation

The purposes of evaluation have important implications for the structure and outcome of evaluation. As concluded by Weiss (1971) the first and foremost step for the evaluator to do is to find out the purposes for having evaluation. Question thirteen was intended to find out what the main reasons were for the sample to evaluate its services. Provided in the list were fifteen reasons for agencies to evaluate their services. Each of these reasons was rated in terms of the extent it applies to their agencies. Many covert reasons for evaluation were not included. Table 16 gives the distribution of the sample by their ratings on the reasons provided.

As a whole all the reasons were rated high. The first six reasons which had the highest mean value of ratings were reasons numbered 7, 10, 11, 13, 15 and 12. These reasons were the common, stated reasons for service evaluation, such as "to find a more efficient way of service delivery" and "to determine how much the clients have benefited from the services." Though these were the highly rated purposes for evaluation, the methods directly related to attaining these purposes were not commonly used or preferred. Similarly, more systematic and refined methods were essential to study the needs situation, relevancy of agency goals and staff quality of work which were the other four highly rated purposes. From the findings in the types
<table>
<thead>
<tr>
<th>Reasons</th>
<th>A Great Deal</th>
<th>Fairly Much</th>
<th>To Some Degree</th>
<th>Little</th>
<th>Comparatively At</th>
<th>Not All</th>
<th>Sure</th>
<th>Not Applicable</th>
<th>No Answer</th>
<th>Value</th>
<th>Mean</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) to satisfy the board of directors</td>
<td>2</td>
<td>4</td>
<td>11</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.0</td>
<td>13</td>
</tr>
<tr>
<td>(2) to satisfy the funding source</td>
<td>6</td>
<td>6</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.8</td>
<td>8</td>
</tr>
<tr>
<td>(3) to satisfy the public and community at large</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>4</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.1</td>
<td>12</td>
</tr>
<tr>
<td>(4) to evaluate the work of the staff</td>
<td>9</td>
<td>4</td>
<td>6</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3.9</td>
</tr>
<tr>
<td>(5) to let the public know how the agency is doing</td>
<td>5</td>
<td>5</td>
<td>9</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3.5</td>
<td>10</td>
</tr>
<tr>
<td>(6) to get some testimonials of success</td>
<td>3</td>
<td>1</td>
<td>8</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2.8</td>
<td>14</td>
</tr>
<tr>
<td>(7) for planning and service improvement</td>
<td>19</td>
<td>2</td>
<td></td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4.8</td>
<td>1</td>
</tr>
<tr>
<td>(8) to get some feedback for records</td>
<td>5</td>
<td>3</td>
<td>9</td>
<td>3</td>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1</td>
<td>3.3</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Reasons</th>
<th>A Great Deal</th>
<th>Fairly Much</th>
<th>To Some Degree</th>
<th>Comparatively Little</th>
<th>Not At All</th>
<th>Sure</th>
<th>Not Applicable</th>
<th>No Answer</th>
<th>Value</th>
<th>Mean</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>(9) to show how effort and time are spent</td>
<td>6</td>
<td>5</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3.7</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>(10) to ascertain the progress towards the agency goals</td>
<td>12</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4.5</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11) to find a more efficient way of service delivery</td>
<td>14</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>4.4</td>
<td>4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(12) to determine to what extent the social needs have been met</td>
<td>10</td>
<td>5</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>3.9</td>
<td>6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(13) to determine how much the clients have benefited from the services</td>
<td>12</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>4.3</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(14) to keep up with the trend</td>
<td>4</td>
<td>5</td>
<td>3</td>
<td>7</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2.5</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(15) to determine if agency goals should be changed or modified</td>
<td>15</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td>4.5</td>
<td>2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
of evaluative information kept, the methods used and the methods preferred, it was clear that the sample had yet to achieve their noble or "high ideal" purpose for evaluation.

The set of reasons ranking immediately after the six were: "to satisfy the funding source," "to let the public know how the agency is doing" and "to show how effort and time are spent." These reasons were very much related to the public relations aspects of the agency function. Accounting for time and effort spent was the most popular method used and preferred.

The reason with the second lowest ratings, "to get some testimonials of success," which was closely associated with the public relations function, was rather common as a method in usage, though not quite preferred. There was a gap between ideal and practice, between what things are and what things should be. Perhaps some respondents answered what the reasons should be instead of what they were.

How in fact did the sample use the results of service evaluation? In the distribution shown in Table 17, "brought to the board of directors" and "reported to the U.C.S." were the most common. The use for planning and staff discussion were the second most common. It was clear then that the most common ways in the use of evaluation results were in fact the second most common reasons for evaluation; whereas the second most common ways of using evaluation were the most common reasons for evaluation.
### TABLE 17

DISTRIBUTION BY THE USE OF EVALUATION RESULTS

<table>
<thead>
<tr>
<th>Ways of Uses</th>
<th>Number of Agencies (N=22)</th>
<th>Percentage</th>
<th>Rank Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) filed in record for director's reference</td>
<td>15</td>
<td>68.2</td>
<td>5</td>
</tr>
<tr>
<td>(2) brought to the administrative level for study</td>
<td>15</td>
<td>68.2</td>
<td>5</td>
</tr>
<tr>
<td>(3) brought to the planning level for study</td>
<td>16</td>
<td>72.7</td>
<td>3</td>
</tr>
<tr>
<td>(4) brought to staff meeting for discussion</td>
<td>16</td>
<td>72.7</td>
<td>3</td>
</tr>
<tr>
<td>(5) brought to the board of directors</td>
<td>18</td>
<td>81.8</td>
<td>1</td>
</tr>
<tr>
<td>(6) filed in record for auditor's reference</td>
<td>4</td>
<td>18.2</td>
<td>8</td>
</tr>
<tr>
<td>(7) reported to the U.C.S.</td>
<td>18</td>
<td>81.8</td>
<td>1</td>
</tr>
<tr>
<td>(8) published in newsletter, local papers, etc.</td>
<td>7</td>
<td>31.8</td>
<td>7</td>
</tr>
</tbody>
</table>
The findings did confirm that the important purposes were for planning and service improvement and for satisfying the funding source and the public at large. It could perhaps be inferred that the U.C.S. had been the strong force in "motivating" and persuading its member agencies to evaluate their services.

Summary

The major purposes for the sample to evaluate their services were for (a) planning and service improvement; and (b) satisfying the public and the funding source. In this connection, it was clear that the U.C.S. had been an influential force for its member agencies to evaluate.

The Value of Service Evaluation

The value of service evaluation held by the respondents was expressed by their ratings on (a) the usefulness of the existing service evaluation to their agencies; (b) the importance of service evaluation as a whole; and (c) the direction of change of the existing service evaluation.

How useful was the existing service evaluation to the agency? Table 18 shows the distribution of the respondents' ratings on the various degrees of usefulness.
TABLE 18
DISTRIBUTION BY RESPONDENTS' RATINGS ON DEGREES OF USEFULNESS OF SERVICE EVALUATION

<table>
<thead>
<tr>
<th>Degrees of Usefulness</th>
<th>Number of Agencies (N = 22)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A great deal</td>
<td>7</td>
<td>31.8</td>
</tr>
<tr>
<td>Fairly much</td>
<td>5</td>
<td>22.7</td>
</tr>
<tr>
<td>To some extent</td>
<td>7</td>
<td>31.8</td>
</tr>
<tr>
<td>Comparatively little</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>Not at all</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not sure</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>No answers</td>
<td>1</td>
<td>4.5</td>
</tr>
</tbody>
</table>

More than one-half of the sample felt that the existing service evaluation was "fairly much" or "a great deal" useful.

How did the respondents rate the importance of service evaluation in general? Table 19 gives the distribution by respondents' ratings on the various degree of importance. The mean value was 4.21 that means as a whole the respondents regarded service evaluation as very important.

Table 20 presents the comparison of the mean values of the respondents' ratings on: (a) the importance of service evaluation in general; (b) the usefulness of the existing service evaluation to the agency; and (c) the adequacy of the existing methods used in service evaluation. The mean value of importance of evaluation was higher than that of either usefulness or adequacy. The findings seemed
TABLE 19

DISTRIBUTION BY RESPONDENTS' RATINGS ON VARIOUS DEGREES OF IMPORTANCE OF SERVICE EVALUATION

<table>
<thead>
<tr>
<th>Degrees of Importance</th>
<th>Number of Agencies (N=22)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A great deal</td>
<td>12</td>
<td>54.5</td>
</tr>
<tr>
<td>Fairly much</td>
<td>4</td>
<td>18.2</td>
</tr>
<tr>
<td>To some extent</td>
<td>5</td>
<td>22.7</td>
</tr>
<tr>
<td>Comparatively little</td>
<td>1</td>
<td>4.5</td>
</tr>
</tbody>
</table>

to suggest that existing service evaluation procedures needed refinement before they could be viewed as important as evaluation might be; that evaluation probably became less ideal in practice.

TABLE 20

COMPARISON OF THE MEAN VALUES OF RATINGS ON IMPORTANCE, USEFULNESS AND ADEQUACY OF EVALUATIONS

<table>
<thead>
<tr>
<th>Category</th>
<th>Mean Value</th>
<th>Rank Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Importance</td>
<td>4.2</td>
<td>1</td>
</tr>
<tr>
<td>Usefulness</td>
<td>3.9</td>
<td>2</td>
</tr>
<tr>
<td>Adequacy</td>
<td>3.9</td>
<td>2</td>
</tr>
</tbody>
</table>

What would the respondents want to see happen to the existing service evaluation? Table 21 shows the distribution by the direction of change desired by the respondents. About 64 per cent of the sample wanted to
improve the existing service evaluation whenever possible. About 14 per cent wanted to improve immediately. No agency wanted to simplify or do away with the existing service evaluation. The findings then documented that at least 77 per cent wanted to improve their existing service evaluation.

**Table 21**

**DISTRIBUTION BY RESPONDENTS' RATINGS ON DIRECTION OF CHANGE**

<table>
<thead>
<tr>
<th>Directions of Change</th>
<th>Number of Agencies (N=22)</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be simplified</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Be done away with</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>Stay as it is for the time being</td>
<td>1</td>
<td>4.5</td>
</tr>
<tr>
<td>Be improved whenever possible</td>
<td>14</td>
<td>63.6</td>
</tr>
<tr>
<td>Be improved immediately</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>Not sure</td>
<td>3</td>
<td>13.6</td>
</tr>
<tr>
<td>No answer</td>
<td>1</td>
<td>4.5</td>
</tr>
</tbody>
</table>

**Summary**

The sample held high value of service evaluation. They considered that (a) the existing service evaluation was useful, and (b) service evaluation in general was very important. They indicated that they would like to improve their service evaluation as soon as possible.
The Kinds of Help Needed

The majority of the sample wanted to improve the existing service evaluation. But what kinds of help did the respondents think their agencies would need to make the improvement possible? Presented in Table 22 was the distribution of the sample on the kind of help needed. What the sample needed most was "some advice on consultation on service evaluation" as well as guidance from some experts. The findings would indicate that the sample had been groping to find ways to improve their service evaluation; that they had been doing by trial-and-error methods without knowledge or proper guidance; that they waited for some guidance to indicate improvement in their evaluation. The findings clearly demonstrated a need for know-how in evaluation research.

There was no clear indication that the size of the annual operating budget had particular relationship with a certain kind of help needed. In fact, the agencies with an annual operating budget between $300,000 and $400,000 indicated what they needed most was "more funds." Those with annual operating budget between $200,000 and $300,000 indicated that what they needed most was more trained staff. Those with smaller budgets wanted expert guidance and some training for their staff.

In the same fashion, the researcher was not able to detect any interpretative relationship between the staff complement and the kinds of help needed.
<table>
<thead>
<tr>
<th>Kinds of Help</th>
<th>A Great Deal</th>
<th>Fairly Much</th>
<th>Some Degree</th>
<th>Little</th>
<th>At Not All</th>
<th>Sure Not</th>
<th>No Answer</th>
<th>Mean Value</th>
<th>Rank Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) More funds</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>2</td>
<td>5</td>
<td>3.9</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Additional trained staff</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>2</td>
<td>1</td>
<td>3.6</td>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Outside expertise to guide and direct your service evaluation</td>
<td>5</td>
<td>9</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>4.0</td>
<td>2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Some training or orientation for your existing staff on service evaluation</td>
<td>7</td>
<td>6</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>3.9</td>
<td>3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Some advice or consultation on service evaluation</td>
<td>5</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4.1</td>
<td>1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Summary

The sample indicated that they needed most some consultation and expert guidance to improve their service evaluation, over more funds or additional staff. The findings indirectly illustrated their need for better know-how of evaluation. There was no interpretative relationship between the kind of help needed and (a) the annual operating budget of the agency, and (b) the size of the staff complement.

The analysis of data was centered on the six aspects relative to the research focus. The major findings were derived mainly from analysis of the frequency distribution and had been summarized separately under each section.

Cross-tabulations yielded few statistically significant relationships. However, due to the small sample size, some phenomena and relationships which did not reach the statistical significance level were reported and interpreted as they appeared to offer reasonable inferences and explanations. Statistical descriptions of some of these relationships used for interpretation were listed in Appendix F for reference. Various statistical measures were used to indicate the level of significance and the degrees of association of these relationships.

The implications of the findings were discussed in the following chapter.
CHAPTER V

CONCLUSION AND RECOMMENDATIONS

The purpose of this research project was to look at certain aspects of the evaluative practices of the U.C.S. member agencies with a view to (a) develop some ideas for the improvement of service evaluation of the U.C.S. member agencies; and (b) formulate specific problems of service evaluation for more precise investigation.

An extensive survey of the literature was conducted to gain familiarity with various aspects of evaluation research in social services. Prior to the formulation of the research focus, an experimental survey was also carried out.

The entire population, that is, all the twenty-seven member agencies of the U.C.S. at the time of the research, were studied as a sample. Data was obtained from twenty-two agencies. It was collected through a structured interview schedule according to a specific procedure.

The data collected was centered on describing the following research focus:

(a) the types of information kept in relation to service evaluation;
(b) the methods used in service evaluation;
(c) the purposes for the use of service evaluation;
(d) the value of service evaluation; and
(d) the kind of help needed for improvement of service evaluation.

The Major Findings

Summarized below are the major findings.

1. The majority of the agencies in the sample was relatively small in terms of staff complement and annual operating budget. The majority had been in existence for over twenty years and were chartered members of the U.C.S.

2. All agencies kept some types of evaluative information, primarily the statistics that account for time and money spent. About one-half of the sample kept information related to service effectiveness. Most did not keep information on service efficiency and service benefits.

3. The methods of service evaluation which were used by all of the sample and preferred by most of them were the accounting methods. The more sophisticated and refined methods which are essential to the evaluation of effectiveness, efficiency and benefits were the least common in usage and the lowest in preference by the sample.

4. A higher percentage of the professional than the non-professional agencies used and preferred the more refined and sophisticated methods of service evaluation. Similarly, there was a higher percentage of the pro-
essional than the non-professional agencies which specified their agency objectives in measurable terms.

5. As a whole the respondents considered the methods used "fairly much" adequate for their purposes.

6. As to preference for sources of evaluators, the sample highly preferred to conduct their evaluation by themselves. The least preferred were those where the evaluation of their services did not involve their agency personnel, such as by the funding source, or some experts outside of the agency.

7. The greatest limitations and difficulties experienced by the sample were those related to the skill and knowledge of evaluation research, as well as those related to financial resources.

8. The main purposes for the sample to evaluate their services were: (a) to improve services, and (b) satisfy the funding source and the public at large.

9. The sample placed a high value on service evaluation. They thought that: (a) the existing service evaluation was useful to their agencies; (b) service evaluation in general was very important; and (c) they would like to improve their existing service evaluation as soon as possible.

10. The kind of help seen as most useful was consultation and guidance from experts to improve service evaluation.

11. There were a few statistically significant relationships found between annual operating budget and (a) types
160

of agencies; (b) the method, "systematic follow up of clients;" (c) degrees of preference for "some experts outside of the agency."

12. There was fear as well as resistance to full acceptance of service evaluation, as indicated by: (a) the reactions of the respondents in the processes of data collection; (b) the relatively low percentage of the sample using and preferring any evaluative methods other than those related to functional budgeting; and (c) the sample's low preference for funding source or outside experts as evaluators.

Implications

Clearly indicated in these findings were the willingness and desire to change and improve, even though the sample considered the existing evaluative practices were adequate for their purposes. They, in general, lacked the skill and knowledge of evaluation research and they needed expert guidance rather than financial help to advance from effort evaluation to efficiency, effectiveness or benefits evaluation. It may be due to the lack of proper understanding or a fear of the unknown that they did not use or prefer the more sophisticated methods.

The fact that the sample were in the second year of a three-year program to establish functional accounting partly explained that the sample used and preferred most
methods essentially related to effort evaluation. They had acquired the knowledge and skill, and, therefore, felt comfortable and confident in using these methods. In other words, most of the sample did not use or prefer the more sophisticated methods because they had not used them and had no proper knowledge of them.

The function of evaluation research as discussed in the literature survey is basically to examine what has and what has not been accomplished in order to look for improvement and changes. Evaluation research also exposes the realistic operations of the agency in contrast to its public appearance. It forces attention to what is actually done rather than what is intended. The whole approach then is to doubt and to be critical. This approach naturally puts the recipients of evaluation in a defensive situation. The common fear of resistance found among the sample and the care for the agency's public image in the respondent's reporting on the interview schedule, could well support what is said in the literature.

Discussed in the following are the specific implications of the findings for (a) the approach to improve service evaluation; (b) the U.C.S.; (c) the agencies; (d) social work education; and (e) further research.

---

192 Please refer to pp. 24-5 and pp. 79-82 of this report for the discussion on this aspect.
For the Approach to Improve Service Evaluation

The foregoing should have some implications on the approach to improve the evaluative practices of the sample. The first basic approach, to the researcher, should be to orient the agency directors to see the real need for and the true function of evaluation research. Evaluation research is necessary in the process to modify agency approaches and services to meet new needs and problems resulting from the continuous changes in social and physical environments. Therefore, the premise for doing service evaluation should be to find out how the agency services could be improved. How the agency directors should be oriented to think (a) that to do evaluation is basically to prevent the agency from becoming obsolete; and (b) that to discover service inadequacy is the basic approach of evaluation and is essential for improvement and change. Though there was no lack of the desire to improve the service evaluation as indicated in the findings, having a proper perspective of evaluation research and its basic approach would partially overcome the fear and resistance as well as to avoid having "eye-wash" "white-wash" sorts of evaluation.

The orientation for agency directors could take the forms of seminars or talks conducted by those who had some

193 For further explanation, please refer to pp.20-3 of this report.
expertise and practical experience in agency service evaluation. However, it is of paramount importance that such orientation discussions should not be too academic nor theoretic; they should be directed toward action and practice as a follow-up.

The second basic approach is to afford the agencies with opportunities to evaluate their services under the guidance of some "experts." The findings indicated strongly such need. The best way to produce confidence in the sample for using more sophisticated methods is to demonstrate how the methods could be applied and for what benefits. No amount of words or theory could probably produce such a confidence level. To afford the opportunity to work under qualified guidance would depend on the availability of the "experts." In this connection, the researcher would suggest that the U.C.S. and its member agencies look beyond their own settings for some expert, or well qualified persons. A list of the experts with their availability for consultation or guidance should be compiled and made public.

The third basic approach should be to strive towards agency coordination and team effort in the field of service evaluation. As indicated by the findings, the majority of the agencies were small in terms of staff complement and budget. In addition, there were agencies which offered similar types of services. As the researcher sees it, a cooperative program shared by a few agencies
for improvement of service evaluation, would: (a) reduce financial and human resources required; (b) provide the kind of stimulation necessary for growth; and (c) lead to better service coordination and better planning. Cramer documents the success of two agencies sharing a cooperative program for staff development.\textsuperscript{194} The successful experience indicates that one agency sharing a program with another is not only desirable but also workable.

The need or even the success of cooperation between agencies in one area does not necessarily indicate or lead to the amalgamation of agencies. In fact, successful cooperation and coordination among agencies on some areas would only indicate that there is no need for amalgamation as each has its distinct contribution to make as well as its ability to coordinate and cooperate. The only snag on the way to such cooperation would be the traditional mentality and habit of non-cooperation especially between agencies offering similar types of services.

For the Windsor United Community Services

For the U.C.S. as a funding source, the findings could suggest that the U.C.S., in addition to its present functions of providing funds and encouraging service evaluation, should attempt to provide some practical advice and

guidance in the improvement of service evaluation. Though the present staff of the U.C.S. did not have the necessary experience and training to directly provide needed expert guidance, it should assist its members to enlist the help of those who have such expertise either as volunteers or by consultation fees. To expect agencies to learn to improve evaluation by trial-and-error or under unqualified guidance would only make the experience unfruitful and frustrating. There would be some qualified persons around should the U.C.S. decide to look for some.

Another implication is for the U.C.S. to play an active role in getting the agencies to start shared cooperative programs for carrying out service evaluation. To play this role, the U.C.S., in the first place, should be convinced of the workability and benefits of such shared programs. In addition, it should be able to accept whatever challenges and potential threat this kind of cooperation and coordination might present to the present status of the U.C.S. and its relationship with its members.

The third implication is that despite the recommendation of the Program and Budget Review Committee,195 the U.C.S. should consider dropping the idea of developing an instrument for service evaluation at this stage. It could begin, should it continue to play its role in encouraging

service evaluation, with the basic approaches suggested earlier.

The fourth implication would be for the U.C.S. to consider employing someone who has strong background in research and policy development. The necessity for a person of such qualifications is self-evident considering: (a) the role expected of the U.C.S., and (b) the present staff complement of the U.C.S.¹⁹⁶

For the Agencies

There were considerable differences in the ratings between professional and non-professional agencies on various aspects of evaluative practices as indicated in the findings. The distinction between professional and non-professional as made by the U.C.S. was based on the number as well as the educational background of the professional persons who occupied the key positions in the agencies. If the agencies are to improve service evaluation, the findings seemed to suggest then that they should move toward more professionalism, that is, employ more persons with professional degrees.

The other major implication would be for the agencies to attempt to develop some kind of shared program with one or more agencies in the evaluation of services.

¹⁹⁶Please refer to p. 99 of this report for information on the staff complement of the U.C.S.
The agencies should also consider looking for some qualified persons to guide their evaluation.

For Social Work Education

It is also evident from the findings and the implications enunciated earlier that in the social work training, the need for the worker to evaluate his practice and for the agency to evaluate its output, should be stressed. The social work students should be oriented to accept and to cultivate a mentality of evaluating their own practice. With such a mentality, they are likely to be change-oriented and future-oriented. It is only when the social workers are prepared and willing to adapt and change to meet new problems and needs that the social work profession will advance with the society.

The findings may also suggest that there is a need for social work educators to (a) put a heavier emphasis on the research component of their curriculum, and (b) produce more students with specialization on research and policy development.

For Further Research

There were a few relationships found statistically significant at 5 per cent or 10 per cent level; those between the size of budget and certain evaluative practices, as well as those between the types of agencies and certain
evaluative practices. As reflected in Appendices D, E, and F, it did not follow that the larger the budget the more likely certain methods would be used. Nor did the relationship vary negatively. What would be the optimum point in budget size, or staff complement for certain types of practices to occur? It is worthy of investigation.

As explained elsewhere, there was a close relationship between the types of agencies and the educational background of the staff members of the agencies. With what types of professional degrees or training are persons more likely to accept evaluation, or do certain types of evaluation?

In actual fact, the types of agencies are to some extent related to the nature of services the agency offered. For example, the counseling services may require of the staff, certain professional qualifications. Therefore, persons with that kind of professional degree would be employed, resulting in the agency being more likely to be classified "professional." Are there any significant relationships between the nature of services and the four types of evaluation or the methods used in evaluation? What are they if there are? This is one of the areas that the researcher had not been able to explore, but felt worthy of further consideration.

One major area of concern for the evaluation researchers is how to popularize evaluation research. In other words, how evaluation research could be presented or explained in order to get easy acceptance. What strategy or approach
should be used? What aspects of evaluation research could be sacrificed or simplified for beginners? These are questions worthy of further consideration.

The data obtained raises many other concerns and doubts for further research.

**Recommendations**

Based on the foregoing discussion on implications of the findings, the researcher would like to make some recommendations to the U.C.S. and its member agencies, as a conclusion to this study.

That the U.C.S. and its member agencies should:

(a) consider having some orientation seminar and talks conducted by experienced evaluation researcher;
(b) look beyond their settings for some who are well-versed with evaluation research to assist them in improving their service evaluation; and
(c) attempt to develop shared, cooperative program in the area of evaluation among the agencies.

That the U.C.S. should consider:

(a) dropping the idea of developing an instrument for service evaluation at this stage; and
(b) employing someone with a strong background in research and policy development.

That the member agencies of the U.C.S. should consider employing more persons with professional degrees.
APPENDICES
INTERVIEW SCHEDULE

1. How long has your agency been in existence? (Please check one).
   (1) Less than a year
   (2) 1 to less than 3 years
   (3) 3 to less than 5 years
   (4) 5 to less than 7 years
   (5) 7 to less than 9 years
   (6) 9 to less than 11 years
   (7) 11 to less than 13 years
   (8) 13 to less than 15 years
   (9) 15 to 20 years
   (10) Over 20 years

   Comments:

2. How long has your agency been a member agency of U.C.S.? (Please check one).
   (1) Less than a year
   (2) 1 to less than 3 years
   (3) 3 to less than 5 years
   (4) 5 to less than 7 years
   (5) 7 to less than 9 years
   (6) 9 to less than 11 years
   (7) 11 to less than 13 years

   Comments:
3. What is your present staff complement?

A) Number of full-time staff

<table>
<thead>
<tr>
<th></th>
<th>Paid</th>
<th>Non-Paid*</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Professional</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Secretarial &amp; Clerical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Maintenance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Other (specify)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

B) Number of part-time staff

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Professional</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Secretarial &amp; Clerical</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Maintenance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Other (specify)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments: __________________________________________

4. What is the approximate annual operating budget of your agency?

$________

Comments: __________________________________________

*Including volunteers, and students on field work placement.
5. Below is one of the classifications of social services. What type(s) of service(s) does your agency primarily offer? (Please check one only)

1) Services for Primary Prevention (services that enhance the general well-being of a population, e.g. social and recreational services)

2) Services for Secondary Prevention (e.g. case-finding, diagnosis and short treatment that reduce symptoms and minimize contagion)

3) Service for Tertiary Prevention (concerned with chronic and irreversible illness)

4) Services for Primary and Secondary Prevention

5) Services for Primary and Tertiary Prevention

6) Services for Secondary and Tertiary Prevention

7) Other (specify)

8) Not sure

Comments: 

__________________________________________________________________________

__________________________________________________________________________
6. What types of information does your agency keep in relation to service evaluation? (Please check each of the following)

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
<th>Not Sure</th>
<th>Not Applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>Staff use of time</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2)</td>
<td>Cost of different services</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3)</td>
<td>Number of clients served</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4)</td>
<td>Characteristics of clients served</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5)</td>
<td>Staff's quality of work</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6)</td>
<td>The efficiency of service delivery (i.e.: quality of services over relative cost)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7)</td>
<td>The effects of the services on the clients</td>
<td></td>
<td></td>
<td>b</td>
</tr>
<tr>
<td>(8)</td>
<td>Client's reactions to the services (e.g.: clients satisfaction and suggestions)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9)</td>
<td>The benefits (usually expressed in monetary terms) the services have brought to the clients</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10)</td>
<td>The need situations towards which the services are directed</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(11)</td>
<td>Other (specify)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments: ____________________________________________
7. Below are some ways (methods) to evaluate service effectiveness. Please check each of the following if your agency has used them.

(1) Statistics on number of clients served

(2) Statistics on characteristics of clients served

(3) Record of time spent delivering service

(4) Experimental or quasi-experimental designs (e.g.: use of control group, "before-after" testing)

(5) Staff observation of client changes

(6) Systematic collection of client responses by questionnaire, interview schedule, or a scale during treatment period

(7) Selective or casual collection of clients' reports (i.e. letters of appreciation, taking the "successful" cases)

(8) Systematic follow up of clients after termination of services

(9) Periodic systematic survey and definition of need

(10) Make some sort of observation or "guess" on need situation periodically

(11) Other (specify)

Comments: ____________________________________________________________

*The word "systematic" here means according to a uniform set of procedures planned and pre-determined.
8. Referring to the method(s) mentioned, which would your agency prefer for evaluating its services? (Please check as many as seem appropriate).

(1) ____ (7) ____
(2) ____ (8) ____
(3) ____ (9) ____
(4) ____ (10) ____
(5) ____ (11) ____

Comments: __________________________________________

9. a) For the purposes of evaluation, does your agency specify its objectives in measureable (that is, something which can be quantified) terms? (Please check one).

   Yes ____  No ____  Not Applicable ____  Not Sure ____

b) If your answer is yes above, how successful do you feel you were?

   (5) A great deal ______
   (4) Fairly much ______
   (3) To some degree ______
   (2) Comparatively little ______
   (1) Not at all ______
   (9) Not sure ______

Comments: __________________________________________
10. As a whole, do you consider the methods used by your agency adequate for evaluating its services? (Please check one).

(5) A great deal
(4) Fairly much
(3) To some degree
(2) Comparatively little
(1) Not at all
(9) Not sure

Comments: ____________________________________________________________
11. Of the following, which would you prefer to conduct your agency service evaluation? (Please rate each in terms of the degree of your preference by circling the appropriate number)

<table>
<thead>
<tr>
<th></th>
<th>A Great</th>
<th>Fairly</th>
<th>Some</th>
<th>Deal</th>
<th>Much</th>
<th>Compar-</th>
<th>Not</th>
<th>At</th>
<th>Not</th>
<th>appli-</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Your agency itself</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Some experts outside the agency</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) The funding source(s)</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) The agency itself with outside experts as advisors</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) A joint team of the agency staff, outside experts and the funding source(s)</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) Other (specify)</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments: ___________________________________________________________
12. What are the major difficulties or limitations that your agency has experienced in service evaluation? (Please rate each of the following in terms of the degree of difficulty or limitations your agency has experienced by circling the appropriate number).

<table>
<thead>
<tr>
<th></th>
<th>Great</th>
<th>Fairly Great</th>
<th>Fairly</th>
<th>Some</th>
<th>Slightly</th>
<th>At Least</th>
<th>Not Much</th>
<th>Not Applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Specifying service goals in measurable terms</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(2) Finding proper scientific ways of yielding reliable data</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(3) Determining what information is required</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(4) Getting feedback from clients</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(5) Meeting resistance from the clients</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(6) Sparing the staff &amp; time</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(7) Preventing hindrance to the smooth functioning</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(8) Lacking financial resources</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>(9) Lacking understanding and enthusiasm among the staff</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th></th>
<th>A Great</th>
<th>Fairly</th>
<th>Some</th>
<th>comparatively</th>
<th>At</th>
<th>Not</th>
<th>Not</th>
<th>Appli-</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>Meeting resistance from the staff</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>11</td>
<td>Lacking encouragement from the board of directors/management</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>12</td>
<td>Having much value judgement or subjectivity in evaluation</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>13</td>
<td>Having nobody or nowhere to turn to for consultation</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>14</td>
<td>Having to assess the effects of the intervening variables</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>15</td>
<td>Meeting complications caused by various stages of program/service development</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>16</td>
<td>Other (specify)</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
</tr>
</tbody>
</table>

Comments: ________________________________
13. What are the main reasons for your agency to evaluate its services? (Please rate each of the following in terms of the extent it applies to your agency by circling the appropriate number).

<table>
<thead>
<tr>
<th>Reason</th>
<th>Great</th>
<th>Fairly</th>
<th>Some</th>
<th>Deal</th>
<th>Much</th>
<th>Comparatively</th>
<th>At</th>
<th>Not</th>
<th>Not</th>
<th>Applied</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) to satisfy the board of directors</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) to satisfy the funding source</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) to satisfy the public and community at large</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) to evaluate the work of the staff</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) to let the public know how the agency is doing</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) to get some testimonials of success</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(7) for planning and service improvement</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(8) to get some feedback for records</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9) to show how effort and time are spent</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(10) to ascertain the progress towards the agency goals</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(continued)
| (11) to find a more efficient way of service delivery | 5 | 4 | 3 | 2 | 1 | 9 | 0 |
| (12) to determine to what extent the societal needs have been met | 5 | 4 | 3 | 2 | 1 | 9 | 0 |
| (13) to determine how much the clients have benefited from the services | 5 | 4 | 3 | 2 | 1 | 9 | 0 |
| (14) to keep up with the trend | 5 | 4 | 3 | 2 | 1 | 9 | 0 |
| (15) to determine if agency goals should be changed or modified | 5 | 4 | 3 | 2 | 1 | 9 | 0 |
| (16) Other (specify) | 5 | 4 | 3 | 2 | 1 | 9 | 0 |

Comments: __________________________________________

________________________________________________________________________
14. How does your agency use the results of service evaluation? (Please check all that are used).

(1) filed in record for director's reference
(2) brought to the administrative level for study
(3) brought to the planning level for study
(4) brought to staff meeting for discussion
(5) brought to the board of directors
(6) filed in record for auditor's reference
(7) reported to the U.C.S.
(8) published in newsletter, local papers, etc.
(9) Other (specify)

Comments: __________________________________________

_________________________________________________________________

15. Generally speaking, how useful is the existing service evaluation to your agency? (Please check one).

(5) A great deal
(4) Fairly much
(3) To some extent
(2) Comparatively little
(1) Not at all
(9) Not sure

Comments: ____________________________________________
16. In general, how would you rate the importance of service evaluation? (Please check one).

(5) A great deal
(4) Fairly much
(3) To some extent
(2) Comparatively little
(1) Not at all
(9) Not sure

Comments: _______________________________________________________

17. What would you like to see happen to the existing service evaluation? (Please check one).

(1) Be simplified
(2) Be done away with
(3) Stay as it is for the time being
(4) Be improved whenever possible
(5) Be improved immediately
(9) Not sure
(6) Other (specify)

Comments: _______________________________________________________
18. If your agency chose to improve its service evaluation, what kinds of help do you think it would need? (Please rate each of the following in terms of the degree of need by circling the appropriate number).

<table>
<thead>
<tr>
<th></th>
<th>A Great</th>
<th>Fairly Some</th>
<th>To Handle</th>
<th>Compar-</th>
<th>Not</th>
<th>Not</th>
<th>At All</th>
<th>Sure</th>
<th>All</th>
<th>Sure</th>
<th>cable</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) More funds</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Additional trained staff</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(3) Outside expertise to</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>direct your service evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Some training or</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>orientation for your existing staff on service evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) Some advice or</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>consultation on service evaluation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(6) Other (specify)</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comments: ________________________________________
THE MEMBER AGENCIES OF THE UNITED COMMUNITY SERVICES
OF GREATER WINDSOR AS AT MARCH 20, 1974

Big Brother Association of Greater Windsor
Boy Scouts of Canada
Canadian Arthritis and Rheumatism Society
Canadian National Institute For The Blind
Canadian Red Cross
Catholic Family Service Bureau
Cerebral Palsy Association of Windsor and Essex County
Community Information Service
Credit Counselling Service of Metropolitan Windsor
Crossroads
Family Service Bureau
Girl Guides of Canada
Goodwill Industries of Windsor
John Howard Society
Mental Health/Windsor-Essex
St. John Ambulance
St. Leonard's House
Salvation Army
Senior Citizens' Centre
Social Planning Division, United Community Services of Greater Windsor
Tel-A-Friend
Victorian Order of Nurses
Windsor Association For The Mentally Retarded
Windsor Group Therapy Project
Windsor Jewish Community Council
Windsor Safety Patrol Association
Windsor Y.M.-Y.W.C.A.
Dear

I am a graduate student of the School of Social Work, University of Windsor, doing my field practice at the United Community Services of Greater Windsor. With the assistance of Mr. Michael Pfaff, the Executive Director of U.C.S., and the approval of the school, I have planned a survey of the service evaluation of U.C.S. member agencies as my research project in fulfillment for the Master of Social Work degree. The project is being carried out under the guidance of a Committee, the chairman and members of which are listed at the close of this letter.

Service evaluation has become very popular as the legislators and citizenry increasingly ask why a service program deserves public support and why it merits priority. Those responsible for budgeting public funds and those responsible for programming services are keen to know what impact the service will have on clients and the community, and to evaluate in terms of cost. Both the reports of the Program and Budget Review Committee and of the Social Planning Review Committee of the U.C.S. recognize the need for and recommend as task the improvement of the evaluation of agency services. There is also strong desire among U.C.S. member agencies striving for "better" evaluation.

It is my belief that a survey of the existing practices with service evaluation in the U.C.S. member agencies would yield some data base helpful to both the agencies themselves and the U.C.S. in their endeavour.
An interview schedule has been drawn up to collect information on various aspects of your agency service evaluation. The interview schedule is designed for the response of the Executive Director of each agency. It is my hope that you will give me your kind co-operation. I will be calling you in the next few days. It will be much appreciated if you will give me an appointment at your earliest convenience. A copy of the interview schedule is enclosed herewith for your reference.

Thanking you and looking forward to meeting you.

Yours sincerely,

Lam Wong

The Thesis Committee:

Chairman - Dr. Lola E. Buckley, School of Social Work, University of Windsor

Members - Mr. Michael J. Pfaff, United Community Services of Greater Windsor
- Prof. Forrest C. Hansen, School of Social Work, University of Windsor
- Dr. Mary L. Dietz, Dept. of Anthropology and Sociology, University of Windsor

cc. Agency President
APPENDIX D

ANNUAL OPERATING BUDGET BY THE METHOD
"SYSTEMATIC FOLLOW UP OF CLIENTS"

<table>
<thead>
<tr>
<th>Budget</th>
<th>Responses</th>
<th>Yes</th>
<th>No</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below $80,000</td>
<td></td>
<td>0</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0)</td>
<td>(100.0)</td>
<td></td>
</tr>
<tr>
<td>Between $80,000</td>
<td></td>
<td>5</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>and $140,000</td>
<td></td>
<td>(83.3)</td>
<td>(16.7)</td>
<td></td>
</tr>
<tr>
<td>Above $140,000</td>
<td></td>
<td>4</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(100.0)</td>
<td>(10.0)</td>
<td></td>
</tr>
<tr>
<td><strong>Column Total</strong></td>
<td></td>
<td>9</td>
<td>6</td>
<td>15</td>
</tr>
</tbody>
</table>

Chi square = 11.52777  2 df  $p < .05$

$c = 0.65921$. 
APPENDIX E

ANNUAL OPERATING BUDGET BY DEGREES OF PREFERENCE FOR "SOME EXPERTS OUTSIDE OF THE AGENCY"

<table>
<thead>
<tr>
<th>Budget</th>
<th>Responses</th>
<th>Not At All</th>
<th>Comparatively Little</th>
<th>To Some Extent</th>
<th>A Great Deal</th>
<th>Row Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below $80,000</td>
<td>0</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Between $80,000</td>
<td>3</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>and $140,000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Above $140,000</td>
<td>0</td>
<td>2</td>
<td>3</td>
<td>0</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td><strong>Column Total</strong></td>
<td><strong>3</strong></td>
<td><strong>3</strong></td>
<td><strong>3</strong></td>
<td><strong>2</strong></td>
<td><strong>16</strong></td>
<td></td>
</tr>
</tbody>
</table>

Chi square = 13.000  6 df  $\alpha = 0.05$

$\chi^2 = 0.66953$
# Appendix F

**Descriptions of the Selected Variable Relationships Using Appropriate Statistical Measures**

<table>
<thead>
<tr>
<th>Variables</th>
<th>Relationship in Statistical Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of Agencies by Evaluative Method Used: &quot;Systematic Follow up of Clients.&quot;</td>
<td>Fisher's Exact Test = 0.08392 Phi = 0.49099 c = 0.44073</td>
</tr>
<tr>
<td>Types of Agencies by Preference for Eevaluators: &quot;Some experts outside the agency.&quot;</td>
<td>Chi square = 6.75555 2 df significance = 0.0801; Cramer's V = 0.64979 c = 0.544386</td>
</tr>
<tr>
<td>Types of Agencies by Difficulty: &quot;Lacking encouragement from the board.&quot;</td>
<td>Chi square = 6.44920 3 df significance = 0.0917; Cramer's V = 0.61593 c = 0.52443</td>
</tr>
<tr>
<td>Types of Agencies by Difficulty: &quot;Meeting complications caused by program development.&quot;</td>
<td>Chi square = 8.5555 4 df significance = 0.0732; Cramer's V = 0.78174 c = 0.61588</td>
</tr>
<tr>
<td>Types of Agencies by Evaluative Information: &quot;Service efficiency.&quot;</td>
<td>Fisher's Exact Test = 0.11905 Phi = 0.63246 c = 0.53452</td>
</tr>
<tr>
<td>Agency Length of Existence by Difficulty: &quot;Having much value judgement or subjectivity in evaluation&quot;</td>
<td>Chi square = 7.8000 3 df significance = 0.0503; Cramer's V = 0.74642 c = 0.59816</td>
</tr>
</tbody>
</table>

(continued)
### APPENDIX F—Continued

<table>
<thead>
<tr>
<th>Variables</th>
<th>Relationship in Statistical Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Annual Operating Budget by Types of Evaluative Information: &quot;The service benefits.&quot;</td>
<td>Chi square = 5.8333 2 df significance = 0.0541; Cramer's V = 0.64550 c = 0.54232</td>
</tr>
<tr>
<td>Annual Operating Budget by Evaluative Method Used: &quot;Experimental designs.&quot;</td>
<td>Chi square = 4.61558 2 df significance = 0.0995; Cramer's V = 0.55470 c = 0.48507</td>
</tr>
<tr>
<td>Annual Operating Budget by Evaluative Method Preferred: &quot;Selective or casual collection of clients report.&quot;</td>
<td>Chi square = 5.73109 2 df significance = 0.0570; Cramer's V = 0.51040 c = 0.45461</td>
</tr>
<tr>
<td>Annual Operating Budget by Difficulty: &quot;Sparing the staff and time.&quot;</td>
<td>Chi square = 13.81239 8 df significance = 0.0068; Cramer's V = 0.63738 c = 0.66953</td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY

Books


---

Articles and Periodicals


Reisman, Arnold; Eisenberg, Norman C.; and Beckman, Allen. "Systems Analysis and Description, the Jewish Communal System." Journal of Jewish Communal Services, XLVI, 1(Fall, 1969).


Unpublished Materials


"Form A and Form B of U.C.S. Agencies Program and Budget Presentation."

VITA AUCTORIS

Mr. Wong Lam Wo was born in Kuala Lumpur, Malaysia in February, 1939. He received his first nine years of education in Kampar, Malaysia, before he continued his education in Singapore. After his graduation in 1963 from the Nanyang University with a B.A. degree in Modern Languages and Literature, he worked first as a high school teacher, and then a senior instructor of the National Youth Leadership Training Institute, Singapore, until he received a Canadian Commonwealth Scholarship to further his studies in Windsor, Canada.

Mr. Wong has keen interest in social development, and has considerable experience in community and labour organizations and leadership development. He represented his country in some international conferences.

He received his B.S.W. degree in 1973, he expects to be conferred the M.S.W. degree in the Fall convocation of 1974 from the University of Windsor. In his B.S.W. field practice, he worked as a counselor in both a high school as well as a community youth outreach program. In the M.S.W. program, he was on placement at the Windsor United Community Services, working with citizen groups and doing planning and policy development.

Mr. Wong has found cross-cultural learning extremely stimulating and inspiring.