Innovation Network Inc

Program Evaluation Practice in the Nonprofit Sector

A study funded by
The Aspen Institute Nonprofit Sector Research Fund
and
The Robert Wood Johnson Foundation

September 30, 1998

Allison H. Fine
Colette E. Thayer
Anne Coghlan
Innovation Network, Inc.
1001 Connecticut Ave. NW
Washington DC 20036

Acknowledgements

This study would not have been possible without the hard work and insightful contributions of numerous individuals.

We are grateful to the members of the study's advisory committee: Alan Abramson, Sylvia Barsion, Janet Carter, Charles Firke, Jennifer Greene, Martha Taylor Greenway, Barbara Kehrer, Kathryn Newcomer, Karen Simmons, and David Stevenson. Their leadership, support, and advice in designing and implementing this study were tremendously helpful.

We would like to offer special thanks to Kathryn Newcomer for contributing considerable time and insight to this study. Dr. Newcomer was instrumental in each stage of our research and her input added significantly to the study. In addition, we are grateful to Martha Taylor Greenway and Karen Simmons for their advice and roles in reviewing preliminary findings and draft reports.

We appreciate the diligent work done by Andrew Goldsmith and Amy Downey entering survey data and by Roseana Bess conducting telephone interviews. We also would like to thank Karen Brotherton and CiCi Hughes for their efforts and tireless support throughout the study.

We are also grateful to The Aspen Institute Nonprofit Sector Research Fund and The Robert Wood Johnson Foundation for supporting this important study.

Most of all, we appreciate the time respondents gave to the study, their willingness to share information, and the thoughtful answers they provided.

Finally, we would like to note that the authors are solely responsible for the accuracy of the statements and interpretations contained within this report. The interpretations do not necessarily reflect the views of the sponsoring foundations or of the advisory committee.

Program Evaluation Practice in the Nonprofit Sector

Executive Summary

The current funding and political environment make it more important than ever that nonprofit organizations demonstrate the effectiveness of their work. Government leaders and the public express skepticism that social interventions have any discernible impact on the problems in our society; public and private funding sources are being eliminated or reduced, and the competition for funds means that nonprofit organizations must offer evidence of their effectiveness. To do so, nonprofit organizations must choose among various evaluation approaches and methods to assess their effectiveness, demonstrate their value to the community, provide useful and usable information, and meet the demands of funders and the public.

In recent years there has been growing debate between two broad approaches to program evaluation. In the more traditional model, an external evaluator is employed as an objective observer who collects and interprets quantitative and qualitative findings, and presents the information to management. A “scientific” paradigm is used that focuses on the quality of the data collected and an evaluation is considered valid to the extent that it meets specific standards of methodological rigor. In the participatory evaluation model, program staff, clients, and other “stakeholders” of a program are engaged in the evaluation process to increase their evaluation knowledge and skills and to increase the likelihood that evaluation findings will be used. Because this approach to evaluation relies on program staff both to formulate evaluation questions and collect data, these investigations may be less objective by the standards of the scientific paradigm. They are valued, however, because they improve the analytic capacity of program participants, and also increase the likelihood that evaluation results will be used to refine and improve programs.

This study of Program Evaluation Practice in the Nonprofit Sector was designed to determine which of these broad approaches are being used by nonprofits, to what extent they are being used, and with what effects. Analysis revealed that the methods and components of past and present evaluations being conducted by nonprofit organizations add up to a complex constellation that does not fall neatly into easily definable categories. Therefore, this study focused on exploring the role of stakeholder participation in program evaluation since it is a key element that distinguishes between traditional and participatory approaches to evaluation. In particular, the study addressed whether different levels of stakeholder participation are associated with various organizational characteristics, evaluation characteristics and evaluation outcomes. The study also sought to determine how organizations used evaluation results and whether the uses are associated with different levels of stakeholder participation.

The study used a multiple and mixed-method research design conducted over three phases. First, a national mail survey of nonprofit organizations believed to have conducted at least one program evaluation in the past three years was administered. Three hundred two surveys were mailed to valid addresses, of which 178 were completed and returned for a response rate of 59%. Next, telephone interviews with a sample of forty mail survey respondents were conducted. Finally, in-depth profiles of four organizations were compiled to examine in greater detail the characteristics and outcomes of particular evaluations reported in the mail survey.

The study reveals that recent evaluations are:

- ➢ Focused on outcome measurement, which is a fairly recent practice that is continuing to grow.
- ➢ Conducted primarily for current funders, followed by program staff, and board members.
- ➢ Include a mixture of quantitative and qualitative data collection methods.

The study also illustrates that regardless of stakeholder participation level:

- ➢ Evaluations were highly likely to use a combination of quantitative and qualitative data. Seventy-one percent of the high participation evaluations used mixed methods, as did 66% of the low participation evaluations.

We found that high stakeholder participation evaluations were:

- ➢ More likely than low stakeholder participation evaluations to be conducted for volunteers (19% compared with 2%).
- ➢ Used more than low stakeholder participation evaluations in nearly all use categories, including:
  - · To improve outcomes or impact (92% compared with 64%);
  - · To promote the program to potential participants (73% compared with 47%);

  - To respond to questions/criticisms about the program (65% compared with 47%); and
  - To decide resource allocation within the organization (58% compared with 22%).

- More likely to be used to design on-going monitoring/evaluation systems than low participation evaluations.

In addition, we found that low stakeholder participation evaluations were:

- More likely than high participation evaluations to be primarily designed for strategic planning (14% compared with 4%).

Our study also reveals that:

- Most respondents (70%) reported that stakeholder involvement plays an important role, especially in increasing the likelihood of programmatic changes being made based on evaluation findings.

- Respondents with high stakeholder participation evaluations were more likely to think their staff, board and funders found the evaluation highly credible than respondents with low-participation evaluations.

The study explored respondents' views of positive evaluation outcomes including usefulness, credibility and satisfaction.

- Factors that contribute to an evaluation's usefulness include the following:
  - Documentation of programmatic success,
  - A useful program planning tool,
  - A focused evaluation design.

- Increased stakeholder involvement was credited with improving evaluation design (by helping to guarantee that relevant questions are asked and appropriate measures are selected), ensuring available resources to implement the evaluation and its recommendations, increasing stakeholders' understanding and appreciation of an agency, and improving stakeholders' understanding of the evaluation.

- Most respondents stated that evaluations conducted by an external contractor are perceived to be more credible. The main reasons given were objectivity and expertise of consultants.

- Respondents with high participation evaluations were more likely to think that funders found the evaluation satisfying than respondents with low participation evaluations (65% compared with 45%).

The field of nonprofit evaluation has grown considerably over the last decade. In fact, increased awareness and resources for evaluation have opened possibilities for evaluation design and use unimaginable five years ago. At a time when program evaluation is evolving quickly, this study presents a window into the growing practice of evaluation in the nonprofit sector.

I. Introduction

Background

The current funding and political environment makes it more important than ever for nonprofit organizations to demonstrate the effectiveness of their work. Government leaders and the public express skepticism that social interventions have any discernible impact on the problems in our society; public and private funding sources are being eliminated or reduced, and the competition for funds means that nonprofit organizations must offer evidence of their effectiveness. To do so, nonprofit organizations must choose among various evaluation approaches and methods to assess their effectiveness, demonstrate their value to the community, provide useful and usable information, and meet the demands of funders and the public.

In recent years, there has been growing debate between two broad approaches to program evaluation. In the more traditional model, an external evaluator is employed as an objective observer who collects and interprets quantitative and qualitative findings, and presents the information to management. A "scientific" paradigm is used which focuses on the quality of the data collected and an evaluation is considered valid to the extent that it meets specific standards of methodological rigor.

More recently, as seen in the work of Patton (1997), Guba and Lincoln (1989), and Fetterman (1996), a participatory evaluation model has been used. The focus of this approach is to engage program staff, clients and other stakeholders in the evaluation process so that the information collected is used to improve the program. Because they rely on program staff both to formulate evaluation questions and collect data, these investigations may be less objective by the standards of the scientific paradigm. They are valued, however, because they improve the analytic capacity of program participants, and also increase the likelihood that evaluation results will be used to refine and improve programs.

Our study has taken place during a period when the field of nonprofit evaluation has blossomed. In recent years, grants made by large foundations for program evaluation have increased both in amount and number. In 1994, large foundations awarded approximately $38 million in grants for program evaluation, representing 0.6% of the total support awarded. In 1996, the amount large foundations awarded for program evaluations jumped to $87 million, accounting for 1.2% of their total support awarded. (Renz, Mandler and Treiber, 1998). In addition, in a recent study of the field, United Way of America cites the growth of nonprofit evaluation efforts including: national organizations conducting research on outcomes, national organizations developing outcome measurement resources for local agencies, managed-care companies stressing service results, accrediting bodies considering outcome-related criteria, and local agencies operating effective outcome measurement systems (Plantz, Greenway and Hendricks 1997).

The increased awareness of and resources for evaluation appear to have affected our study. For instance, more survey respondents than anticipated had completed evaluations within the last three years; more stakeholders were involved in the design and implementation of evaluations than we expected, and more internal staff are leading evaluation efforts than we anticipated. Upon further analysis, we found that some of our results were counterintuitive. We did not find that participation levels were consistent with any particular organizational characteristics (e.g. budget, type of services, staff size), nor did we find approach or design associated with whether an inside staff member or an outside consultant was primarily responsible for the evaluation. The lack of any predictable pattern between the purpose of the evaluation, characteristics of the organization, evaluator and evaluation design leads us to believe that the dramatically changing environment, particularly the evaluation pressure brought to bear by grantmakers, is causing significant changes in the numbers of agencies conducting evaluations. The determination of these agencies to designate staff time to participate in the evaluation and to use evaluation results for a variety of purposes is also likely to have been affected. Overall, this study reflects a sample of nonprofit agencies that had positive evaluation experiences and many of them are continuing their evaluation efforts.

Purpose of the Study

At the study's outset, our goal was to develop a clear set of criteria upon which to classify an evaluation approach as "traditional," "participatory" or some hybrid of the two. Subsequently, we intended to determine the extent to which each approach was being used and the outcomes it engendered. We hypothesized that natural clusters of evaluations would empirically emerge from the data based on an evaluation's purpose, intended audiences, methods, who had primary responsibility for conducting it, and level of stakeholder participation.

Our study revealed that the methods and components of evaluations recently and currently being conducted by nonprofit organizations add up to a complex constellation that does not fall neatly into easily definable categories. Therefore, throughout our study, we placed an emphasis on exploring stakeholder participation since it is a key element that distinguishes between traditional and participatory evaluation approaches.

In addition to exploring which evaluation models are being used by nonprofits, the study was designed to examine the associations between evaluation approaches and the various organizational characteristics, and evaluation characteristics, outcomes, and uses. Again, we distinguished these results by level of stakeholder participation.

Specifically, the following narrative will inform the field about:

- evaluation practices among nonprofit organizations,
- the extent to which program stakeholders are involved in recent nonprofit organization evaluation practice, and
- the associations between levels of stakeholder participation and specific positive evaluation outcomes.

II. Literature Review

Evaluation Approaches

By looking at evaluation practice over time, scholars have devised numerous typologies to organizing and describing evaluation theories, approaches and models. House (1980), in using the earlier work of Stake (1973), Popham (1975) and Worthen and Sanders (1973), identifies a taxonomy of eight major evaluation models or approaches, including system analysis, behavioral objectives, decision-making, goal-free and case study. Each of these approaches differs in its audience, methodology, outcome and the typical questions it asks.

Building on House's work, Greene (1990) identifies four major genres of evaluation research. As shown in Table 1, the first is Postpositivism, which investigates efficiency, accountability and causal linkages for high-level policy and decision makers, using experimental or quasi-experimental designs with quantitative data collection methods. The next approach is Pragmatism, which looks at issues of management and utility mostly for program decision-makers and often using a mix of quantitative and qualitative data collection techniques. Less common is the Interpretivist approach to program evaluation, which seeks to understand program experiences through diverse views, and is done for multiple program stakeholders by using primarily qualitative methods. Least common is Critical Theory evaluation, done to empower program beneficiaries through the participation of stakeholders using a mix of data collection methods.

Further building on this work, Greene and McClintock (1991) identify a repertoire of eight to ten distinct approaches evaluators can choose from which together represent a widening range of philosophical principles and scientific methodologies. This substantial transformation of evaluation purpose and methodology results from two major forces: a) recognition that observations are shaped by the values of the observer and b) new "organizing frames" for evaluation (ibid.). Specifically, the new frames include: a broadened conception of how evaluation results are used; recognition that different phases of program development have different data needs; attention to program components other

These recent changes have resulted in more diverse and responsive sets of evaluation approaches which, according to Shadish et al. (1991), have become more inclusive in focus, methods and audience and have a greater concern for context. More specifically,

> exclusive reliance on studying outcomes yielded to inclusive concern with examining the quality of program implementation and the causal processes that mediated any program impacts. Exclusive reliance on quantitative studies yielded to including qualitative methods. Using policy makers as both the source of evaluation questions and the audience for the results yielded to consideration of multiple stakeholder groups. Concern with methodology gave way to concern with the context of evaluation practice and fitting evaluation results into highly politicized and decentralized systems (ibid. 32).

As Shadish et al. (1991) point out, a growing trend in evaluation practice has been the inclusion of multiple stakeholders groups. Indeed, a number of evaluators (Patton, 1997; Guba and Lincoln, 1989; Fetterman et al., 1996) advocate extensively for the use of participatory evaluation or the involvement of program stakeholders in designing and implementing an evaluation.

In a different typology, Kaminsky (1993) identifies participatory evaluation as one of four major evaluation approaches. Accordingly, participatory evaluation as an evaluation approach is usually done for the purpose of collaborative education for multiple stakeholders, it asks primarily process-oriented evaluation questions, and it uses experiential knowledge. This approach is in sharp contrast to the more conventional experimental approach to evaluation conducted by an objective external evaluator for the purpose of instrumental change and which uses scientific knowledge and methods to determine outcomes and causality (ibid.).

Participatory-oriented Evaluation

Participatory-oriented evaluation refers to the participation of program stakeholders in the various phases and tasks of an evaluation. The ways in which stakeholders participate and the reasons for their participation vary from evaluation to evaluation.

Nevertheless, nearly all evaluation is participatory in some way and, thus, several theorists conceptualize participation as a continuum (Kaminsky 1993, Rebien 1996, Whitmore 1988). One end of the continuum represents a low degree of participation where stakeholders merely provide data and are essentially evaluation objects. On the other end, stakeholders are in full control and are actively involved in all stages of the evaluation.

Participatory-oriented approaches evolved largely in response to the non-use of evaluation results. In the early 1970s, evaluation theorists began to criticize evaluation practices as overly mechanistic and insensitive, and resulting in unused findings (Worthen et al., 1997). Specifically, evaluation results were not being used because the information supplied by evaluators was either not appropriate, not available when needed, or not presented in useful or understandable ways for decision-makers (Gold, 1983). Furthermore, evaluations were too narrow in scope. They contained unrealistic standards for success, were often unfair to program staff and participants, and were largely irrelevant to decision-makers (Weiss, 1983).

To overcome these shortcomings, and particularly to increase the use of evaluation results for decision making, a number of participant-oriented evaluation approaches evolved, including responsive evaluation (Stake, 1972; 1983), stakeholder-based evaluation (Gold, 1983), illuminative evaluation (Parlett and Hamilton, 1976), transactional evaluation (Rippey, 1973), democratic evaluation (MacDonald, 1974) and utilization-focused evaluation (Patton, 1986, 1997). A common rationale to all these approaches was that participation would increase the usefulness of evaluation results (Whitmore, 1988).

More recently, Cousins and Earl (1995) altered the stakeholder-based evaluation model. Their rationale for what they term “participatory evaluation” is still to enhance evaluation utilization. But unlike most previous stakeholder-based models, their version involves smaller numbers of primary users who are much more involved in actually doing the evaluation. In addition, the researcher and stakeholders are in more of a partnership, where both parties share responsibility for conducting the evaluation and the researcher serves as a coordinator and technical advisor.

Several studies show evaluations that involve stakeholders are better able to “stimulate knowledge production and promote positive reactions to evaluation findings” (Shulha and Cousins, 1997: 195). In addition, Greene (1988) finds stakeholder participation in evaluation to be an important way to increase the use of evaluation findings, specifically: instrumental use for program changes; conceptual use for a broader understanding of the program; persuasive use or citations of the results in reports and proposals to external audiences; and symbolic use for enhancing program prestige and visibility. Cousins and Earl (1992, 1995) also promote stakeholder collaboration as a vehicle for capacity building and organizational learning. Lastly, Fetterman et al. (1996) find numerous ways in which involvement in evaluation empowers stakeholders. As Weiss (1997) concludes in a recent review of evaluation use, “Collaboration between evaluators and program staff all through the evaluation process tends to increase the local use of evaluation. Not always . . . But participatory methods do seem to help.”

Evaluation Practice in the Nonprofit Sector

The term nonprofit sector is used to describe institutions or organizations that are neither government nor businesses and do not exist to make money for owners or investors. The nonprofit sector in the United States is vast and diverse with more than one million organizations that spend nearly $500 million annually (National Center for Nonprofit Boards and Independent Sector, 1997). Under Section 501(c)3 of the tax code, there are more than 25 classifications of nonprofit organizations including charities, foundations, social welfare or advocacy organizations and professional and trade associations. Of those with incomes above $25,000, roughly 50,000 are classified as human service organizations (Plantz, Greenway and Hendricks, 1997). As in earlier investigations (ibid), the term “nonprofit” in this study refers to human service organizations which offer direct services such as youth development, employment assistance and substance abuse counseling.

Although there is not a great deal of literature describing specific evaluation approaches for the nonprofit sector, Sumariwalla and Taylor (1991) describe how conventional evaluation approaches grounded in the hard sciences have not been successful in evaluating human service programs. More appropriate for decision making are less formal approaches such as responsive evaluation and naturalistic inquiry that focus on the interests and concerns of stakeholders, use ordinary language, and are based on everyday reasoning (ibid).

Several studies investigate evaluation practice in the nonprofit sector. An area that has received considerable study is evaluation practice in foundations. For example, in 1993 the Council on Foundations compiled a book to orient grantmakers and grantees to different kinds of grant evaluation, consisting of nine case studies illustrating a variety of evaluation approaches. In 1994, the Council also conducted a foundation management survey which explored grantmakers' grant monitoring and evaluation practices (Kroll, n.d.). Within this survey, nearly all the respondents reported that they evaluate the projects they fund, of which almost four-fifths is handled by staff or board members. Most respondents used evaluation findings to assist them in making future funding decisions or to measure program effectiveness. Almost half reported to have used evaluation findings to capture program results or to demonstrate the strategic value of a program.

Another study investigating evaluation practice and use in small and mid-sized foundations found similar results (McNelis and Bickel, 1996). From a national mail survey, most of the 170 responding foundations reported that they evaluate funded programs as part of their standard operating practices, but mostly in the form of grant monitoring by foundation staff and grantee self-reports. Evaluation was mostly done to hold grantees accountable, assess program impact, and inform strategic thinking in foundations. The most commonly cited barrier to conducting evaluation was lack of resources, including money and qualified staff, followed by organizational beliefs and attitudes (ibid).

In a follow-up study, McNelis, Bickel and Connors-Gilmore (1996) find that small and mid-sized foundations do not usually evaluate their own programs but rather focus on their grantees' programs.

Looking at a broader sample of nonprofit organizations, Sumariwalla and Taylor (1991) ask purposively selected funding agencies and national representative organizations about the evaluation activities of their members, affiliates or grantees. Although answers varied by type of organization, respondents overall reported that the most common type of evaluation conducted was "Measurement of Volume of Program Delivery," followed by "Compliance with Program Delivery Standards" and "Tracking of Characteristics of Program Participants." The least frequent types of evaluation were assessing "Program Outcomes/Results" and "Participant Satisfaction." The two greatest barriers to conducting assessment of program outcomes were lack of financial support and lack of qualified staff. In addition, a number of respondents stated the need to identify valid models and accurate measures for program evaluation. More broadly, the authors concluded

> A new paradigm for program evaluation separate from that of rigorous evaluation research models needs to be developed for nonprofit organizations. . . . there are many quasi-experimental and qualitative evaluation methods that can be routinely and realistically applied in non-profit agencies. The development of these models should be furthered, rather than continuing a focus on "pure" evaluation research (ibid. 11).

In another report on evaluation practice in the nonprofit sector, The Nonprofit Management Association and Support Centers of America (1998) recently conducted a baseline study of current evaluation practices and barriers among management support organizations (MSOs). By using a mail survey and follow-up phone interviews, the investigators found that although a common evaluation framework seems to be missing among MSOs, respondents focused on measuring customer satisfaction and were not, on the whole, measuring program outcomes. Nearly all the respondents reported using an internal evaluator and just over half reported using external evaluators. Nearly half reported conducting evaluation in order to improve services, followed by nearly 20% conducting evaluation to fulfill funding requirements. Cited barriers to conducting evaluation included insufficient time, funds and staff capabilities. Respondents expressed a need to integrate evaluation practices into the delivery of services. A major recommendation of the study was the development of a common evaluation framework for MSOs.

In summary, from the few studies conducted, it appears that evaluation practice in the nonprofit sector has changed significantly. At one time, it was based on scientific models that proved to be inappropriate and not useful to local decision-making. Evaluation practice then became more inclusive and focused on process, or implementation assessment. In recent years, and particularly among MSOs, there has been a focus on customer satisfaction.

But even more recently, a major development in the nonprofit sector is the growth of outcome measurement or assessing “the benefits or changes for participants during or after their involvement in a program” (Hatry, van Houten, Plantz and Greenway, 1996: 2). As a result of increasing budget cuts, nonprofit agencies must demonstrate greater accountability and worth. Thus, national organizations such as Big Brother, Big Sisters of America and Girl Scouts of the U.S.A. are researching program outcomes (Plantz, Greenway and Hendricks, 1997). Several national organizations, including United Way of America and Goodwill Industries, have developed resources for local agencies to develop outcome measurement systems. In addition, thousands of local service agencies have developed outcome monitoring systems. Although relatively new, some studies are finding this current form of outcome measurement to be useful to both local and funding agencies (ibid).

Overall, however, there is little in the literature on evaluation practice and use in human service nonprofit organizations and the effectiveness of different types of evaluation approaches. There is even less on determining the use and effectiveness of participatory evaluation or stakeholder participation in evaluation of nonprofit organizations. This study attempts to fill this gap by identifying evaluation practices among a sample of human service nonprofit organizations and specifically to determine the effects and value of stakeholder involvement.

III. Methodology

The study used a multiple and mixed-method research design conducted over three phases. First, a national mail survey was sent to nonprofit organizations believed to have conducted at least one program evaluation in the past three years. Program evaluation was defined as a one-time, systematic assessment of any program or service an organization delivers that includes an overall design, data collection and analysis, and reporting of results. In Phase 2, telephone interviews with a sample of mail survey respondents were conducted. In Phase 3, in-depth profiles were compiled of four organizations to examine in greater detail the characteristics and outcomes of particular evaluations reported in the mail survey.

Phase I: National Mail Survey. The mail survey was primarily designed to obtain quantitative data describing characteristics and outcomes of evaluations recently completed by a wide range of U.S. nonprofit organizations. To better target the survey effort, a purposive sample was devised by asking knowledgeable individuals from foundations, national support organizations, management support organizations, and independent evaluators to nominate nonprofit organizations they believed to have completed a program evaluation in the past three years. After eliminating duplicate or non-valid nominations, a list of 308 organizations was compiled.

Among the 308 questionnaires sent, six were returned unanswered primarily because of invalid addresses. From the revised sample size of 302, we received 178 completed questionnaires for a response rate of 59%. Of the 178 completed questionnaires, 140 were from nonprofit service delivery organizations that had completed at least one evaluation, and thus, were valid to include in the analysis.

The questionnaire was designed by InnoNet with the input and feedback of the study's advisory committee members (see Appendix I) and pre-tested with organizations similar to those on the survey mailing list.

The questionnaire consisted primarily of closed-ended questions asking about the respondents' organizations and the characteristics and outcomes of their organization's most recently completed evaluation within the last three years. (For a copy of the questionnaire, see Appendix II.) The quantitative results of the survey were analyzed by frequency counts, percentages, means and cross-tabulations. Because the sample for the survey was purposive and not meant to represent a broader population, tests for statistical significance, although calculated, cannot be used to generalize to the entire nonprofit community. (Organizational characteristics of survey respondents are included in Appendix III.)

Phase II: Telephone interviews. Telephone interviews lasting approximately fifteen minutes were conducted with forty respondents randomly selected from the original survey respondent list of 178 organizations. These interviews were done to determine current evaluation practices, to explore issues surrounding outcome measurement, and to understand what makes evaluations useful, credible and satisfying. The instrument consisted of closed-ended questions that asked respondents about their organizations' current evaluation efforts (as of May 1998) and open-ended questions soliciting their opinions on what makes evaluations useful, credible and satisfying. Frequencies and percentages summarized quantitative data, and qualitative data were coded and summarized by common and divergent themes. (Organizational characteristics of interview respondents are included in Appendix III.)

Phase III: In-depth Profiles. To gain a better understanding of the evaluations that were rated highly useful, credible and satisfying by their staff in the mail survey, four evaluation efforts were profiled by using a purposively selected sample of mail survey respondents. Methods used to compile the profiles included document review and in-depth interviews. Twenty-seven evaluations received the highest possible ratings by respondents on usefulness, credibility and overall satisfaction of the 140 completed mail surveys. Of those, we asked respondents who had not been contacted for the previous telephone interviews to send us their evaluation report and information on the evaluated program. After reviewing these documents, four respondents were purposively selected for in-depth interviews, representing two evaluations with low levels of stakeholder involvement and two evaluations with high levels of stakeholder involvement. Participants were asked open-ended questions about how their evaluation was conducted and particularly about stakeholder involvement, how the evaluation results were used, and what, in their opinion, made their evaluations highly useful, credible and satisfying. Narrative summaries for each set of documents and interviews were written which were then analyzed and summarized by common and divergent themes.

IV. Findings

A. Recent and Current Evaluation Practices Among Nonprofit Organizations

Evaluation practice varies widely among nonprofit organizations. Specifically, agencies select among different purposes, audiences, evaluators, and data collection methods, in addition to a host of other variables, in designing an evaluation. In this study, program evaluation was defined as a one-time, systematic assessment of any program or service an organization delivers that includes an overall design, data collection and analysis, and reporting of results. Although the practice of evaluation is complex, a number of key findings emerged from our analysis:

80% of the respondents currently conducting evaluations are primarily measuring outcomes.

✓ Outcome measurement is a fairly recent practice that is continuing to grow.

✓ Evaluations are being conducted primarily for current funders, followed by program staff and board members.

✓ Evaluations tend to include a mixture of quantitative and qualitative data collection methods.

Purpose – Why was the evaluation conducted?

In both recent and current evaluations, the most commonly cited purpose for conducting an evaluation was to measure program outcomes or impact. More than half (56%) of the recently completed evaluations (N=140) were designed primarily to measure outcomes or impact while 80% of the organizations currently conducting evaluations (n=35) have designed their evaluations to do so.

As Figures 1 and 2 illustrate, considerably fewer organizations are conducting their evaluation for the primary purpose of assessing program implementation, informing strategic planning, assessing the quality of operations, and measuring client satisfaction.

Figure 1. Primary Purpose of Most Recently Completed Evaluations (N=140)

Primary Purpose

Figure 2. Primary Purpose of Current Evaluations (n=35)

Primary Purpose

Note: Total exceeds 100% due to some respondents conducting multiple evaluations

Nearly half of the respondents currently evaluating outcomes are doing so as a funding requirement.

Because of the high proportion of organizations reporting that their most recently completed evaluation measured program outcomes or impact, this issue was probed more deeply in the telephone interviews. Respondents who are currently measuring or monitoring outcomes (n=33) were asked to describe the types of outcomes they are evaluating. Analysis revealed that nearly three-quarters (70%) are indeed examining outcomes that are "benefits or changes for participants during or after their involvement with a program." (Hatry, van Houten, Plantz, and Greenway, 1996: 2) or customer satisfaction. Eighteen percent listed a mix of outcomes and outputs, and one organization listed only outputs.[1] Examples of mentioned outcomes included increased parenting awareness, decreases in teen pregnancies, decreases in drug and alcohol use, and educational progress. Examples of mentioned outputs included the number of people exposed to the program and the number of businesses opened by participants.

Nearly half of the respondents measuring outcomes (43%) reported doing so because it is a funding requirement. Of these 17 respondents, 20% specifically mentioned a United Way requirement for outcome measurement.[2] A fifth (20%) also said they are measuring outcomes to see if the program is working or making a difference, and at least another 15% said to improve program planning. Several respondents also explained that their organizations are measuring outcomes for fundraising purposes, because it is "a smart thing to do," and "because it is the wave of the future."

Of the 33 organizations currently measuring or monitoring outcomes, 24% started to do so between 1976 and 1989, 24% between 1990 and 1994, and 45% since 1995[3].

Recently completed evaluations (N=140) were done for numerous audiences, the most common being current funders (69%), program staff (61%),

[1] Because of ambiguous answers, it was not possible to tell if three organizations are measuring outcomes or not.
[2] There is a strong probability that the United Way of America had an influence on our sample's evaluation efforts. Of the 40 interview respondents, 65% are members of the United Way. And of the 33 organizations that are currently measuring outcomes, more than half (58%) said they have been influenced by the United Way.

board members (56%) and potential funders (44%). The evaluations were also done for government agencies (28%), program participants or clients (24%), and other nonprofit organizations (9%).

Recently completed evaluations were done for numerous audiences, the most common being current funders (69%), program staff (61%), board members (56%), and potential funders (44%).

Implementation – How was the evaluation conducted?

It was reported that internal staff had the primary responsibility for conducting 51% of recent evaluations (N=140) and 40% of current ones (n=35). External contractors were responsible for leading 40% of recent evaluations and only 20% of current ones. A combination of internal staff and an outside contractor lead 9%[4] of the recent evaluations and 29% of the current ones.

Respondents were asked to indicate the data collection methods used in their evaluations. When survey respondents' replies were coded by data type, we found that more than two-thirds (67%) of the most recently completed evaluations used a mix of qualitative and quantitative data collection methods. Similarly, more than three-quarters (89%) of the interview respondents reported using mixed methods.

For survey respondents, the most commonly reported data collection method was review of program documents (84%), followed by individual interviews (66%), observation (50%), written questionnaires (42%), focus groups or group discussions (33%), and tests of knowledge or skills (19%). In addition, 15% of the respondents reported that a control or comparison group was used in their most recently completed evaluation.

Continuing Evaluation Practice

Most respondents are continuing to conduct evaluations. Organizations participating in the survey (N=140) reported completing between one and 50 program evaluations in the past three years, with a median of completing three. Of interview respondents (n=40), 40% are currently conducting one evaluation, 15% are conducting two evaluations, 28% are conducting between three to ten evaluations and 5% are conducting more than ten evaluations.

---

[3] Six percent of the respondents were unsure when their organization began measuring outcomes.

[4] However, this may be an underreported figure because respondents may not have been aware that multiple responses (i.e. a mix of internal and external) were allowed.

The remaining 13% of the respondents reported that their organization is not currently conducting any evaluations.

B. Involvement of Program Stakeholders in Recent Nonprofit Organization Evaluations

As the field of participatory evaluation grows, the roles that stakeholders are playing in program evaluation efforts need to be analyzed to determine whether various levels of stakeholder involvement are affecting the implementation and outcomes of evaluations. In particular, it is important to explore whether the level of stakeholder involvement is associated with organizational and/or evaluation characteristics. Our analysis revealed the following key findings:

- ✓ Staff was the stakeholder group most involved in all stages of an evaluation.
- ✓ Board members' largest role came in planning/designing and analyzing/ interpreting data.
- ✓ Participants were involved in data collection more than any other stage of the evaluation.
- ✓ Nearly a quarter of the respondents involved funders in planning/designing their evaluation.
- ✓ Evaluations were more likely to have high stakeholder participation than low in organizations with budgets from $1 million to $10 million, while evaluations were more likely to have low stakeholder participation than high in organizations with budgets below $1 million.
- ✓ Organizations that conducted their evaluation for volunteers or the United Way were more likely to have high participation than low.
- ✓ Organizations that designed their evaluation primarily for strategic planning were more likely to have low participation than high.

Levels of stakeholder involvement

To investigate levels of stakeholder involvement, we divided recently completed evaluations (from the survey) into low, medium and high levels of stakeholder participation. The division was based on scores that were computed from reported levels of stakeholder involvement in response to a survey question (See Appendix II, Survey, Question 15). As Figure 3 illustrates, the question asked whether staff, board members, funding agency personnel and program participants were involved in designing the evaluation, collecting data, analyzing or interpreting the results and/or writing the evaluation report.

For the entire sample, most organizations (79%) involved staff in data collection and more than half (61%) involved staff in planning the evaluation.

Figure 3. Stakeholder Participation Level Scoring Matrix

Evaluation Stage	Stakeholder Group: Staff	Board Members	Funders	Program Participants
Design				
Data collection				
Data analysis or Interpretation				
Report writing				

The level of stakeholder involvement could range from zero (no stakeholder involvement) to sixteen (each stakeholder group involved in each stage of the evaluation). Each evaluation was "scored," and then the scores were divided as evenly as possible into thirds representing low, medium and high levels of stakeholder participation.

Using this process, 41% of the sample fell within the low stakeholder participation level with a score ranging from 0 to 2, 40% fell into the medium level with a score of 3 or 4, and 19% fell into the high level with scores ranging from 5 to 11. It is important to note that these levels were not categorized as low, medium or high by pre-defined participation criteria, but rather they were defined empirically based on the range of stakeholder participation in our sample.

Types of stakeholder involvement

For the entire sample (N=140), most organizations (79%) involved staff in data collection and more than half (61%) involved staff in planning the evaluation. Although fewer organizations involved staff in data analysis or interpretation (39%) and report writing (29%), roughly a third of the organizations did so.

Few organizations involved their board in data collection (5%) or report writing (4%). Nevertheless, almost a fifth of the organizations involved their board members in planning the evaluation (16%) and analyzing or interpreting data (18%).

Program participants collected data in 3% of the low participation evaluations and in 38% of the high participation evaluations.

The largest role played by participants was in data collection (12%). In fact, less than a tenth of the respondents involved participants in planning the evaluation (6%), analyzing or interpreting data (4%) or report writing (2%).

Nearly a quarter of the organizations (21%) involved funders in planning their evaluation. To a lesser extent, funders were also involved in analyzing and interpreting data (11%). In far fewer instances, funders participated in data collection (7%) and report writing (6%).

The types of stakeholder involvement by participation level are illustrated in Figures 4, 5, and 6 (and appear in Appendix V).

Low-participation evaluations (n=58) mainly involve staff. Staff generally participates in planning the evaluation and/or collecting data; however, they do not analyze or write reports in many instances. Other stakeholder groups are involved to a very limited extent in low participation evaluations.

Medium-participation evaluations (n=56) tended to involve staff in all stages of the evaluation. In addition, there is some board and funder participation in planning as well as board involvement in data analysis and interpretation.

By definition, high-participation evaluations involve more stakeholders in each stage of the evaluation. In these evaluations, (n=26) staff are highly involved in all stages. Nearly half of the organizations involved their board in planning the evaluation and analyzing data. More than a third involved participants in collecting data. Almost half involved funders in planning the evaluation and nearly 40% involved them in analyzing or interpreting data.

Figure 4. Low-Participation Evaluations (n=58)

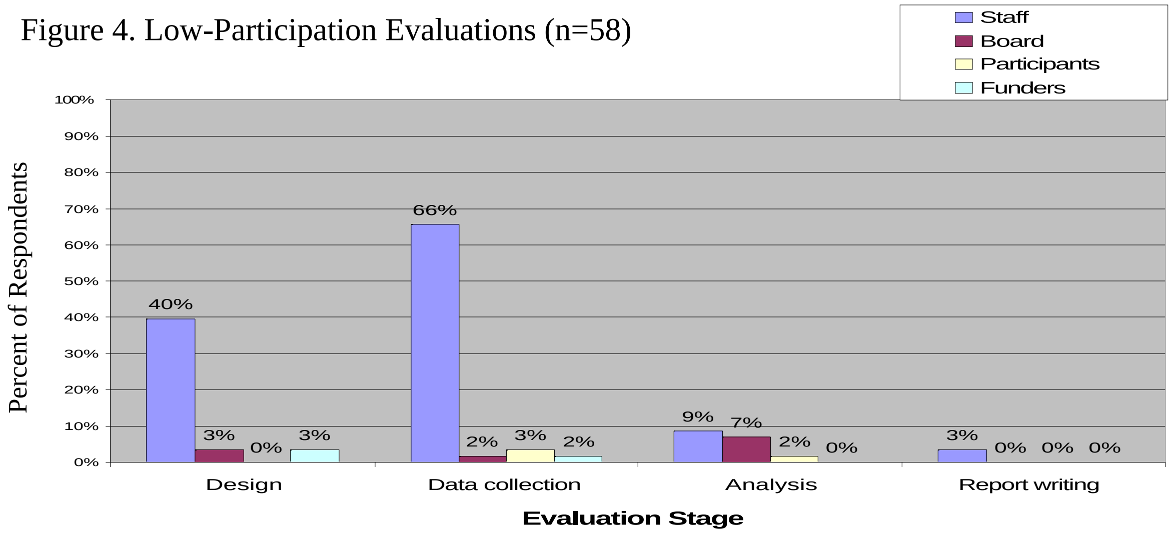


Figure 5. Medium Participation Evaluations (n=56)

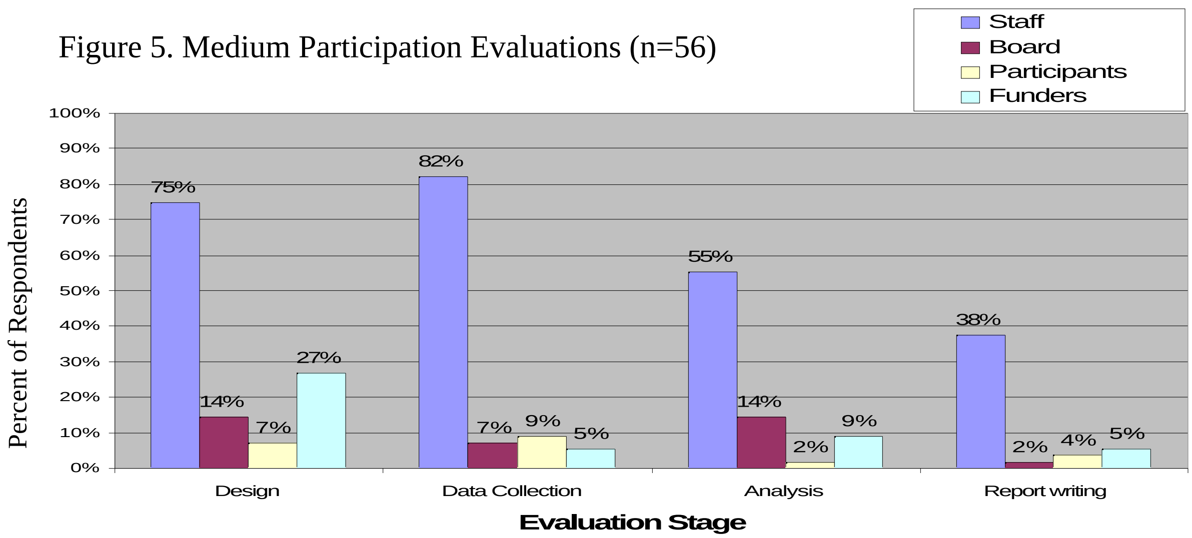
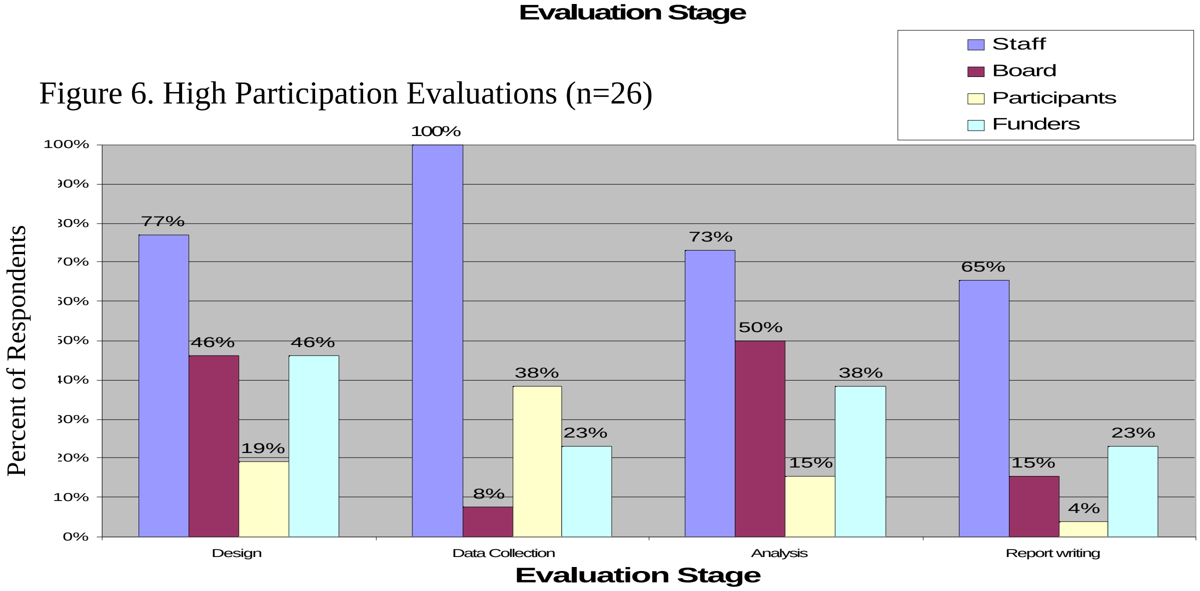


Figure 6. High Participation Evaluations (n=26)



Respondent Characteristics by Level of Stakeholder Participation

As seen in Table 2, this study examined the relationship between stakeholder participation level and an organization's primary service, full-time equivalent staff size, annual operating budget, in-house evaluation staff size, and evaluation budget. While our study revealed that none of the relationships were statistically significant, a few of them are worth mentioning because the differences are large enough between stakeholder groups that they may warrant further study.

Evaluations were more likely to have high-participation in:

- Human service organizations (79% compared with 65%).
- Organizations with budgets from $1 million to $ 10 million (58% compared with 45%).
- Organizations that have evaluation staff and have one to five people who devote at least half of their time to evaluation (45% compared with 29%).

Evaluations were more likely to have low-participation in:

- Organizations with budgets under $1 million (43% compared with 31%).
- Organizations that have staff with a primary responsibility for evaluation but do not have anyone that devotes at least half of his/her time to evaluation (69% compared with 55%).

Table 2. Percentage of Survey Respondents by Organizational Characteristics and Level of Stakeholder Participation for Most Recently Completed Evaluation

		Level of stakeholder participation		
Organizational Characteristics	All	Low	Medium	High
Primary Service	(N=134)	(n=57)	(n=53)	(n=24)
Human Services	67%	65%	64%	79%
Health, Mental Health	11	12	13	4
Education	9	9	8	13
Arts, Culture, Humanities	3	5	0	4
Environment, Animals	2	0	4	0
Other	8	9	11	0
Full-time Equivalent Staff	(N=140)	(n=58)	(n=56)	(n=26)
< 25	49%	52%	45%	50%
26 – 50	19	21	21	12
51 – 150	18	16	20	19
>151	14	12	14	19
Annual Operating Budget	(N=140)	(n=58)	(n=56)	(n=26)
< $1,000,000	39%	43%	38%	31%
$1,000,000 – 10,000,000	47	45	45	58
> $10,000,000	14	12	18	12
Have Evaluation Staff	(N=140)	(n=58)	(n=56)	(n=26)
Yes	81%	81%	82%	81%
Those who have evaluation staff, # of staff who devote 50% of time to evaluation	(n=110)	(n=45)	(n=45)	(n=20)
0	61%	69%	56%	55%
1 – 5	37	29	42	45
> 6	2	2	2	0
Have Budget for Evaluation	(N=140)	(n=58)	(n=56)	(n=26)
Yes	51%	53%	46%	54%
Amount for those who have Annual Evaluation Budget	(n=71)	(n=31)	(n=26)	(n=14)
< $5000	17%	23%	15%	7%
$5001 – 10,000	14	10	12	29
$10,001 – 20,000	18	19	23	7
$20,001 -- 50,000	17	13	15	29
> $50,001	34	36	35	29

Evaluation Characteristics by Level of Stakeholder Participation

High stakeholder participation evaluations were more likely than low participation evaluations to be conducted for volunteers.

Low stakeholder participation evaluations were more likely than high participation evaluations to be designed for strategic planning.

This study investigated the associations between levels of stakeholder participation and the evaluation's primary purpose; the evaluation's audiences; who had primary responsibility for conducting the evaluations; and data collection methods.

Statistically significant relationships (at the 0.05 level of significance) were found between level of stakeholder participation and some audiences. In particular:

- High participation evaluations were more likely to be done for volunteers than low participation evaluations (19% compared with 2%).
- High participation evaluations were more likely to be done for the United Way[1] than low participation evaluations (12% compared with 0%). In addition, although the relationship was not statistically significant, high participation evaluations were more likely to be primarily designed to measure outcomes or impact than low participation evaluations (65% compared with 56%). This is not surprising considering the United Way model for outcome measurement stresses stakeholder involvement.
- High participation evaluations were more likely than low participation evaluations to be led by internal staff (54% compared with 33%) and to be led by a mix or internal staff and an outside contractor (21% compared with 6%). Low participation evaluations were more likely to be led by an external contractor than high participation evaluations (61% compared with 25%).

In addition, there are a number of non-statistically significant relationships worth mentioning.

- Low participation evaluations were more likely to have been primarily designed for strategic planning than high participation evaluations (14% compared with 4%).
- Regardless of stakeholder participation level, evaluations were very likely to use a combination of quantitative and qualitative data. Seventy-one percent of the high participation evaluations used mixed methods, as did 66% of the low participation evaluations.

Table 3. Percentage of Survey Respondents by Evaluation Characteristics and Level of Stakeholder Participation for Most Recently Completed Evaluation

Evaluation Characteristics	All	Level of stakeholder participation: Low	Medium	High
Primary Purpose	N=140	n=58	n=56	n=26
To measure outcomes/impact	56%	52%	55%	65%
To assess implementation	9	7	11	12
For strategic planning	9	14	5	4
To assess quality of operations	7	7	9	4
To measure client satisfaction	4	7	4	0
Other	8	3	11	12
No primary purpose designated	7	10	5	4
For Whom was Evaluation Done [a]	N=140	n=58	n=56	n=26
Current funders	69%	66%	68%	81%
Program staff	61	52	68	69
Board	56	57	50	69
Potential funders	44	41	39	58
Government agencies	28	24	23	46
Program participants/clients	24	21	23	35
Other nonprofit organizations	9	5	14	8
Volunteers*	6	2	5	19
United Way (written in)*	4	0	5	12
National Office (written in)	3	5	0	4
Other	9	14	4	8
Prime Responsibility for Conducting Evaluation*	N=131	n=54	n=53	n=24
Internal staff	51%	33%	68%	54%
External contractor	40	61	25	25
Combination of internal and external	9	6	8	21
Data Collection Methods	N=132	n=56	n=52	n=24
Qualitative only	19%	16%	23%	17%
Quantitative only	14	18	12	13
Mix of quantitative and qualitative	67	66	65	71

[a] Multiple responses allowed.
*Pearson Chi-Square is statistically significant at 0.05 level.

C. Evaluation Uses

A common rationale for using participatory evaluation is that stakeholder involvement increases the likelihood that evaluation results will be used and that evaluation processes will continue. Our study of the relationship between stakeholder participation levels and reported uses revealed the following key findings:

✓ High-participation evaluations were used more than low-participation evaluations in nearly all use categories.

✓ High-participation evaluations were more likely to be used to design on going monitoring/evaluation systems than low-participation evaluations.

All Use Categories

As shown in Table 4, high-participation evaluations were used more than low-participation evaluations in nearly all use categories. The statistically significant relationships between stakeholder participation level and use include the following:

- To improve outcomes or impact
  (92% compared with 64%)

- To promote the program to potential participants
  (73% compared with 47%)

- To respond to questions/criticisms about the program
  (65% compared with 47%)

- To decide resource allocation within the organization
  (58% compared with 22%)

Other uses where differences among stakeholder participation emerged, although the relationships were not statistically significant, include the following:

- To change the program design or operation
  (85% of high participation compared with 67% of low participation)

- To plan future programs
  (85% of high participation compared with 67% of low participation)

- To improve outreach and public relations
  (54% of high participation compared with 36% of low participation)

- To train or reassign staff
  (46% of high participation compared with 35% of low participation)

A category where low participation had slightly higher use was “to report to a funder.” As shown in Table 4, 78% of the low-participation evaluations were used to report to a funding agency while only 73% of the high-participation evaluations were used to do so.

On-going Monitoring/Evaluation Systems

High-participation evaluations were more likely than low-participation evaluations to be used to design an ongoing monitoring/evaluation process (73% compared with 47%). A surprising finding was that there was little difference between high- and low-participation evaluations being used to gain support for evaluation among staff and board.

Table 4. Percentage of Respondents by Evaluation Outcome and Level of Stakeholder Participation for Most Recently Completed Evaluation

		Level of stakeholder participation		
Evaluation Outcome	All (N=140)	Low (n=58)	Medium (n=56)	High (n=26)
Evaluation used to:				
Report to a funder	76%	78%	75%	73%
Improve outcomes or impact*	75	64	79	92
Change program design or operations	73	67	73	85
Plan future programs	71	67	70	85
Get additional funding	54	47	61	58
Design ongoing monitoring/ evaluation process*	51	47	46	73
Promote the program to potential participants/clients*	51	47	46	73
Respond to questions/criticisms about the program*	48	47	41	65
Improve outreach and PR	41	36	39	54
Decide resource allocation within organization*	36	22	41	58
Train or reassign staff	36	35	32	46
Gain support for evaluation among staff or board	34	36	32	35
Other	4	3	7	0

*Pearson Chi-Square is statistically significant at 0.05 level.

D. Evaluation Outcomes

Practitioners of participatory evaluation believe that involving stakeholders will positively influence evaluation outcomes such as usefulness, credibility and satisfaction. This belief is predicated on the notion that involvement creates buy-in and reduces the likelihood that the findings will be deemed irrelevant. Our study revealed the following key findings:[5]

- ✔ Respondents cited a variety of factors that contribute to an evaluation's usefulness, including recommendations, documentation of success, providing a planning tool, and having a focused design.
- ✔ Respondents reported that stakeholder involvement plays an important role in increasing the likelihood that programmatic changes will be made based on evaluation findings.
- ✔ Respondents with high-participation evaluations were more likely than those with low-participation evaluations to think their staff, board, and funders found the evaluation highly credible.
- ✔ Most respondents stated that evaluations conducted by an external contractor are perceived as more credible than those conducted by internal staff.
- ✔ Respondents with high participation evaluations were more likely to think that funders found the evaluation highly satisfying than respondents with low-participation evaluations.

Usefulness

Overall, most survey respondents (77%) found their evaluation to be highly useful. However, as Table 5 shows, respondents with high-participation evaluations were more likely than those with low-participation to rate their evaluations highly useful (85% compared with 74%).

To help explain evaluation usefulness, we asked interview respondents what they think makes an evaluation useful to their organization.

[5] Although no statistically significant relationships exist between the outcomes and levels of stakeholder participation, these results are worthy of elucidation and suggest issues for further study.

The respondents reported a variety of factors. Most often they cited:

- **Recommendations for program improvement.** Several respondents (28%) mentioned recommendations and their connection to practical action as a key element of a useful evaluation. As explained by one respondent, "[A useful evaluation] includes constructive criticism in order to modify programs."
- **Documentation of program successes.** Documenting success was mentioned by respondents (28%) interested in demonstrating their achievements to both internal and external audiences. For example, "[An evaluation is useful] if it validates programs for external funders/stakeholders" and "[A successful evaluation] lets the board of directors and the community know we are doing a good job."
- **A useful planning tool.** Several respondents (20%) stated that useful evaluations are those that can be used for planning service delivery, future programs, resource allocations and/or for fundraising purposes. Comments that were typical of these respondents included the following: "[A useful evaluation has] the ability to give us direction; a useful tool to understand where we need to make changes and to see how to improve services" and "[A useful evaluation] helps us focus on mission and purposes and to organize everything (staff hiring, training, structure) around that."
- **A focused evaluation design.** A number of respondents (18%) found it important to have a focused evaluation design because they were interested in ensuring that the evaluation would answer their key questions. As several respondents explained, "The evaluation needs to be focused. There needs to be agreement ahead of time about what measures are central to the evaluation." In addition, "You need specific research/evaluation questions to make the evaluation useful."

When asked specifically whether who conducts an evaluation makes a difference to its usefulness, more than half of the respondents (63%) indicated that it does. According to respondents, internal staff and external contractors bring different assets to an evaluation effort.

"[A useful evaluation] helps us focus on mission and purposes and to organize everything (staff hiring, training, structure) around that."
Interview Respondent

In particular, external contractors' strengths are expertise and objectivity while internal staffs are knowledgeable about their programs and services. Several respondents indicated that the actual choice of an evaluator depends on the purpose of the evaluation. However, some respondents also suggested that useful evaluations result from using a combination of internal staff and an external contractor.

"Without stakeholders on board and understanding it from the outset, it's another piece of data that can sit in a desk drawer."
Interview Respondent

Respondents were also specifically asked if they have found that stakeholder involvement makes a difference to an evaluation's usefulness. Seventy percent of the respondents reported that stakeholder involvement plays an important role, especially in increasing the likelihood of programmatic changes being made based on evaluation findings. Respondents indicated that involvement creates buy-in and therefore creates incentives for implementation. In addition, increased stakeholder involvement was credited with improving evaluation design (by ensuring that relevant questions are asked and appropriate measures are selected), ensuring available resources to implement the evaluation and its recommendations, increasing stakeholders' understanding and appreciation of an agency, and improving stakeholders' understanding of the evaluation. As a respondent explained, "You need internal staff involved, particularly in collecting the data, for useful and reliable data and for changes to be implemented." In addition, "Without stakeholders on board and understanding it from the outset, it's another piece of data that can sit in a desk drawer."

Table 5: Evaluation Outcomes by Stakeholder Participation Levels

Evaluation Outcome	All (N=140)	Level of stakeholder participation: Low (n=58)	Medium (n=56)	High (n=26)
Usefulness				
Highly useful	77 %	74 %	77 %	85 %
Credibility				
Staff found evaluation highly credible	71	71	70	77
Board found evaluation highly credible	64	64	61	73
Funders found evaluation highly credible	63	57	66	69
Satisfaction				
Staff found evaluation highly satisfying	66	64	68	69
Board found evaluation highly satisfying	64	66	61	65
Funders found evaluation highly satisfying	54	45	59	65

"[A credible evaluation has] a good design that has multiple strategies, a design that confirms information from different perspectives."
Interview Respondent

Credibility

A concern organizations and evaluators often express about using participatory evaluation approaches is the credibility of data, the findings, and the overall evaluation. Our sample of respondents did not find these issues problematic. As Table 5 shows, respondents with high-participation evaluations were more likely to think their staff, board and funders found the evaluation highly credible than respondents with low-participation evaluations. A possible explanation for this result is that these stakeholder groups, by definition, were more likely to be involved in high-participation evaluations and therefore had the opportunity to voice their concerns in time for corrections to be made.

To explore the issue of credibility further, interview respondents were asked what they thought made an evaluation credible to program stakeholders. Respondents cited the following factors:

- **Design.** Half of the respondents) reported that an evaluation's design affects its credibility. For example, one respondent mentioned the importance of "collecting evidence instead of anecdotes [and] doing the evaluation systematically." Another respondent explained, "[A credible evaluation has] a good design that has multiple strategies, a design that confirms information from different perspectives."

- **Outside evaluator.** Several respondents (18%) noted the appearance of credibility that an outsider brings to an evaluation. For example, some respondents felt, "If the evaluator is an external contractor, the evaluation is looked upon as less biased."
- **Stakeholder involvement.** Several respondents (18%) also cited the role that increased stakeholder involvement plays in augmenting an evaluation's credibility. One respondent stated, "An evaluation is more credible to stakeholders if they participated in the creation of the it." In addition, "the process that was used to create an evaluation [is what makes an evaluation credible to stakeholders]. If they were incorporated into the process, they will believe in the process and will consider the source or process more credible no matter what the findings."
- **Reputation of the agency or evaluator.** Some respondents (13%) felt that the reputation of the agency or of the evaluator affects credibility. As one respondent explained, "Funders believe the results because of the good reputation of the organization." Another respondent stated, "Good agency reputation gives weight to an internal evaluation."

"It used to be key to have an external contractor but now that is less so the case as people realize the importance of process evaluations and expertise being made available to organization staff."
Interview Respondent

In addition, some respondents believed that perceived credibility is influenced by "the presentation of results since graphs and numbers are eye-catching," "when stakeholders hear what they want to hear," and "stakeholders' openness to constructive criticism."

When asked specifically whether who conducts an evaluation affects stakeholders' perceptions of credibility, slightly more than half of the respondents (53%) said that it does. In fact, most respondents stated that evaluations conducted by an external contractor are perceived to be more credible. The main reasons given were objectivity and expertise of consultants.

In contrast, some respondents believed that affiliation with an organization was less important than the evaluator's skills, abilities and knowledge of evaluation, as well as his/her familiarity with the program being evaluated.

For example, "Whoever conducts the evaluation needs to have a real understanding of the programs" and, "You need someone who understands different models [of evaluation] and can communicate ideas. It comes down to the education of the evaluator." Yet, another respondent stated, "We believe that doing an evaluation internally makes it more credible. The university researcher who is so concerned with rigor has missed the boat."

> *"The evaluation is satisfying if needs for improvement are found, as well as finding good work is being done and work is being done correctly."*
> Interview Respondent

In explaining the different views, a respondent described how perceptions of credibility have changed in recent years: "It used to be key to have an external contractor but now that is less so the case as people realize the importance of process evaluations and expertise being made available to organization staff."

When asked specifically whether stakeholder involvement makes a difference to an evaluation's credibility, many of the respondents (63%) said that it did. In particular, respondents explained that stakeholder involvement in the design and implementation of evaluations has a positive effect on their perceptions of credibility. In addition, some respondents mentioned that having stakeholders provide data increased their perceptions of its credibility. Nevertheless, some respondents believed that stakeholder involvement did not make a difference to an evaluation's credibility. For example, one respondent said, "It has more to do with who's doing the evaluation and the results gotten. Stakeholders have much less to do with credibility once the evaluation is under way."

Lastly, a respondent explained that the changing notion of stakeholder involvement and credibility is in part "[because] the United Way model includes stakeholder involvement [and] stakeholder involvement will increase credibility because everyone else in Kansas City is using that model. If an organization doesn't use this model, the evaluation seems less credible."

Satisfaction

Overall, more than half of the survey respondents thought their staff, board members and funders found the evaluation satisfying (66%, 64%, and 54% respectively). Although there appeared to be only slight differences in perceptions of staff and board members' satisfaction for different levels of stakeholder participation, there appears to be a greater difference in perceptions of funders' satisfaction for high- versus low-participation evaluations.[6] Respondents with high-participation evaluations were more likely to think that funders found the evaluation satisfying than respondents with low participation evaluations (65% compared with 45%). Involvement of stakeholders may have increased this outcome; however, this issue requires further study for better understanding.

Again, to help explain these results, we asked interview respondents what they think makes an evaluation satisfying. The most common responses pertained to demonstrating program effectiveness (35%) and identifying areas for program improvement (25%). For example, one respondent said, "[An evaluation is satisfying to stakeholders when it] validates that the mission is being accomplished … and exposes programs which need re-mediation or improvement." Another explained, "The evaluation is satisfying if needs for improvement are found, as well as finding that good work is being done and work is being done correctly." In addition, some respondents (13%) defined a satisfying evaluation as one which contains an honest appraisal of the program or agency, "not [one that is] glowing or glossed over."

Surprisingly, only a few respondents (5%) specifically mentioned the ability to use the evaluation for fundraising or marketing as a factor affecting satisfaction. Those who did make such references commented that a satisfying evaluation contains "[an] informative display of information so that it can be used for fundraising or marketing" and "brings in new resources."

Additional elements of a satisfying evaluation that were mentioned by some respondents included the following: an evaluation that is integrated into regular program activities rather than being "added on;" an evaluation process that involves stakeholders; an evaluation design that provides timely results; and one that includes qualitative as well as quantitative data.

[6] The difference is not statistically significant.

E. Reflections on Evaluations Receiving High Outcome Ratings

To gain a better understanding of the evaluation experiences of survey respondents, four in-depth profiles were composed of evaluations that were self-reported to have been highly useful, credible and satisfactory. These four evaluations were purposively selected from among twenty-seven that were reported as highly useful, credible, and satisfying to staff. To explore differences in stakeholder participation levels, two agencies with high-participation were selected and two with low-participation were selected.

Table 6 presents the respondents whose evaluations were profiled as well as a brief description of the program evaluated. In some cases, the respondent's evaluation covered multiple programs. Appendix VI contains in-depth profiles that include a description of the respondent, the evaluated program, the purpose and focus of the evaluation, the evaluation procedures, and the uses of the evaluation.

The following analysis of the profiles identifies and discusses common themes and lessons learned. Although clearly not a representative sample, these results are reflective of the findings from the interview and survey responses.

Table 6. Respondents and the Program(s) on Which Evaluation Pofiles Are Based.

Respondent	Program(s) Evaluated
The Salvation Army Hope Center for Children Offers six interrelated child and family programs that provide prevention and treatment services for abused and neglected children and their families.	**Residential Treatment Program** Inpatient program for children **Day Treatment Program** Structured program for children **Day Care Program** Primarily for low-income and single parent households **Therapeutic Foster Care and Adoption Program** Recruits, trains, and licenses families; matches children with families; and provides on-going support services.
The Girl Scouts of Genesee Valley, Inc. Is a youth-serving organization that serves 14,000 female members.	**Operation Stephne** Provides middle school girls with conflict resolution and anger management skills.
Metro Ministries of Corpus Christi Provides a variety of services to homeless people including food, transitional shelter, clothing, referrals, employment opportunities, and training.	**Rainbow House** Provides emergency shelter and transitional housing for homeless women and their children under twelve. **Rustic House** Provides emergency shelter and transitional housing for homeless mentally or physically disabled men and men over 60 years old.
Ad Hoc Group Against Crime Combats and prevents crime, violence and drug abuse by building individual and community self-esteem, pride and positive values through education, spirituality, morality, prevention and intervention, with the focus on youth and young adults in metropolitan Kansas City.	**Summer '94 Mentor Program** Provides mentoring program for mentees referred due to prior problems with violence and mentees who were ex-offenders, recovering drug addicts/ alcoholics, or victims and/or perpetrators of violent crimes.

Common Themes

➢ **The Environment is Changing**

Pressure from funders led three of the four evaluations profiled to be conducted. While some agencies described a growing interest in outcome measurement by funders, others mentioned actual funding requirements specified by grantmakers. Nevertheless, despite these evaluations being born out of necessity, each one profiled was perceived by its agency as extremely useful, credible, and satisfying.

Deborah Long, chief executive officer of the Girl Scouts of Genesee Valley said that "historically we've been given money to provide direct services as a not for profit organization and that's where our major concern has been. But the tide has turned in the last four years or so around responding to outcomes."

Through Metro Ministries of Corpus Christi's evaluation process, Executive Director Reverend Edward Seeger learned the importance of measuring outcomes. He grew to believe that, "Something needs to be different [in clients] to justify being in business and being on the receiving end of charitable dollars" and stated that the community is increasingly holding service providers to that standard.

➢ **Evaluations Were Used for a Variety of Purposes**

Regardless of the level of stakeholder involvement, all four of the profiled evaluations served multiple purposes. Among the most commonly mentioned purposes were raising funds, satisfying reporting requirements, providing insight into a program, refining programs, and documenting program successes.

The Salvation Army Hope Center for Children Center found its evaluation extremely useful because, as Administrator Michael Fitzgerald explained, it provided, "a tool for any funding stream." In addition, Mr. Fitzgerald stated that it "gives people better insight into what we are trying to do with kids."

The Girl Scouts found its evaluation useful because the agency gained the ability to measure outcomes in a way it felt was appropriate and in a way that would satisfy multiple funders. Ms. Long explained that prior to the evaluation, the Girl Scouts "had been teetering about what appropriate evaluation is and what methods were appropriate methods." The evaluation was also useful because it was a learning experience for staff. Ms. Long appreciated the opportunity for staff "to experience what evaluation is all about."

Metro Ministries found its evaluation useful because the organization could use it to document that their program "is working." Reverend Seeger was "able to go into the community and say we are effective stewards of your dollars. And we think we are efficient stewards of your dollars." The evaluation instilled staff with a sense of pride in a job well done and allowed Metro Ministries to define what well done meant.

The Ad Hoc Group Against Crime found its evaluation extremely useful because, as Vice President Vivian Jones stated, "it gave us ideas about what kinds of interventions are useful in working with high-risk youth. Specifically, for example, we found that one-on-one activities with a caring adult is very positive for youngsters." The evaluation was also useful because it confirmed for staff that mentoring is an effective intervention.

- **Successful Evaluation Efforts Are Manageable**

Two of the agencies strove to keep their evaluations very "simple," while two of the agencies employed designs that were more elaborate. Nevertheless, each agency found their evaluation manageable and worth the effort required to complete it.

The Hope Center found its evaluation useful because the process was neither long nor complicated. In fact, Mr. Fitzgerald explained that data collection, “is now rolled into treatment planning for children.”

By employing a participatory approach to evaluation and dedicating two staff members to spearhead the evaluation effort, the Girl Scouts ensured that a manageable evaluation design was created.

Metro Ministires described its evaluation as “simple and manageable.” The evaluation focused on only one outcome and it relied primarily on two data collection methods (document review and observation).

Ad Hoc enlisted the services of the Ewing Marion Kauffman Foundation to conduct its evaluation. Therefore, Ad Hoc was not faced with manageability issues related to the evaluation.

➢ **Perspectives Differ on Causes of Credibility**

Two agencies felt that the participation of an outside evaluator enhanced their effort's credibility. Mr. Fitzgerald believed that the credibility of the Hope Center's evaluation was augmented by having a professor's name on the report cover. Ms. Jones at Ad Hoc believed that staff, board members, and volunteers found the evaluation extremely credible since a foundation's research staff conducted it. Ms. Jones explained that having someone from the outside conduct the evaluation "brought a high level of objectivity therefore reducing potential bias."

Both of the agencies that conducted their evaluation internally believed their evaluations were perceived as highly credible. Reverend Seeger at Metro Ministries stated that he received no criticism for conducting the evaluation in-house. He cited the staff's integrity as well as their lack of incentive to "cook the numbers" since compensation is not tied to evaluation findings. Furthermore, Reverend Seeger believed that stakeholders found the evaluation highly credible because "people know it's an uphill battle do deal with an overwhelming social problem [i.e., homelessness] and no one expects us to succeed terribly well. People are pleased that we're helping some people who are homeless."

Some of the staff members interviewed felt that funders found their evaluations very credible rather than extremely credible. For example, the Girl Scouts felt that funders' impressions of credibility may have been influenced by their limited evaluation experience. Ms. Long explained that "Funders are struggling just like agencies are and I'm not sure they know, collectively, what are the most appropriate models of evaluation. They are mandating that we do evaluation and expecting us to do it and it's linked to funding dollars, but I don't have a clear sense that they know what they want." She added that "funders are not really clear yet about how you get to what works and what doesn't work.

For example, we haven't defined as a community what a healthy young person looks like or what a healthy family looks like."

Lessons Learned

1. Evaluation Efforts Can Be Integrated Into Programmatic Activities.

The Hope Center integrated data collection into regular job responsibilities. As Mr. Fitzgerald stated, "Before it was another thing to do, but now it's just a part of their jobs."

Metro Ministries noted that the evaluation process "helps staff 'notice' when someone is ready to go" (i.e., leave the transitional housing shelter) by having them consider whether the resident's case is considered successful as defined through their evaluation effort if they were to leave at that point.

2. An Evaluation Needs To Be Systematic, Not Fancy.

The Hope Center chose to "focus on things that were not real complicated but on things that we needed to do." Together, the staff articulated a definition of a successful residency and then created a process for measuring its success rate.

Based on training in participatory evaluation, program staff at the Girl Scouts designed and implemented their own evaluation. Without prior experience in evaluation, the staff members articulated outcomes and indicators along with targets. In addition, they involved program participants in survey design.

Metro Ministries' evaluation was based on a review of client records and observation by program staff. It contained neither advanced statistics, charts, graphs or illustrative quotes from clients. Nevertheless, the evaluation provided useful data that was systematically collected.

3. Enlisting An Outside Professional Evaluator May Add to an Evaluation's Credibility without Significantly Impacting Its Cost.

At the Hope Center, a board member led the evaluation effort. While this enabled the Hope Center to avoid paying for a consultant, it also allowed them to find a reputable person to "drive the evaluation, take some responsibility, and put their name on it." Since many nonprofits cannot afford to hire an outsider, they may do well to consider enlisting board members or other volunteers in their effort.

Ad Hoc approached a prominent local foundation, The Ewing Marion Kauffman Foundation, a supporter of other Ad Hoc programs, to conduct an evaluation of its Mentoring Program. The foundation assigned a research team that conducted the evaluation almost entirely alone.

4. Outcome Focused Evaluations Are Receiving Considerable Community Attention.

There was substantial interest in the outcome measurement work done at the Hope Center. In fact, Mr. Fitzgerald has been surprised by the attention his evaluation received from staff, funders, and outsiders. "I knew it was coming, and I knew it was important but [I'm surprised] that there's such a high degree of interest in the community."

The Girl Scouts' participation in the Rochester Effectiveness Partnership led to a great deal of recognition for the Operation Stephne program. In fact, participation combined with the outcomes on which it focused, enabled the Girl Scouts to "get in with violence prevention" and to "position [itself] on some of the critical issues affecting girls."

5. The Timing of an Evaluation Can Affect Its Impact and Uses.

The Girl Scouts found it useful to design its evaluation and program at the same time. Ms. Long explained that "you really have to give more careful thought to what you're trying to accomplish" in a program by designing a program and evaluation simultaneously. She added, "it is easy to cream and to get into all of those things that really don't prove outcomes if you wait too long down the road and you want to say something positive to your funders . . . . If you build it in up front, then you're more prepared to have a solid evaluation plan to move forward in proving your outcomes."

Ad Hoc had its evaluation conducted by outsiders after its first year of program operations. Evaluating the program early in its life cycle equipped Ad Hoc with documentation of a successful intervention to aid in fundraising. It also provided recommendations for mid-course program improvements.

V. Conclusion

At a time when program evaluation is evolving quickly, this study presents a window into recent and current evaluation practice in the nonprofit sector. The increased awareness, pressure and resources for program evaluation have created new opportunities and possibilities for creating useful, credible and satisfying efforts that were unimaginable just a few years ago.

This study will aid evaluators and evaluation practitioners in the future design and implementation of their evaluation efforts. It is clear that an increasing number of agencies will be undertaking evaluation efforts due to more stringent funding requirements and critical public scrutiny. This report will aid these efforts by providing a glimpse of how and why some practitioners had positive evaluation outcomes.

Our analysis indicates that the participation of stakeholders in a variety of ways increases the satisfaction, and improves the design and usefulness of the evaluation. The study also articulates that a useful evaluation is one that begins with a focused evaluation design, provides recommendations for program improvement, documents success, and contributes to an agency's planning effort. While findings from our sample cannot be generalized to the entire nonprofit sector, they provide considerable insight for nonprofit agencies and program evaluators into ways to plan and implement evaluations that are likely to result in positive outcomes of usefulness, satisfaction and credibility.

In addition to providing insight into evaluation practice, this study suggests several areas for further consideration. In particular, it raises critical questions about the role of stakeholder involvement:

- Why do some organizations involve stakeholders in their evaluations? What are their expectations, and to what extent are these expectations met? Why or why not?
- What is the quality of stakeholder participation? What exactly do stakeholders do and how does this involvement affect the evaluation?

- In specific evaluations, what were the benefits of or value added by involving stakeholders? What were the disadvantages of involving stakeholders?

- What are some of the barriers to involving stakeholders in evaluations? What are the costs associated with conducting a highly participatory evaluation? Are agencies being required to conduct evaluations without compensation likely to conduct participatory evaluations?

Our study also suggests the need to investigate the influence funding agencies have on agencies' evaluations. Specifically:

- What influence do funding agencies have on an organization's choice of evaluation approach and design, investigated outcomes, the way the evaluation is used and the organization's perceptions of the evaluation?

Appendices

Appendix I

Advisory Committee Members

Alan Abramson
Director of the Nonprofit
Research Fund
Aspen Institute

Sylvia Barsion
Director of Research and
Evaluation
Girl Scouts of the U.S.A.

Janet Carter
Saugerties, NY

Charles Firke
Director of Research and
Evaluation
YMCA

Jennifer Greene
Associate Professor
Human Service Studies
Cornell University

Martha Taylor-Greenway
Vice President of Impact
Measurement
United Way of
Metropolitan Atlanta

Barbara Kehrer
Barbara Kehrer Consulting
Palo Alto, CA

Kathryn Newcomer
Department Chair
Graduate School of Public
Administration
George Washington University

Karen Simmons
Director of Nonprofit
Management Development
Center
La Salle University

David Stevenson
Director, National Center for
Charitable Statistics
Urban Institute

Appendix II

Data Collection Instruments

1. Mail Survey Questionnaire
2. Telephone Interview Schedule
3. In-depth Profile Interview Protocol

InnoNet

A Study for the
Robert Wood Johnson Foundation
&
The Aspen Institute

Robert Wood Johnson Foundation and the Aspen Institute
SURVEY OF EVALUATION USE IN NONPROFIT ORGANIZATIONS

Would you like to know how to ensure that your organization's future evaluations are useful and lead to positive benefits for your programs? So would we!

The Robert Wood Johnson Foundation and the Aspen Institute have commissioned Innovation Network, Inc. (InnoNet) to conduct a study to determine the factors that make evaluations by nonprofit organizations useful. Please help us by completing the survey below. As thanks, you will receive:

- automatic entry into a random drawing for $100
- a reprint of the 1997 article, "Outcome Measurement: Showing Results in the Nonprofit Sector"
- a copy of the results of this study.

Your participation is voluntary and all responses will remain confidential. Completing the survey should take only about 15 minutes of your time. Please return the questionnaire in the envelope provided to InnoNet, 1001 Connecticut Ave., NW, Suite 1138, Washington, DC, 20036 by February 2, 1998. If you have any questions, please call Ana Coghlan at (202)728-0727. Your participation is greatly appreciated!

Your Name: ______________________________

Title: ______________________________

Organization: ______________________________

Address: ______________________________

______________________________ Phone: ______________

City State Zip Code

Section I. To begin, we have a few questions about your organization.

1. What sorts of services does your organization provide? (Please check all that apply.)
   - ❑ Arts, culture, and/or humanities
   - ❑ Education/schools
   - ❑ Environment/animals
   - ❑ Health, mental health
   - ❑ Human Services, specifically:
     - ❑ Children and youth services
     - ❑ Community development
     - ❑ Crime, legal related
     - ❑ Domestic violence
     - ❑ Economic development
     - ❑ Employment, job related
     - ❑ Family services
   - ❑ Food, agriculture, and nutrition
   - ❑ Homelessness
   - ❑ Housing, shelter
   - ❑ Immigration
   - ❑ Public safety, disaster preparedness
   - ❑ Recreation, sports
   - ❑ Senior services
   - ❑ Women's issues
   - ❑ Youth development
   - ❑ Other ______________________________
     ____
     ____

2. Of those checked above, please underline the **one** primary service your organization provides.

3. How many full-time staff does your organization have? ________________

4. What is the approximate annual operating budget of your organization? $ _____________

5. Are any staff members responsible for conducting program evaluation as part of their regular duties?

   Yes No

   If yes, how many staff members devote at least 50% of their time to evaluation? ______________

6. Does your organization budget funds annually to support evaluation?

   Yes No

   If yes, what is the approximate annual allocation? $ ___________________

Section II. We are interested in knowing how many formal program evaluations your organization has either conducted or undergone. By program evaluation we mean a one-time, systematic assessment of the operations or results of any program or service your organization delivers, which included an overall design, data collection and analysis, and reporting of results.

7. How many program evaluations has your organization either conducted or undergone during the last three years? __

8. Of the evaluations initiated, how many were completed? _____ *(If your answer is "0," you are finished with this survey. But please return this form in the envelope provided to InnoNet, 1001 Connecticut Ave., NW, Washington, DC, 20036)*

We now want to ask several questions about your organization's most recent fully completed evaluation.

9. Please briefly describe the program that was evaluated. ______________________________

____

____

10. What was/were the main purpose(s) of the evaluation? (Please check all that apply.)
- ❑ To measure the program's outcomes or impact on participants/clients
- ❑ To describe program operations
- ❑ To describe program participants/clients
- ❑ To measure costs and benefits of the program
- ❑ To assess the implementation of the program
- ❑ To measure participant/client satisfaction
- ❑ To assess the quality of program operations
- ❑ For organizational strategic planning
- ❑ Other? (please describe) ____

  ____

11. Of those listed above, please underline the **one** primary purpose of the evaluation.

12. For whom was the evaluation primarily done? (Please check all that apply.)
- ❑ The governing board of the organization
- ❑ Current program funders
- ❑ Potential funders
- ❑ Program staff
- ❑ Program participants/clients
- ❑ Volunteers
- ❑ Governmental (state, county and/or federal) agency(ies)
- ❑ Other nonprofit organizations
- ❑ Other (please describe) ____

  ____

13. What methods were used to collect the information? (Please check all that apply.)

- ❑ Review of program documents
  If yes, please check specific documents reviewed:
  - ❑ Program description/curriculum
  - ❑ Internal participant/client records
  - ❑ Internal financial records
  - ❑ Other documents (please describe) ____

- ❑ Mail questionnaires
  If yes, please check the type of questions asked:
  - ❑ Closed-ended questions
  - ❑ Open-ended questions
  - ❑ Other (please describe) ____

- ❑ Face-to-face or telephone interviews
  If yes, please check the type of questions asked:
  - ❑ Closed-ended questions
  - ❑ Open-ended questions
  - ❑ Other (please describe) ____

- ❑ Observation of program staff/participants/activities, etc.
  If yes, please check how observations were done:
  - ❑ Checklists used by observers
  - ❑ Staff members conducted the observation

- ❑ Focus group(s)/group discussion(s)
- ❑ Tests of knowledge or skills
- ❑ Review of existing research from other programs
- ❑ Use of a control or comparison group which did not receive program benefits/services
  If yes:
  Was this group randomly assigned? Yes No Don't know
- ❑ Other (please describe) ____
  ____

14. Who had the primary responsibility for conducting the evaluation?
- ❑ Internal staff
- ❑ Outside contractor
- ❑ Volunteers
- ❑ Other (please describe) ____

15. We would like to know how <u>program stakeholders</u> participated in the evaluation. Please check all activities in which each group participated.

Evaluator
- ❑ Planning/designing the evaluation
- ❑ Collecting data
- ❑ Analyzing/interpreting results
- ❑ Writing report
- ❑ Receiving report
- ❑ Attending briefing on results

Program Staff
- ❑ Planning/designing the evaluation
- ❑ Providing data
- ❑ Collecting data
- ❑ Analyzing/interpreting results
- ❑ Writing report
- ❑ Receiving report
- ❑ Attending briefing on results

Board Members
- ❑ Planning/designing the evaluation
- ❑ Providing data
- ❑ Collecting data
- ❑ Analyzing/interpreting results
- ❑ Writing report
- ❑ Receiving report
- ❑ Attending briefing on results

Funder's Staff
- ❑ Planning/designing the evaluation
- ❑ Collecting data
- ❑ Analyzing/interpreting results
- ❑ Writing report
- ❑ Receiving report
- ❑ Attending briefing on results

Program Participants/Clients
- ❑ Planning/designing the evaluation
- ❑ Providing data
- ❑ Collecting data
- ❑ Analyzing/interpreting results
- ❑ Writing report
- ❑ Receiving report
- ❑ Attending briefing on results

Volunteers (not program clients)
- ❑ Planning/designing the evaluation
- ❑ Providing data
- ❑ Collecting data
- ❑ Analyzing/interpreting results
- ❑ Writing report
- ❑ Receiving report
- ❑ Attending briefing on results

16. Which of the following were included in the evaluation report and/or presentation? (Please check all that apply.)
- ❑ Illustrative quotes or comments from respondents
- ❑ Stories or case examples
- ❑ Basic statistics (counts, percentages, averages)
- ❑ Advanced statistics (correlation, analysis or variance, regression, factor analysis)
- ❑ Charts or graphs
- ❑ Recommendations for future action
- ❑ Other (please describe) ____

17. About how much money did this evaluation cost when completed? (Total should not include your organization's staff time.)

____

Section III. Now we would like to ask you about the benefits of the same evaluation you have described above, i.e. your organization's most recently completed evaluation.

18. Please rank the extent to which your organization found the evaluation to be useful on a scale of 1 to 5 where 1 denotes not useful at all and 5 denotes extremely useful.

Not useful at all				Extremely useful
1	2	3	4	5

19. How have the evaluation process and/or results been used? (Please check all that apply.)
- ❑ To change the program design or operation
- ❑ To improve the outcomes or impact of the program
- ❑ To report to a funder
- ❑ To respond to questions or criticisms about the program
- ❑ To promote the program to potential participants/clients
- ❑ To get additional funding
- ❑ To plan future programs
- ❑ To train or reassign staff
- ❑ To decide resource allocation within the organization
- ❑ To improve outreach and public relations
- ❑ To design an ongoing monitoring/evaluation process
- ❑ To gain support for evaluation among staff or governing board
- ❑ Other (please describe) ____

____

20. Have any of the recommendations offered in the evaluation report been implemented?

Yes

No

21. On a scale of 1 to 5, how credible do you think the following program stakeholders found the results of the evaluation?

	Not credible at all				Extremely credible	Don't know	Not applicable
Board members	1	2	3	4	5	DK	NA
Funders	1	2	3	4	5	DK	NA
Program staff	1	2	3	4	5	DK	NA
Program participants	1	2	3	4	5	DK	NA
Volunteers	1	2	3	4	5	DK	NA

22. In general, on a scale of 1 to 5, how satisfied would you say each of the following stakeholders were with the evaluation?

	Not satisfied at all				Extremely satisfied	Don't know	Not applicable
Board members	1	2	3	4	5	DK	NA
Funders	1	2	3	4	5	DK	NA
Program staff	1	2	3	4	5	DK	NA
Program participants	1	2	3	4	5	DK	NA
Volunteers	1	2	3	4	5	DK	NA

23. Looking back on the evaluation described above, are there things your organization wishes had been done differently?

Yes

No

If yes, what? ____

____

____

THANK YOU VERY MUCH FOR PARTICIPATING IN THIS SURVEY!

Please add on the back side of this page any comments you have that would help us better understand the type of evaluation conducted and its benefits and/or shortcomings.

Please return your completed survey in the enclosed envelope by February 2, 1998 to InnoNet, 1001 Connecticut Ave., NW, Suite 1138, Washington, DC, 20036.

Optional

Please use the space below to add any comments that might help us better understand the type of evaluation your organization conducted and its benefits and/or shortcomings. Thank you.

Telephone Interview Protocol
Survey of Evaluation Use in Nonprofit Organizations

May 10, 1998

Respondent's Name: ____ ID No. ____________

Title: ____________________________ Organization: ________________________________

Phone No.: ______________________

Call Record				
Interviewer	Date	Time	Person Contacted (Resp. or Other)	Result
1.				
2.				
3.				
4.				
5.				
6.				

Introduction:

Hello, my name is _______________ and I'm calling from Innovation Network in Washington, D.C. We sent you a letter last week asking if you would be willing to be interviewed on the phone as a follow-up to a mail survey you completed several months ago. If you can recall, that survey asked you about one of your recently completed evaluations. I'm calling today to learn more about your organization's current evaluation efforts. Do you have about 10 minutes now to answer a few questions?

Yes, let's begin 1
That's great. Before we begin, please understand that your participation is completely voluntary and everything you tell me will be kept confidential. If I come to any question that you prefer not to answer, just let me know and I'll skip over it.
Okay? PROCEED WITH INTERVIEW BELOW

No, I don't have time right now 2
All right, I understand that this isn't a good time. Is there a time I should call back?
Date and time to call back: __________________ → WRITE DATE AND TIME ABOVE IN "RESULT"

No, I do not wish to participate . . . . . 3 → STOP THE INTERVIEW AND THANK THEM FOR THEIR EARLIER PARTICIPATION

Section 1: **First, I'd like to ask you about current evaluation practices in your organization.**

1. **Does your organization have any program evaluations underway at this time? (***By program evaluation I mean a one-time, systematic assessment of any program or service your organization delivers. In addition, by program evaluation I mean an effort that includes an overall design, data collection and analysis, and reporting of results.)*

   No . . . . 0 → IF NO, SKIP TO Q. 9
   Yes 1

2. **How many evaluations does your organization currently have underway**? ________
   (By current, I mean an evaluation that is at least designed but is not yet completed.)

3. Please tell me briefly what the current evaluation(s) is (are) about. IF THE ORGANIZATION HAS MORE THAN THREE EVALUATIONS UNDERWAY, ASK THE RESPONDENT TO REPORT ON ONLY THREE. THEY ARE FREE TO CHOOSE WHICH THREE TO REPORT ON.

Evaluation #1: ____

____

____

Evaluation #2: ____

____

____

Evaluation #3: ____

____

____

I would now like to ask more specific questions about your current evaluation(s).

4. **Which of the following is (are) the primary purpose(s)of the current evaluation(s)?** *(If you have more that one current evaluation, please state one primary purpose for each.)* READ THE LIST BELOW AND CIRCLE ALL THEY SAY YES TO.

   To measure program outcomes or impact 1
   To assess program implementation 2
   For strategic planning 3
   To assess quality of operations. 4
   To measure client satisfaction. 5
   Other ______________________________________ 6 (WRITE IN)
   Don't know. 7
   Refuse to answer. 8

5. **Who has the primary responsibility for conducting the current evaluation(s)? Internal staff, an outside contractor, or a combination of both?** *(If respondent says combination or both, ask him or her if they mean a combination in the same evaluation or some evaluations are internal and some are external.)* DO NOT READ RESPONSE CATEGORIES. CIRCLE ONLY ONE ANSWER.

   Internal staff only 1
   External contractor only 2
   Combination of internal and external in
   same evaluation(s). 3
   Some evaluations are internal and
   some are external 4
   Other __________ ____________________________ 5 (WRITE IN)
   Don't know. 6
   Refuse to answer. 7

6. **What types of data collection methods are being used in the current evaluation(s)? Quantitative, qualitative, or a combination of both? By quantitative I mean methods that result in numbers such as closed-ended surveys and numeric tests. For qualitative I mean such methods as open-ended interviews, observation of activities, and focus groups**. *(If respondent says combination or both, ask them if they mean a combination in the same evaluation or quantitative in some evaluations and qualitative in others.)* DO NOT READ RESPONSE CATEGORIES. CIRCLE ONLY ONE ANSWER.

   Quantitative methods only 1
   Qualitative methods only 2
   Mix of quantitative and qualitative methods in
   same evaluation(s). 3
   Some evaluations will use quantitative methods
   and some will use qualitative methods 4
   Other ______________________________________ 5 (WRITE IN)
   Don't know. 6
   Refuse to answer. 7

7. **Are program stakeholders involved in designing and conducting the current evaluation(s)? By "stakeholders" I mean staff, board members, funders and/or program participants.** DO NOT READ RESPONSE CATEGORIES. CIRCLE ONLY ONE ANSWER.

No. . . . . 0 → IF NO, SKIP TO Q.9
Yes 1
Stakeholders are involved in some evaluations
and not in others 2
Don't know . . . . .3→ IF DK, SKIP TO Q. 9
Refuse to answer 4

8. **In general, to what extent are stakeholders involved in designing and conducting the current program evaluation(s)? A little, some or a lot?** DO NOT READ RESPONSE CATEGORIES. CIRCLE ONLY ONE ANSWER.

A little 1
Some. 2
A lot 3
Will vary by evaluation 4
Don't know 5
A. Refuse to answer. 6

~~~~~~

In the mail survey, there was a high occurrence of <u>outcome-focused evaluation</u>. I would like to explore this a bit with you.

9. Is your organization currently measuring or monitoring program outcomes? By program outcomes I mean changes in program beneficiaries due to the program. ***(This measuring of outcomes may or may not be part of a current evaluation you have already described.)***

No . . . . 0 → IF NO, ASK Q. 10, THEN SKIP TO SECTION II, p. 6
Yes 1
Don't know 2

10. **As you may know, United Way has been encouraging organizations to conduct outcome measurement. Is your organization a member of the United Way?**

No 0
Yes 1
Don't know 2
Refuse to answer . .3
~~~~~~

IF RESPONDENT IS NOT CURRENTLY MEASURING PROGRAM OUTCOMES (I.E., SAID NO TO QUESTION #9), SKIP TO SECTION II, P. 6

11. Has your organization been influenced by the United Way to measure outcomes?

No 0
Yes 1
Don't know 2
Refuse to answer . .3

12. In what year did your organization start measuring or monitoring program outcomes?

___________ [WRITE IN YEAR]
Don't know 2
Refuse to answer 3

13. What program outcomes are you measuring or monitoring?

14. Why did your organization start measuring outcomes?

~~~~~~

Section II. **In this last section, I would like to ask you a few questions about what makes a good evaluation.**

~~~~~~

15. In general, what have you found makes an evaluation useful to your organization?

Have you found that the primary purpose makes a difference to an evaluation's usefulness? (Why, how, what primary purpose do you think is most useful?)

Have you found who conducts an evaluation to make a difference to its usefulness? By who conducts, I mean internal staff, an external contractor or a combination of both.

Have you found that stakeholder involvement makes a difference to an evaluation's usefulness?

16. In general, what do you think makes an evaluation credible to program stakeholders? By "credible" I mean believable or trustworthy.

Have you found that the primary purpose of an evaluation makes a difference to an evaluation's credibility to stakeholders? (Why, how, what primary purpose do you think is most credible?)

Have you found who conducts an evaluation to make a difference to its credibility to stakeholders? By who conducts, I mean internal staff, an external contractor or a combination of both.

Have you found that stakeholder involvement makes a difference to an evaluation's credibility?

17. **For the last question, in general what do you think makes an evaluation satisfying to stakeholders?** *(By satisfying, I mean fulfilling stakeholders' information needs and expectations.)*

Thank you very much for participating in this interview!
Innovation Network will send you a copy of the study's results early next fall.

In-depth Profile
Interview Protocol

Hello, I'm _____________ calling from Innovation Network. Thank you for agreeing to answer some questions for us. Let me begin by reminding you of the reason for my call. We are studying approaches to evaluation in the nonprofit sector, and we have selected a few organizations to interview that self-reported extremely high ratings for its evaluation's usefulness, credibility, and satisfaction. Through this interview, we hope to gain an understanding of the process of your evaluation as well as to understand why your evaluation had such positive outcomes.

Is this still a good time for you to answer some questions for me? If so, let's begin.

Process

1. What lead your agency to evaluate your _______________ program?

2. What was your agency's previous experience with evaluation?

Stakeholder Involvement

3. Who was involved in designing the evaluation? How were they involved?
   Probe for staff, board members, program participants/clients, and funders

4. Who was involved in collecting the data? How were they involved?
   Probe for staff, board members, program participants/clients, and funders

5. Who was involved in analyzing/interpreting the data? How were they involved?
   Probe for staff, board members, program participants/clients, and funders

6. Who was involved in writing the report? How were they involved?
   Probe for staff, board members, program participants/clients, and funders

Explanation of usefulness, credibility, and satisfaction

7. Can you tell me, in your opinion, what made the evaluation so highly useful?

8. What do you think made the evaluation so credible to staff? Why (or why not) did funders and board members find it equally credible?

9. What do you think made the evaluation so satisfying to staff? Why (or why not) did funders and board members find it equally satisfying?

Use of results

10. Can you tell me how the results of the evaluation were used?

11. Have you made any changes to your program or organization based on the evaluation results?

12. Did your organization change any evaluation practice as a result of the evaluation?

Lessons Learned

13. What surprised you the most about the evaluation?

14. What advice would you give to other groups who are interested in conducting a similar evaluation?

Those are all of the questions I have for you. We appreciate your taking the time to answer them and look forward to sharing the report with you in the Fall.

Appendix III

Respondent Characteristics

Percentage of Survey and Interview Respondents by Organizational Characteristics

Organizational Characteristics	Survey	Interview
Primary Service	(N=134)	(N=37)
Human Services	67%	60%
Health, Mental Health	11	11
Education	9	8
Arts, Culture, Humanities	3	5
Environment, Animals	2	3
Other	8	14
Full-time Equivalent Staff	(N=140)	(N=40)
< 25	49	48
26 – 50	19	13
51 – 150	18	30
> 151	14	10
Annual Operating Budget	(N=140)	(N=40)
< $1,000,000	39	40
$1,000,000 – 10,000,000	47	48
> $10,000,000	14	13
Have Evaluation Staff	(N=140)	(N=40)
Yes	81	65
Those who have evaluation staff, # of staff who devote 50% of time to evaluation	(n=110)	(n=25)
0	61	60
1 – 5	37	32
> 6	2	8
Have Budget for Evaluation	(N=140)	(N=40)
Yes	51	51
Amount for those who have Annual Evaluation Budget	(n=61)	(n=15)
< $5000	20	20
$5001 – 10,000	16	7
$10,001 – 20,000	21	33
$20,001 - 50,000	20	13
> $50,001	23	27

Appendix IV

Evaluation Characteristics

Survey Respondents by Evaluation Characteristics for Most Recently Completed Evaluation

Evaluation Characteristics	Percent of Respondents
Primary Purpose	N=140
To measure outcomes/impact	56%
To assess implementation	9
For strategic planning	9
To assess quality of operations	7
To measure client satisfaction	4
Other	8
No primary purpose designated	7
For Whom was Evaluation Done [a]	N=140
Current funders	69%
Program staff	61
Board	56
Potential funders	44
Government agencies	28
Program Participants/Clients	24
Other nonprofit organizations	9
Volunteers	6
United Way (written in)	4
National Office (written in)	3
Other	9
Prime Responsibility for Conducting Evaluation	n=131
Internal staff	51%
External contractor	40
Combination of internal and external	9
Data Collection Methods	n=132
Qualitative only	19%
Quantitative only	14
Mix of quantitative and qualitative	67

[a] Multiple responses allowed

Appendix V

Types of Stakeholder Involvement by Stakeholder Participation Level

Survey Respondents by Type of Stakeholder Participation and Level of Stakeholder Involvement

Ways Stakeholders Participated in Evaluation	All (N=140) 100%	Low (n=58) 41%	Medium (n=56) 40%	High (n=26) 19%
Staff				
Planning/designing	61%	40%	75%	77%
Collecting data	79	66	82	100
Analyzing/interpreting	39	9	55	73
Writing report	29	3	38	65
Board				
Planning/designing	16%	3%	14%	46%
Collecting data	5	2	7	8
Analyzing/interpreting	18	7	14	50
Writing report	4	0	2	15
Participants				
Planning/designing	6%	0%	7%	19%
Collecting data	12	3	9	39
Analyzing or interpreting data	4	2	2	15
Writing report	2	0	4	4
Funder				
Planning/designing	21%	3%	27%	46%
Collecting data	7	2	5	23
Analyzing/interpreting	11	0	9	39
Writing report	6	0	5	23

Appendix VI

In-depth Profiles

The Salvation Army Hope Center for Children

Michael J. Fitzgerald
Administrator
The Salvation Army Hope Center for Children
3740 Marine Avenue
St. Louis, MO 63118

Annual budget:	$ 3,000,000
Annual evaluation budget:	$ 0
Staff:	100 fte
Stakeholder participation level:	High

Description of Organization

The Salvation Army Hope Center for Children offers six interrelated child and family programs that provide prevention and treatment services for abused and neglected children and their families. The Hope Center serves approximately 150 children from birth to age twelve each day and has a staff that consists of social workers, nurses, family therapists, educational specialists, experienced child care workers, and a consulting child psychologist and pediatrician. The six programs are: Residential Treatment, Soulard School, Day Treatment, Day Care, Family Therapy/Parenting Education, and Therapeutic Foster Care and Adoption Program.

Evaluated Programs

The Hope Center conducted an evaluation of four of its programs. The programs evaluated were: Residential Treatment, Day Treatment, Day Care, and Therapeutic Foster Care and Adoption.

The Residential Treatment program is a forty-bed inpatient program, serving children whose abusive environments have created severe social, emotional, behavioral and developmental problems.

The Day Treatment program is a structured program for children ages two to six with emotional, behavioral, or developmental disorders related to a history of abuse or neglect. The program includes pre-school curriculum, behavior management program, play therapy, family therapy and support services.

The Day Care program serves children ages twelve months to six years who typically come from low-income and single parent families.

The Therapeutic Foster Care and Adoption Program recruits, trains and licenses families through a unique program designed by the Hope Center. The program matches children with families and provides ongoing support services.

Purpose and Focus of the Evaluation

The Hope Center conducted an outcome focused evaluation to assess placement stability, child functioning, and child safety in four of its programs. According to the Hope Center's survey, in addition to measuring outcomes, the evaluation's purpose was to measure client satisfaction and assess the quality of program operations.

The Hope Center's primary incentive for conducting an evaluation was to comply with United Way requirements. As Hope Center administrator Michael J. Fitzgerald explained, "All of the 134 United Way agencies in the St. Louis area were told that they need to do this as part of the funding process."

Previously, the Hope Center did not evaluate outcomes. Instead, they reported statistics such as the number of children served by each program. Mr. Fitzgerald stated that the Hope Center viewed the prospect of measuring outcomes as a positive challenge.

Evaluation Procedures

Management staff and a board member designed and implemented the evaluation. Mr. Fitzgerald met with the management staff to share what he learned at a United Way training and to discuss long range goals as well as what the evaluation should measure. With emphasis on designing an evaluation that made sense to staff, "it became a team process." The board member who led the evaluation effort is a Washington University professor of social work who is "a guru in research and numbers."

To assess stability of placements, staff obtained statistics from the Division of Family Services in Jefferson County on Hope Center placements and comparable placements in the St. Louis metro area. The board member then calculated the percentage of children achieving stable placements (no additional state placements) from foster care and residential care programs.

To measure child functioning in the Daycare, Day Treatment, and Foster Care programs, staff evaluated the children every three months using a 33-item instrument designed to assess the child's cognitive, social, and emotional development. The instrument was previously designed by a Master's student in 1994. For children in daycare and foster care programs, the board member calculated the mean net change in cognitive, social, and emotional scale scores for a six month and a year-long interval. There were not enough data to analyze children in the day treatment program.

Also to measure child functioning, Hope Center staff tracked twenty recent twenty graduates from the Daycare program and obtained their Missouri Kindergarten Inventory of Developmental Skills: (MO-KIDS) score which is a test that measures their preparedness for school. Again, this data was compared with the St. Louis area Early Entry program passing rate.

Additional data was gathered by staff on the St. Louis public schools Early Childhood Screen for Daycare program graduates entering non-magnet schools. The frequency of screens receiving a score in the "refer" category (i.e. "failed") were calculated.

To assess the appropriateness of behavior among children in the residential treatment program, a 43-item five-point scale instrument was employed. The mean net change in child functioning was calculated for a six-month and year-long period.

To measure child safety, data on accidents and injuries for children in the Day Treatment, Day Care, and Residential Treatment programs were recorded by staff using a five-point scale. The scale consisted of the following categories: observation, minor injury, moderate injury, out-patient hospitalization, and in-patient hospitalization. The board member calculated the frequency of each type of accident/injury over a seven month period.

In addition, staff tracked the frequency and duration of physical restraints performed by staff. Restraints are employed only as a last resort if a child is in danger of inflicting harm on himself, others or property.

To assess foster home recruitment, the staff obtained data from the City of St. Louis on the number of its licensed foster homes in 1995 and 1997. Staff also tracked the number of its own licensed foster homes in 1995 and 1997. The board member compared these data.

The report, written by the board member, included an executive summary, overview, analysis of outcomes (i.e., placement stability, child-functioning, child safety, and foster home recruitment) that included a narrative as well as bar charts, conclusion, and plans for future research.

Program participants and funders were not involved in conducting the evaluation.

Uses of Evaluation Results and Processes

The Hope Center used the evaluation findings in a variety of ways including:

- To improve the outcomes or impact of the program,
- To report to a funder,
- To respond to questions or criticisms about the program,
- To promote the program to potential participants,
- To plan future programs,
- To improve outreach and public relations,
- To design an ongoing monitoring/evaluation process,
- To report to the board,
- To report to the corporate and national offices,
- To strengthen proposals,
- To show staff they are doing a good job,
- To inform parents and clients,
- To train new foster care parents,
- To recruit families for the programs, and
- To make adjustments to current programs.

Mr. Fitzgerald explained that the evaluation is “A valuable tool in the recruitment of foster parents because we have hard facts saying that this is a successful program. [Parents should] come be a foster parent for us because we do a good job.”

An example of how the evaluation has been used to plan future programs can be seen with the Therapeutic Foster Care Program. Mr. Fitzgerald used information from the evaluation when he sought approval from the board and corporate office to expand the staff size. "I showed the board and corporate Salvation Army that we're doing a good job at this. Therefore, that's one more reason that they should allow me to do it." The program expanded from two to five full-time workers in the last year and a half.

The Girls Scouts of Genesee Valley, Inc.

Deborah Long
Chief Executive Officer
Girl Scouts of Genesee Valley, Inc.
1020 John Street
West Henrietta, NY 14586

Annual operating budget**:**	$ 3,000,000
Annual evaluation budget:	$ 0
Staff:	38 fte
Stakeholder participation level:	High

Description of Organization

The Girls Scouts of Genesee Valley, Inc. is a nonprofit youth serving organization. Its staff and volunteers corps of 4,000 serve 13,700 female youth members.

Program Evaluated

Operation Stephne is designed to provide middle school girls with conflict resolution and anger management skills. The participating girls are expected to attend weekly small group meetings, participate in workshops for skill and knowledge building, field trips which promote the development of social/life skills, and design and implement a community education project. Students with a minimum of three disciplinary incidents are referred to Operation Stephne by the school guidance counselor.

The program was designed in cooperation with Frederick Douglass Middle School and the Catholic Youth Organization.

Purpose and Focus of the Evaluation

The purpose of the evaluation was to determine the impact and effectiveness of Operation Stephne during its first year. Specifically, the evaluation was designed to answer the following questions:

1. How many of the 25 participants continued throughout the duration of the 10-month program period? What are the main reasons for attrition?

2. How and to what extent has the girls' participation in the program enabled girls to design a quality community education program/quality service project focusing on violence reduction?

3. What impact has Operation Stephne had on program participants?

    a. How have the anger management and conflict management skills of participating girls changed?

    b. What percentage of girls achieve a score of 80% or better on a post-test evaluation of learning?

    c. How have the number of overall school incidents changed overall at the target schools, and how have they changed for participating girls?

The evaluation of Operation Stephne followed a participatory model that staff members learned through their involvement in the Rochester Effectiveness Partnership (REP). REP began in 1996 as a two-year pilot project that brought together funders, evaluators, and nonprofit human service agencies seeking to determine and improve the effectiveness of their work. During the REP project, staff attended hands-on training sessions during which they learned about and practiced evaluation planning techniques and methodologies and designed a comprehensive evaluation of Operation Stephne. In addition to releasing two staff members for training, the Girl Scouts were required to have a leadership staff member participate on a Governance Committee. The Girl Scouts submitted an application to participate in the REP and received compensation for their involvement.

Chief Executive Officer Deborah Long explained that her agency wanted to join the partnership "…because more and more funders at that point, two to two and a half years ago, were requiring that we report on outcomes and results." Ms. Long added that "When your funding base is…asking for outcomes, it seems important to me that we need to be thinking as an organization about what the most efficient way is to report outcomes."

The Girl Scouts has a history of collecting qualitative data through market research, focus groups, and self-report surveys. However, their experience in carrying out full scale program evaluation was limited prior to joining REP.

Evaluation Procedures

Staff spear-headed the evaluation under the guidance of an evaluator who provided evaluation training and technical assistance throughout the evaluation effort. To begin, the staff created a logic model for Operation Stephne, selected the evaluation strategies, and designed the data collection instruments with some assistance from program participants.

Staff collected data to answer the evaluation questions through pre- and post-program surveys of girls, post-program survey of adults, observations, interviews, review of participant portfolios, school records, attendance records and other documentation related to the participation of girls in the program. The evaluation report organized findings by data collection method, rather than by evaluation question, and it included statistics as well as quotes from respondents.

Program participants were involved in designing questions for a survey of adults (school personnel or parents) that asked about behavior changes in the girls. They were also responsible for administering the survey, which entailed having an adult complete and sign it and having the youth return it to the program coordinator. Staff members were solely responsible for data analysis and report writing.

Board members were not involved in the evaluation design, implementation, or analysis since program evaluation falls outside the purview of governance.

Funders of REP participated in governance meetings as well as other training and project evaluation activities in addition to providing financial support. Girl Scout funders, however, did not participate directly in the design or implementation of the evaluation of Operation Stephne.

The report included an introduction (containing an overview of REP, a statement of program need, and a description of program design), evaluation design, evaluation strategies, findings, conclusions, and next steps, as well as recommendations and issues for further consideration.

Uses of Evaluation Results and Processes

The Girl Scouts reported using the evaluation for a variety of reasons, including the following:

- To change the program design or operation,
- To improve the outcomes or impact of the program,
- To report to a funder,
- To plan future programs, and
- To design an ongoing monitoring/evaluation process.

The two staff members involved in REP used their newly acquired knowledge and experience to lead training for other Girl Scout staff, including the executive director, on use of the logic model and on participatory evaluation. In fact, the Girl Scouts are interested in using a participatory evaluation process with all of its other projects since numerous funders, including the United Way, are interested in outcomes.

The evaluation has been used for fundraising by Ms. Long who has been “beating the streets” since the evaluation report was written. In particular, Operation Stephne has been awarded funding from the Youth Bureau for three months to increase the project staff from part-time to full-time. The Bureau will consider an application from the Girl Scouts in their next three-year funding cycle for Operation Stephne. In addition, the school district has approached the Girl Scouts about the possibility of building Operation Stephne into the mediation process in schools.

Programmatic changes were made based on evaluation findings. For example, girls expressed an interest in meeting longer than the 16-week cycle. The program was changed to enable the girls to continue meeting with adult leadership around issues of interest to them.

The evaluation findings and the attention received from participating in REP have given the Girl Scouts recognition for dealing with serious issues affecting girls. Ms. Long stated that the community is “recognizing the Girl Scouts not just for cookie sales, which is a great thing, or Girl Scout camping but we are on the cutting edge of providing innovative programming for girls.”

Metro Ministries of Corpus Christi

Reverend Edward Seeger
Executive Director
Metro Ministries of Corpus Christi
P. O. Box 4899
Corpus Christi, TX 78469

Annual budget:	$ 1,000,000
Annual evaluation budget:	$ 0
Staff:	30 fte
Stakeholder participation level:	Low

Description of Organization

Corpus Christi Metro Ministries is a nonprofit organization that provides a variety of services to homeless people including food, transitional shelter, clothing, referrals, employment opportunities and training. Formally organized in 1982, Metro Ministries began as a coordinated response by local clergy and religious lay leaders to the problem of homelessness. Currently, there are 100 supporting congregations and several denominational bodies that form the core of Metro Ministries.

Programs Evaluated

Rainbow House provides two-week emergency shelter and a longer-term transitional program for homeless women, and their children under twelve. The program provides counseling, job search assistance, access to day care, and other forms of assistance. Transitional housing is available for up to 24 months, depending upon a woman's needs and progress.

Rustic House is a transitional shelter for homeless mentally or physically disabled men and men over 60 years old. Men may stay indefinitely and are provided professional mental health services.

Purpose and Focus of the Evaluation

Metro Ministries' evaluation was designed to measure the outcomes of women in Rainbow House and men in Rustic House. The primary reason cited by the agency for conducting an outcome-focused evaluation was pressure from funders, in particular The United Way. As Executive Director Reverend Edward Seeger explained, "About nine years ago we thought it was sufficient to just be good people doing good things….The community was not holding us to any standards – it had no expectations for any outputs or outcomes" In the mid-1990s, the United Way of the Coastal Bend convened the executives of affiliate agencies and informed them that "when the local United Way allocations panel came to visit, they wanted to hear not simply how good we are and how many people in the community were applauding us, but what bonafide difference we could point to in individual's lives as the result of the activities we undertook."

Evaluation Procedures

To measure Metro Ministries' success rate in helping residents improve their conditions, the executive director designed an evaluation in consultation with staff. Data collection consisted of staff members documenting each residents' homeless condition at intake, documenting his/her outcome at dismissal, and determining whether the outcome was successful or not. The criteria for success, which assumed no time limit, was whether a resident departed to live independently or interdependently. The staff completed the assessments as a regular part of their case management. The data was then compiled on a quarterly basis and aggregated annually.

The report did not contain any narrative. Instead, it included determinations of success/not success for each individual resident as well as the overall percentage of success.

Board members and residents were not involved in the evaluation design, implementation, data analysis, or report preparation.

Uses of Evaluation Results and Processes

Metro Ministries reported using the evaluation in a variety of ways, including:

- To improve the outcomes or impact of the program,
- To report to a funder,
- To respond to questions or criticisms about the program,
- To promote the program to potential participants,
- To get additional funding,
- To decide resource allocation within the organization, and
- To improve outreach and public relations.

According to Executive Director Reverend Edward Seeger, Metro Ministries used the evaluation findings as marketing, fundraising, and internal appraisal tools. Although "the primary consumer for the evaluation is the United Way, Metro Ministries has come to use the data with a variety of audiences" Reverend Seeger explained. In fact, the agency used data from the evaluation in grant proposals, at speaking presentations, in reporting to the board, and in acknowledgement letters to donors. Reverend Seeger noted, "Now we can talk to service clubs or religious congregations or private benefactors and use the same language. The organization that you stand behind as a friend, as a supporter, as a donor, is in fact working and your dollars are buying community change. Financial support does get translated into differences in people's lives and therefore arguably into differences in the life of this community."

Reverend Seeger noted that the evaluation has also been used in an unintended way. It has generated a process that "helps staff 'notice' when someone is ready to go" by having them consider whether the resident's case would be considered a success were they to leave at that point.

In addition, Metro Ministries has used its evaluation to increase the community's understanding of "who [we] are and what [we]'re getting done" since the evaluation enables staff to articulate how residents have changed since receiving Metro Ministries' services.

Ad Hoc Group Against Crime

Vivian Jones
Vice President
Ad Hoc Group Against Crime
3330 Troost
Kansas City, MO 64109

Annual budget**:**	$ 688,172
Annual evaluation budget:	$ 0
Staff:	22 fte
Stakeholder participation level:	Low

Description of Organization

The Ad Hoc Group Against Crime's mission is "To combat and prevent crime, violence, and drug abuse by building individual and community self-esteem, pride and positive values through education, spirituality, morality, prevention and intervention, with the focus on youth and young adults in metropolitan Kansas City." To this end, Ad Hoc provides a variety of programs and services including Project Inspire (mentoring programs and services), Substance Abuse-HIV/AIDS Outreach Program, Runaway Youth Prevention/Intervention Project, Youth and Gang Services, Secret Information and Victim Assistance Program, Tutoring Program, Saving Youth N' Crisis (a support group), Grief Support Program (for members of homicide victims' families), Anger Control and Conflict Resolution Programs, and a Summer Youth Program.

Program Evaluated

Ad Hoc's Summer '94 Mentor Program, "Adolescents of Color at Risk…Mentored by Those Who Have Been There" served 100 youth between the ages of 12 and 18. Most youth were referred due to prior problems with violence and were selected from among youth participants from other Ad Hoc programs. Additional referrals were received from parents, police, and the Family Court. Mentors, all except one of who were ex-offenders, recovering drug addicts/alcoholics, or victims and/or perpetrators of violent crimes, were nominated from among 40-50 individuals participating in other Ad Hoc programs. Each mentor was assigned several mentees (from 6 to 14) and received a small stipend.

Purpose and Focus of the Evaluation

The evaluation was designed to measure the Mentor Program's outcomes, to assess implementation, and to measure participant satisfaction. In addition to studying the program's impact on participants, Ad Hoc undertook the evaluation to "leverage additional funding."

Prior to the evaluation, Ad Hoc had conducted some assessments of the program. Specifically, several families were contacted to discuss changes they saw in their children and mentors were interviewed regarding the activities they conducted with their mentees.

Evaluation Procedures

Ad Hoc approached the Ewing Marion Kauffman Foundation, a funder of other Ad Hoc programs, and asked the Foundation to evaluate its Mentor program. The Foundation agreed, and at no cost to Ad Hoc, led the evaluation effort building on data collection activities previously conducted by program staff.

The evaluation effort addressed numerous questions, such as:

Mentees

- What self-reported problems do mentees face?
- How much do mentees think the program helped them in a variety of areas?
- Why do mentees enjoy having a mentor?
- How does the number of non-status offenses change for mentees over the course of the program?
- What types of problems do parents identify their child having before and after the program?

Mentors

- What areas were mentors able to work on with their mentees?
- To what extent did mentors think the program helped them to work with their mentee in a variety of areas?

- What do mentors think of the training they received? What aspects were most and least helpful?

Mentors and Mentees

- What activities were most valuable to the mentees and/or to the mentors?
- What did mentors and mentees discuss?
- What recommendations for program improvement do mentees and/or mentors have?

Prior to the program, Ad Hoc staff designed and administered a pre-program questionnaire of mentees. The staff also designed and administered pre- and post-program questionnaire of parents.

Foundation staff designed and conducted interviews with mentees and mentors in addition to collecting data from the Juvenile Justice Department on the number of non-status offenses attributed to the mentees.

Board members, program participants, and funders did not participate in the evaluation effort.

The report included an executive summary, background on the program, a description of evaluation methods, a presentation of findings by data collection method, a conclusion, and recommendations. It also contained statistics (percentages of responses given by mentees, mentors, and/or parents), bar charts (of the percentages), and illustrative quotes from respondents.

Uses of Evaluation Results and Processes

Ad Hoc reported using its evaluation in a variety of ways, including:

- To improve the outcomes or impact of the program,
- To report to a funder,
- To respond to questions or criticisms about the program,
- To promote the program to potential participants,
- To get additional funding,
- To plan future programs,
- To improve outreach and public relations,
- To design an ongoing monitoring/evaluation process, and
- To gain support for evaluation among staff or the governing board.

According to Ms. Jones, the evaluation was used to assess how the program was working. In addition to measuring outcomes, the evaluation was used to understand program activities such as how often mentors and mentees were meeting, what they were discussing, and what activities the mentees enjoyed.

The evaluation provided the staff with ideas about what kinds of interventions are useful in working with high-risk youth. Based on evaluation findings, other programs were modified. In addition, following the evaluation Ad Hoc reassessed how they evaluated some programs and increased their focus on outcomes in some of their program areas.

Ms. Jones believed the evaluation helped Ad Hoc to leverage additional funding for the program. In fact, a slightly modified form of the program was funded for three years following the evaluation.

References

Coghlan, A.T. (1998). *Empowerment-oriented Evaluation: Incorporating participatory evaluation methods to empower Ugandan communities to prevent HIV/AIDS*. Ph.D. Dissertation, Cornell University.

Council on Foundations (1993). *Evaluation for Foundations: Concepts, cases, guidelines and resources.* San Francisco: Jossey-Bass.

Cousins, J.B., & Earl, L.M. (1992). The Case for Participatory Evaluation. *Educational Evaluation and Policy Analysis*, *14,* 397-418.

Cousins, J.B., & Earl, L.M. (1995). The Case for Participatory Evaluation: Theory, research and practice. In J.B. Cousins & L.M. Earl (Eds.), *Participatory Evaluation in Education: Studies in Evaluation Use and Organizational Learning.* (pp. 3-18). London: Falmer Press.

Fetterman, D.M., Kaftarian, S.J. & Wandersman A. (Eds.), (1996). *Empowerment Evaluation: Knowledge and tools for self-assessment and accountability*. (pp. 379-384). Thousand Oaks, CA: Sage Publications.

Gold, N. (1983). Stakeholders and Program Evaluation: Characterizations and reflections. *New Directions for Program Evaluation*, *17,* 63-72.

Greene, J.C. (1987). Stakeholder Participation in Evaluation Design: Is it worth the effort? *Evaluation and Program Planning*, *10*: 379-394.

Greene, J.C. (1988). Stakeholder Participation and Utilization in Program Evaluation. *Evaluation Review*, *12*(2): 91-116.

Greene, J.C. (1990). The Politics and Science of Program Evaluation. *Human Ecology Forum*, *18,* 25-27.

Greene, J.C., & McClintock, C. (1991). The Evolution of Evaluation Methodology. *Theory into Practice*, *XXX,* 13-21.

Guba, E.G. (1990). The Alternative Paradigm Dialog. In E.G. Guba (Ed.), *The Paradigm Dialog.* (pp. 17-27). Newbury Park, CA: Sage Publications.

Guba, E.G., & Lincoln, Y.S. (1989). *Fourth Generation Evaluation.* Newbury Park, CA: Sage Publications.

Gray, S.T. (ed.) (1993). *A Vision of Evaluation: A report of learnings from Independent Sector's work on evaluation*. Washington, D.C.: Independent Sector.

Hatry, H., van Houten, T., Plantz, M.C., and Greenway, M.T. (1996). *Measuring Program Outcomes: A practical approach.* Alexandria, VA: United Way of America.

House, E.R. (1980). *Evaluating with Validity*. Beverly Hills, CA: Sage Publications.

Kaminsky, A.L. (1993). *Participatory Evaluation in Hierarchies: Practical dilemmas of implementation (A case study).* Ph.D. Dissertation, Cornell University, Ithaca, NY.

Kroll, J. (n.d.). Evaluation Practices Among Grantmakers. Unpublished paper. Washington, D.C., Council on Foundations.

MacDonald, J.B. (1974). An Evaluation of an Evaluation. *Urban Review, 7,* 3-14.

McNelis, R. H. and Bickel, W.E. (1996). Building Formal Knowledge Bases: Understanding evaluation use in the foundation community. *Evaluation Practice, 17*(1): 19-41.

McNelis, R. H., Bickel, W.E. and Connors-Gilmore, M. (1996). A National Survey of Small and Mid-size Foundations: The role of evaluation. Unpublished paper. Pittsburgh: University of Pittsburgh.

National Center for Nonprofit Boards and Independent Sector (1997). *What You Should Know About Nonprofits.* Washington, D.C.: NCNB.

Nonprofit Management Association and Support Centers of America (1998). Evaluation in Management Support Organizations: Baseline study of current p ractices, February, 1998. Unpublished paper.

Parlett, M., & Hamilton, D. (1976). Evaluation as Illumination: A new approach to the study of innovatory programs. In G.V. Glass (Ed.), *Evaluation Studies Review Annual* (Vol. 1). Beverly Hills, CA: Sage Publications.

Patton, M.Q. (1997). *Utilization-focused Evaluation, 2nd Edition.* Newbury Park, CA: Sage Publications.

Plantz, M.C., Greenway, M.T. and Hendricks, M. (1997). Outcome Measurement: Showing Results in the Nonprofit Sector. In K.E Newcomer (Ed.), *Using Performance Measurement to Improve Public and Nonprofit Programs*. San Francisco, CA: Jossey-Bass Publishers.

Rebien, C.C. (1996). Participatory Evaluation of Development Assistance. *Evaluation, 2,* 151-171.

Renz, L., Mandler, C. and Treiber, R. (1998). Foundation Giving. USA: The Foundation Center, 92-93.

Rippey, R.M. (Ed.). (1973). *Studies in Transactional Evaluation*. Berkeley, CA: McCutchan.

Sanders, J.R. (1995). Book Review: Evaluation for Foundations: Concepts, cases, and resources, by the Council on Foundations. *Evaluation Practice 16*(3): 283-285.

Shadish, W.R., Cook, T.D., & Leviton, L.C. (1991). *Foundations of Program Evaluation: Theories of practice.* Newbury Park: Sage Publications.

Shulha, L.M. and Cousins, J.B. (1997). Evaluation Use: Theory, research, and practice since 1986. *Evaluation Practice, 18*(3): 195-208.

Smith, N.L. (1994). Evaluation Models and Approaches. In T. Husen & T.N. Postlethwaite (Eds.), *The International Encyclopedia of Education.* (pp. 2101-2109). Oxford: Pergamon Press.

Stake, R.E. (1983). Program Evaluation, Particularly Responsive Evaluation. In G.F. Madaus, M.S. Scriven, & D.L. Stufflebeam (Eds.), *Evaluation Models: Viewpoints on educational and human services evaluation.* (pp. 287-310). Boston: Kluwer-Nijhoff Publishing.

Sumariwalla, R.D. and Taylor, M.E. (1991). The Application of Program Evaluation in the Management of the Non-profit Sector: An exploratory study. A paper prepared for the 1991 Spring Research Forum "Leadership and Management."

Weiss, C.H. (1982). Measuring the Use of Evaluation. *Evaluation Studies Review Annual*, *7*, 129-145.

Weiss, C.H. (1983). The Stakeholder Approach to Evaluation: Origins and promise. *New Directions for Program Evaluation*, *17,* 3-14.

Weiss, C.H. (1997). Have We Learned Anything New about the Use of Evaluation? Plenary speech at the 1997 Annual Meeting of the American Evaluation Association, San Diego, CA.

Whitmore, E. (1988*). Participatory Evaluation Approaches: Side effects and empowerment.* Ph.D. Dissertation, Cornell University, Ithaca, NY.

Worthen, B.R., Sanders, J.R., & Fitzpatrick, J.L. (1997). *Program Evaluation: Alternative Approaches and Practical Guidelines.* New York: Longman.

than goals; and legitimacy of the interests and information needs of varied evaluation audiences (ibid.).

Table 1. Major Approaches to Program Evaluation

Philosophical Framework	Key Values Promoted	Key Audiences	Preferred Methods	Typical Eva
Postpositivism	Systems theory/efficiency, accountability, theoretical causal knowledge	High-level policy- and decision-makers	Quantitative -- experiments and quasi-experiments, systems analysis, causal modeling, cost-benefit analysis	Are desired o and attributed Is the progran alternative?
Pragmatism	Management/effectiveness, practicality, quality control, utility	Mid-level program managers, administrators, other decision-makers	Eclectic, mixed -- structured and unstructured surveys, questionnaires, interviews, observations	Which parts o work well an improvement How effective with respect t organization's beneficiaries
Interpretivism	Pluralism/understanding, diversity, solidarity	Program directors, staff and beneficiaries	Qualitative -- case studies, interviews, observations, document review	How do vario experience th
Critical, normative science	Emancipation/empowerment, social change	Program beneficiaries, their communities, other ✌powerless✍ groups	Participatory -- stakeholder participation in varied structured and unstructured, quantitative and qualitative designs and methods; historical analysis, social criticism	In what ways goals, or activ program mair resource ineq society?

Adapted by Greene from Greene (1990).