Evaluation for Program Improvement: Sharing a Model Used for Youth Services

Trish Oberweis and Matthew Petrocelli

Program evaluation has become part and parcel to how city departments do business. There is great potential for evaluators to be innovative in designing the research, and attentive to the needs of city decision-makers. One innovative model is represented by the IDEAL[1] Project in the American Northwest, modeled after the ProDES Project in the East. The project design builds in a means to use outcomes information to improve program performance, creating a learning environment in which those providing services for youth could be able to achieve ever-greater results with their clients. The evaluation system is understood as a learning tool to increase program effectiveness in two ways: by relying on data to assess strengths and weaknesses and by collaboration between program planners and the evaluation staff to use the data to enhance performance. The paper concludes with some comments about the distance between design and accomplishments.

Introduction

Agencies providing social services for youth are increasingly being challenged to be more accountable than ever before (Hodges and Hernandez, 1999, Cherin & Meezan, 1998). Even before cities faced the fiscal cutbacks outlined since 9/11, Joseph Wholey wrote, "Our context today is one of increasing demands for effective public services, limited resources for government agencies and programs, declining public confidence in and support for government, and increasing demands for accountability by those providing public services" (1996:145). Such an increase in demands for accountability is now paired with a decrease in available funds.

In the face of these two simultaneous challenges, social service providers must establish the nature of their own impact in order to continue to be competitive for funds. Performance and outcome measurement are two vehicles that afford the opportunity to both respond to calls for accountability and simultaneously create information for program development, refinement and/or expansion. According to Hodges and Hernandez, many children's mental health service systems "are implementing systems of outcome accountability with the express purpose of using outcome data to improve service planning and delivery as well as cost efficiency" (1999:184). This close attention to performance and outcome measurement is intended to have the effect of improved abilities of programs to accomplish their goals.

Such systems may indeed be able to use data to help programs learn how to achieve their objectives increasingly well. Reporting of data-driven outcomes may also enhance the credibility of the program and, especially important during times of cutback, performance or outcomes monitoring systems may help "communicate the value of their programs to policymakers or to the public" (Wholey, 1991:55). Creating a data system that enables accountability has advantages for human service systems. By creating a learning environment and building a dataset to guide that learning, human service departments, including justice systems, have the ability to design a powerful planning tool. Although the model described here

---

[1] An acronym for the project's pseudonym.

serves a juvenile justice system, the concept transfers to any system using a range of programs to serve a particular population.

A Model for Evaluation and Development

In 1999, the Juvenile Probation Department in a large Northwestern city set out to build a tool that would feed a culture of learning. An information system was planned to assist with an array of decisions related to referrals and placements of youths in programs, from planning the range of services, to selecting specific agencies to provide those services, to actual referrals of young people. Planners envisioned an information system that would enable program evaluation as a primary function, among the other data services, but wanted to ensure that evaluation information would be put to use to improve the effectiveness of the programs the Department funded.

To this end, high-ranking probation staff identified a research model whose output appealed to their collective sense of the types of questions they wanted to answer. This model then became the starting point for the development of a local project, later named Program IDEAL (an acronym for Information for Development, Evaluation and Learning). The design was to be implemented in phases, beginning with Department contracts with community-based organizations that provide services to delinquent and predelinquent youths living at home. For youths placed on probation and returned to the home of a relative or other caregiver, the Department holds contracts with about 30 community-based programs. Each of these programs is contracted to serve a specified number of clients annually. These clients, with the help of the program staff, became the first group of participants in the Program IDEAL Project. The project then expanded to include county placement programs, and was intended to expand again until every youth receiving any service funded through the Department will participate in the outcomes monitoring system[2].

Among the goals of key probation staff for the information system was a desire to learn in more rich detail about the achievements of each program with its clientele. Recidivism was but one of many dimensions to be monitored over time. Other areas of interest were intended to paint a more robust picture of treatment effects than merely recidivism rates could. These other dimensions represent intermediate client changes that are related to program interventions, such as improvement in school attitude or improved family relationships. These are measured quantitatively, relying for the most part on pre-validated psychosocial scales.

Such monitoring of intermediate (in-program) and ultimate (recidivism) outcomes was expected to enable a continuous flow of a wide range of information, which should reveal patterns in the various programs' strengths and weaknesses. The nature of the design fosters data-driven program refinements. It enables program planners to receive quantitative feedback on any programming changes, comparing aggregate data on client outcomes before the programming change to data collected after such a change. Although the Juvenile Probation Department at that time was already engaged in an evaluation of many of the programs for in-home supervision, some stakeholders wanted something more complicated than a snapshot evaluation.

---

[2] The author left the project in 2002 and shortly thereafter, California's budget problems escalated drastically in the fallout of both the Enron scandal and the simultaneous security crisis. The Program IDEAL was absorbed into the Probation Department, but was not expanded as the original plan had aspired.

The goal of the Program IDEAL project was to create a tool for refining programs, improving placement decisions, and for learning the complexities of what works, with whom, under what circumstances. Department representatives wanted a tool for *developing* programs and services for youth as well as for *evaluating* them. They wanted a means to discover the variety of factors that contributes to overall success (meaning a lack of recidivism). But system planners also desired a way to translate that information into real decision-making about which youth should go to which program and receive which services. Planners wanted a system capable of testing program theory (Blalock, 1999). The model for this kind of information development and processing existed in the work undertaken as the ProDES Project.

Building on an Innovative Model

Two information systems, which formed the starting point for the Program IDEAL Project, had already been constructed under the directorship of Drs. Phil Harris and Peter Jones of Temple University. The original information system, ProDES (Program Development and Evaluation System) was developed mostly (but not exclusively) for residential treatment programs serving delinquent youth.   A second information system, PrOMIS (Program Outcomes Monitoring Information System) was developed for research with delinquency prevention programs. These systems have now been in operation for nearly 15 years, and have become trusted sources of information for various units within its own juvenile justice practitioner community.   The project was a finalist in 1999 for Harvard University's Kennedy School Innovations in American Government Award.

Design for Data Collection and Use

The Program IDEAL system was planned to begin operation with local day treatment (non-residential) programs and, over time, to expand to include all residential facilities as well as various entities within Probation Services to which local youths were sent. The research design consists of as many as eight data collection points for any given youth and relies on up to six sources of information about each client (See Chart 1). Background and situational information is gathered from the cumulative manual files kept by probation officers. Juvenile and adult arrest records and/or court data are extracted both prior to (juvenile records only) and several months after program interventions are administered. Data are collected using a pre- and post-test design at intake to and discharge from the program. (Roberts-Gray, Steinfeld & Bailey, 1999.) The pre- and post-tests include both a staff assessment of the youth and his/her self-report data as captured by a number of psychometric scales. The four categories of measurement dimensions are self esteem, school bonding, family bonding and an assortment of personality attributes. (See Chart 2 for a complete list of these scales.) Finally, months after program participation has ended, the former client and his/her caregiver are interviewed separately by telephone. The measurement dimensions and how they were determined are discussed in a later section. Involving a variety of system users in the development of the system and its design was critical to the success of the project.

Chart 1: Data Collection Points in the Program IDEAL Project

Data Collection Form	Point in Time	Data Source
Intake Record Data	As close to formal adjudication as possible	Probation files and the existing offense-based data
Intake Staff Assessment	Within 3 weeks of program entry	Program staff person/case manager
Intake Self Report	Within 3 weeks of program entry	Youth-client
Discharge Staff Assessment	As close as possible to the end of service delivery	Same program staff person as above, often a case manager
Discharge Self Report	End of service delivery	Youth-client
Follow up Interview 1	6 months after services end	Youth's parent or caregiver
Follow up Interview 2	6 months after services end	Youth/former client
Recidivism Record Check	9 months after services end	City-maintained databases and the statewide warrant/arrest-tracking database.

Although the Program IDEAL system is able to provide system information and some information intended to facilitate case management, this discussion focuses exclusively on the system's propensity to evaluate and facilitate program refinement. The purpose of this article is to describe the potential of this system as a public administration tool to assist with development and refinement of human service programs and to discuss the specific features of the systems' development that enable its potential power as a learning tool.

Much of the potential of Program IDEAL stems from a basic feature of its design: from inception it was organized around the framework of utilization. Evaluation studies in the past have suffered greatly from under-utilization, meaning that even after data were available for use in decision –making, still decisions were made without them.   Studies consistently point to a problem with evaluation information going unused in actual decisions and policy-making (House, 1972; Wholey et al., 1970; Cherin & Meezan, 1998; Patton, 1997:7-10).

Two important facets of the structure of the project were defined through this focus on utilization. First, the data flow was to be continuous and the data collection was to be permanent. Second, the contents of the database were to be determined according to the perspectives of all system users rather than by justice literature or by relying solely on the perspective of the funding agency—in this case the Department of Juvenile Probation.

A Continuous Flow of Information

Although the Program IDEAL system would be able to perform the function of evaluation, by no means is the system a simple tool for process or impact evaluation. Instead, the system is intended to provide a *continuous flow of information* to those being evaluated and the system as a whole to enable (as nearly as possible) *real-time examinations of outcomes.* Program IDEAL is an outcomes monitoring

system. This is to say that IDEAL can provide information that is timely and that allows for assessment of program refinements as often as desired, and certainly within the time frame of an annual contract.

The design of the Program IDEAL system guards against two common problems with typical evaluation. The first of these common problems relates to *when* the research information becomes available during the process of completing an evaluation. At the risk of over-simplifying, let's consider that a typical impact evaluation consists of three parts, which can reasonably be expected to take a total of one year:

1) identification or construction of a data set,
2) the analysis of those data, and
3) the reporting of evaluation findings.

In many instances, the final report is delivered too late to be of use in program planning for the next budget cycle and too late to assist funding agencies with the decision of whether and how to fund the program (Henry & Rog, 1998). From a program planning perspective, or from the perspective of those responsible for making annual decisions about a contract, having one-shot impact or process evaluation reports too often lends little or no guidance for future planning . There many valid reasons why this is so, but the result remains frustrating. After a good amount of effort and expense, the program may receive evaluation information that might not apply to the program as it stands at that time and the information may not (or may no longer be) particularly useful in planning, adjusting, refining or expanding the program. Thus, even rigorously devised evaluation information can miss the mark of usefulness, due to the timing of results combined with the ever-evolving nature of programming.

The Program IDEAL system guards against problems such as this one by providing a *continuous flow* of information. This means that once Program IDEAL achieves a level of data that will support a program's evaluation, new evaluation outputs will be available as often as desired.

The usefulness of a continuous flow of information becomes particularly apparent in the context of program development or refinement. Let's suppose that a program participating in the Program IDEAL project wants to add a component to its program. Within a relatively short period of time (say one or two durations of the average program length), Program IDEAL would be able to demonstrate the effect of that new component on clients' in-program change. Armed with a continuous flow of information, the program is able to learn from Program IDEAL reports, adjust its course of action in terms of service delivery, and then, shortly thereafter, obtain an analysis of the impact on in-program changes, as well as ultimate outcomes, such as recidivism.

For example, a program may adjust its staffing patterns and capabilities, such as adding non-English speakers or drug counselors or it may change program components and services, such as introducing a tutoring component. By comparing the data for in-program changes among clients *before* the change was made to the data for clients *after* the program was changed, coordinators and planners can quantitatively see the impact of programmatic changes. In this way, programs participating in Program IDEAL have the advantage of being able to learn about how changes in service delivery relate to the program's ability to meet its goals. For the more sophisticated program, there is even an opportunity to test hypotheses about what kinds of services or staff members are the most effective with (particular kinds of) clients.

Similarly, programs may use data to understand systematically how the program differs in practice from how it works theoretically. For example, using Program IDEAL outputs, a program may learn about subgroup variations in its outputs.   Harris and Jones argue that, "By ignoring difference among treatment clients, important patterns of outcomes are missed, leaving policymakers with outcome results that are often too superficial for useful prediction" (1999:405).

A simple example of the benefit of examining subgroup variations might be to compare the group of clients with the best outcomes against the intended target population.   The characteristics of those two groups--those with the best outcomes and those intended to benefit from services--should theoretically be the same. If not, adjustments can then be made either to the defining of the target population, or to the services provided.   Perhaps a program might find that although its target population is youths aged 13-19, it has good outcomes with 17-19 year olds and poor outcomes with 13-15 year olds. Information could be used either to adjust the target population to focus on the older group, or to develop new programming to improve the outcomes with the younger group.

By enhancing the level of harmony between characteristics of clients served and services provided that are known to be effective with those client characteristics, programs and clients are likely to enjoy improved outcomes. Certainly, Program IDEAL analyses could then be used to learn whether and how changes in target population or services actually impact the achievement of program objectives. Additional adjustments to the program's target population or array of services can then be made according what the data show. The continuous flow of information facilitates this culture of learning, in which programs look at outcomes data to identify patterns of clients, and program strengths and weaknesses. Programs can make changes based on data and can follow up on the changes by revisiting outcome data to examine the impact of those changes.

The primary advantage of having a continuous flow of information, then, is that *evaluation outputs arrive at the program in time to adjust the course of action to improve program effectiveness* usually within a single budget year. However, there is more to be considered in making the move from timely information to actual program improvement.

The notion of "improvement," depends on having a well-defined set of objectives. How would one notice improvement without knowing what the goals were? This brings us to the second issue of the typical evaluation against which Program IDEAL's design guards.

Goals and evaluation: Making meaningful information for all system users

In addition to the problem of receiving evaluation reports after the opportunity to adjust a program's course has passed, there is a second risk involved with many typical evaluations. Utilization of evaluation data is of critical importance in the planning stages of an information system (Wholey & Newcomer 1997; Patton, 1997; Poulin, Harris, & Jones, 2000). There are several barriers to utilization, each of which must be addressed, but this discussion is oriented toward a single utilization concern: selecting evaluation criteria that yield meaningful outcome information to system users.

Evaluation research is often oriented toward measurement of goal achievement, in one way or another. *It is of fundamental importance to be cognizant of all objectives and who set them.* Outcome criteria should not be derived only from the literature, or only from the perspective of the funding agency.

The evaluation should be linked to the goals and values of the agency being evaluated as well (Papineau & Kiely, 1992). Wholey argues, "[program] managers can and should play central roles in evaluation of the programs for which they are responsible" (1991:56), including participation in the selection of criteria for measurement.

Without involvement from the programs' staff members, the data are far less likely to be meaningful to the program decision-makers who might be in a position to make more informed decisions, if appropriate data were available. Insightfully, Poulin, Harris, and Jones argue, "Without measuring the stated goals of the organizations within the system, we risk two things: (a) not learning key pieces of information about the desired outcomes of organizations and (b) producing an evaluation that will be ignored by the organizations we seek to aid in development" (2000:517).

If the information is not meaningful, it will go unused, despite the costs in terms of both money and other resources drained in the process. Inviting members from the organizations being evaluated as well as other system users to participate in the planning process not only sets the stage for creating a learning environment, but also ensures that the data will be meaningful to those system users and that the evaluation output is useful and credible to them. Henry and Rog attest, "Real involvement in the design and ongoing decision-making of stakeholders and potential users of the evaluation has been shown to be related to the perception that the information is useful" (1998:100).

Clearly, discussion about the set of criteria on which to base the evaluation is necessary and should include representation from all intended users, if possible. For example, with the Program IDEAL project, the current set of stakeholders includes the Juvenile Probation Department (in particular, roughly 8 key staff members have been actively involved) and nearly 30 community-based programs that are being evaluated and that have the potential benefit of being able to learn from the system.

It would be perhaps the simplest design for the research team to rely on justice literature to determine what programs *ought* to be doing with and for young people. It would also be expected that the funding agency, in this case the Juvenile Probation Department, might set some objectives for program achievement. There is nothing inherently problematic with either of these approaches and, indeed, both were employed at various stages in the development of this system.

However, one critical perspective is missing. In order for IDEAL evaluation outputs to be meaningful to the executive staff at community-based organizations, the outcome dimensions that are measured must be closely tied to the programs' self-selected goals. Because this tool is being made available to executive staff, they must be able to have output information on the dimensions that have meaning for them in terms of their own program's self-improvement. Adding this perspective at the planning phase elevates an outcome-based evaluation system into a potential tool to refine and develop, to expand and nurture the very programs themselves.

Meaningful evaluation output is of critical importance, a point that seemingly cannot be stressed enough. By developing an evaluation project, identifying or creating a dataset and analyzing it, resources must be used, both in terms of funding and in labor power. These resources will be depleted in vain if the evaluation output is not useful for program planning. To this end, a great proportion of Program IDEAL's own developmental resources were spent on a process to determine what data the system would collect, seeking to maximize useful output yet minimize the data collection burden placed on programs. One

process for determining how to evaluate a program and what data to collect that relies on program planners' ideas about meaningful information is called "evaluability assessment."

Evaluability Assessment

Evaluability assessment is a way to involve program planners in the development of an evaluation model that yields meaningful information. Wholey and Newcomer write, "Securing agreement among the key stakeholders regarding agency mission and strategic objectives is the first step" (1997:97). Program goals are specifically identified during the evaluability assessment process, and the contents of the program are spelled out clearly. Sometimes, a formal logic model may be created to demonstrate how the program's content will produce the social effect it seeks (Julian, 1997). The process provides the research team with the program's perspective on what it seeks to accomplish and how it endeavors to do so.

Program IDEAL's Developmental Methodology

Evaluability assessments were undertaken with each of 35 programs at community-based organizations (CBOs), all of which had contracts with the Juvenile Probation Department in 1999-2000. The evaluability assessment meetings included staff from multiple levels within the program or agency, and were the research team's basic introduction to each program. Prior to these meetings, the researchers had little or no conception of each program's content or objectives.

In the meetings, each program staff group was asked to enumerate the program content and goals, and discussion remained open until consensus was reached. Generally, there was less debate about objectives than was anticipated, but this may have related to a decision to allow an expansive list of objectives for each program, rather than forcing agreement on just the most important few. Rather than having the group select the most important two or three goals, the lists tended to be long and included a wide range of program goals. The opportunity to shorten the list would come later.

One main function of the evaluability assessment interviews was to enable the research team to understand which facets of the clients the program sought to change so that those dimensions would be measured would frame the feedback based on Program IDEAL data. Because all the programs participating had received contracts through the Department of Juvenile Probation, the research team was aware that all programs—and the Juvenile Probation Department-- share the common overarching goal of preventing any future delinquency or recidivism. However, as more intermediary goals, programs sought to change attitudes, perceptions, and behaviors of their clients in specific ways that were assumed to have an impact on future delinquency.

For example, most programs commonly assumed that improvement in school attendance and achievement would have the effect of preventing future law violations and so focused on changing client attitudes toward and behaviors in school. These intermediate goals, as specified by the actual programs involved in this evaluation, were intended to form the basis of the evaluation system.

By gaining knowledge about each of the programs in depth, evaluators will also have the added benefit of being able to use the system to examine how changes in programming affect changes in outcomes. The qualitative work undertaken to achieve a thorough understanding of program goals was

important to the research team's ability to accurately report how well various programs were doing in achieving their own, self-selected goals.

Equally important, the process allowed the program personnel to take ownership of the evaluation process. Not only were their own unique goals to form the basis of the evaluation, but the research team has given ample opportunity to program staff members of all rank to join in the process. After the qualitative interviewing process was completed, program staff members were inclusively invited to read the descriptive reports that highlighted what the research team understood to be the content of those interviews. Any and all staff members were invited to discuss, edit, refine or otherwise alter the content of the descriptive reports.

The editing process represented a final opportunity to clarify program goals and other descriptive elements. All suggested revisions were made, and no staff member from any program objected to any content in any report. Instead, there were simple factual changes (such as a target population that listed one age group rather than the correct one), enhancements (particularly with regard to direct questions inserted into the draft text), and clarifications to the writing itself. Although no program changed factual information except to add clarification, program representatives seemed to be comfortable suggesting changes and refinements to the description of their programs. The fact that all program staff members were so heartily invited to participate in the planning of their evaluation may have enhanced their interest in the project as well as its credibility with the executive staff (Wholey & Newcomer, 1997; Poulin, Harris, & Jones, 2000).

After evaluability assessments and the follow on reports were completed with 33 programs in 31 agencies, the research team found a space to begin envisioning the evaluation project—one whose outputs would be closely related to the program's goals. The next step was to identify from the evaluability assessments those common elements of intermediary change that many programs sought to achieve with clients. Thus, if multiple programs identified better school attendance as a first step toward reducing future delinquency, then school attendance counted among the outcomes the system would potentially monitor.

Two kinds of responses about program's self-definition demanded consideration for measurement in the system. Sometimes, programs responses revealed already-identified short term and middle term objectives. For example, if program personnel were explicit about their assumptions that changing a given dimension like school attendance would decrease future delinquency, and the program spent time and energy trying to increase clients' school attendance, then those assumed correlates—in this instance, school attendance—demanded consideration as useful to measure. In short, where the assumptions were clear, those dimensions became intermediate outcomes and were likely candidates for our monitoring system.

Similarly, programs offered ideas not as objectives of the program, but as indicators that a client was doing well in the program and was expected to be successful after participation ended. Although not necessarily specific objectives of the program, these indicators served as feedback that the client was succeeding, learning the lessons that the program offered and was seen as more likely to avoid future delinquency. These success indicators also were placed on the list of item being considered for measurement.

What is important about this methodology is that local justice professionals—in programs and official positions—decided what was meaningful. Their understandings of client change and delinquency

reduction have driven the framework for this evaluation information system from the beginning. The assumptions and hypotheses of the Department and the community-based programs delivering services drive this system. As described earlier, this is critical to the overall usefulness of the information being created or processed through the Program IDEAL database.

Evaluability assessment served as a means to gain the participation and support from the program staff members themselves, and this support will prove critical to the lasting success of this project. Program personnel were actively participating in the design of an evaluation study that they recognized could impact their relationship with the funding agency, the Department of Juvenile Probation. Ambiguity and confusion about how the data would be used by the Probation Department and the research team contributed to programs participating hesitantly, and data flow was difficult to establish.

However, the fact of their cooperation in creating the list of dimensions of client change to be examined empirically sets the stage for their accepting as valid, useful, and relevant whatever the output may be, regardless of how it looks for any given local program. It was a means of agreeing to a particular set of informational items that made sense as criteria to use as intermediate outcomes for evaluation purposes. The list of potential items to monitor were deemed meaningful by participating programs specifically because they collectively created that list rather than having the parameters of the research be determined by probation staff members, research staff members, or other similar projects, such as ProDES.

The final step in creating the first draft of data collection instruments involved operationalizing the outcomes dimensions to be measured. That is, vehicles to measure those dimensions had to be either created or selected. To this end, a research subcommittee was formed and included program leaders and staff who then assisted in identifying suitable measures (with the guidance of the research staff) to assess each youth's progress toward a given objective. Again the intended effect of intensive participation by stakeholders was that any findings resulting from these measures would be accepted as solid indicators of change.

The fact that local professionals played such a central role in shaping the design of the system would hopefully have significant implications for how the information will be used in the future. Because the data were designed to be meaningful to service providers, the system should have added value and could even be a capacity-builder for the city, enabling it to demonstrate the value of its services more rigorously than other cities competing in for the same state and private funds. However, simply to have data is not to use it to make empirically driven programmatic decisions, refinements and expansions (Patton, 1997). With measurement underway, the focus must then turn to building a community of data consumers, facilitating the use of research data to make program decisions. At this juncture, serious and potential fatal problem developed, to which we will return in a few paragraphs.

Information Consumption: A future in bridging the gap between data and planning

Even years of excellent planning and careful data collection could potentially be rendered powerless if stakeholders ignore a final but critical element: users of information must be empowered to use data through training that is directed at the appropriate levels (Smith, 1998). Program administrators are not necessarily, or even usually, data experts and nor are they paid to be. Moreover, within the constraints of an eight hour day, data may seem less relevant to a program's success than a client in need of service or

less urgent than a proposal deadline. Other things may seem more important than reading research reports, and some of those who do take the time might not be trained to connect the findings to specific courses of action for the program.

Cherin and Meezan write, "Conventional wisdom in the field of evaluation utilization attributes most of the underutilization of evaluation data to, among other things, the lack of training in empirical methods at both the administrative and staff levels of the organization" (1988:3). The Program IDEAL research team was intended to work closely with program coordinators and directors around the evaluation findings. A member of the research team would deliver findings in person, holding meetings with key program personnel to go over quarterly or bi-annual reports about program outcomes. Reports may still go unread, but the findings they contain would nonetheless be shared with the program staff. The information was to be explained and clarified for program personnel, who would have the opportunity to raise questions and direct future research about their own program outcomes.

The goal of the Program IDEAL system was to learn about outcomes and to use data to improve outcomes. To this end, any number of outcomes reports can and will be generated to look for variations among sub-groups of youths, such as how the service needs of young teens compares to the service needs of late teens. This type of inquiry could be general (e.g. for the whole system) or specific (e.g. for the clients of a given program).

Strength and Development through Collaboration

Collaboration between the research team and the program development staff enables data driven program planning. It ensures that information generated through the research process is useful for assisting in program planning. But there is a larger collaboration involved in the Program IDEAL Project, one between the community programs and the Department of Juvenile Probation. The Department needs evaluation to ensure its own fiscal responsibility, but it does not need to provide such a powerful learning tool to programs fortunate enough to have contracts. However, by investing in a system like IDEAL, one that can provide information to help programs learn how to be more effective by empirically looking at the impact of programmatic shifts, or trends among sub-groups of clients, the department is investing in the future.

These evaluation dollars mark an investment in education; in learning in as much detail as possible what kinds of programs and services work with which kinds of youths. They open the possibility to achieve ever-greater levels of success with youths sooner than ever before. By collaborating with the programs, the top levels of the Probation Department attempted to create a win-win arrangement in which programs have the opportunity to become increasingly better at serving youths and in which the Department is able to contract with those increasingly more effective programs.

From Plan to Practice: Policy, Politics and Putting It Into Practice

The Program IDEAL System had some important characteristics: it was well planned, well designed, backed by the highest levels of authority within the Juvenile Probation Department, and reasonably well funded. But it lacked a critical component: not all of the middle and line level staff supported the data collection. Surveys were long and moderately intrusive. Although each question on the survey was there because the information it yielded was important to assessing the outcomes that programs themselves had identified, there was a predictable resistance to using the data collection forms.

When case working program staff voiced their concerns about data collection, Program IDEAL staff listened and responded with information about why each question mattered. Some of mid-level supporters within the Probation Department did the same, giving the project the essential backing it needed. Meetings were held to shorten and alter the data collection forms. But, when case working staff issued their concerns to line level and low level supervisory staff with the Probation Department, mutiny prevailed. Some of the Probation personnel supported the project emphatically, but this support was inconsistent across personnel and this proved to be a fatal problem.

Despite direct orders from the highest levels to try out data collection, some middle and low level staff within the Probation Department were not subtle in their rejection of the research project. They supported the resistance, encouraged acts of civil disobedience in the form of refusing to comply with data collection--even refusing to acknowledge that a new youth had begun treatment at a program--and data collection never became routine. The project was doomed to failure. All the collaborative meetings and the three years of developmental work was undone within a year when some Probation personnel undermined the authority of the Chief, aided and abetted program staff who were uncertain about collecting data on clients, and supported the outright refusal to provide data to the research team.

It was a tragic demise to a system intended to benefit programs, as well as delinquent youths and those at risk for delinquency, and the larger municipal structure designed in part to assist them. A short time after the author employed by the project left the Program IDEAL environment, the project was reduced, then subsumed into the Project Department itself and defunded. The largest lesson learned is that in that place where the rubber meets the road, the integrity of the ideal must be retained. Program IDEAL had some lofty ideals. Having seen the success of these same ideals in another research setting invigorated both the research staff and many of the Probation Department supporters of the project and gave hope that early resistance would be overcome. In this case, a lack of support from some among the middle and lower ranks within the Probation Office was enough to substantially undermine the project.

Author Information

Trish Oberweis, Ph.D.
Southern Illinois University Edwardsville
Department of Sociology
Box 1455
Edwardsville, IL 62026
toberwe@siue.edu

and

Matthew Petrocelli, Ph.D.
Southern Illinois University Edwardsville
Department of Sociology
Box 1455
Edwardsville, IL 62026
mpetroc@siue.edu

In memory of Lawrence Bennett

References

Blalock, A. B. (1999). Evaluation and the Performance Management Movement: From Estrangement to Useful Integration? *Evaluation, 5*(2), 117-149.

Cherin, D., & Meezan, W. (1998). Evaluation as a Means of Organizational Learning, *Administration in Social Work, 22*(2), 1-21.

Harris, P. W. & Jones, P.R. (1999). Differentiating Delinquent Youths for Program Planning and Evaluation. *Criminal Justice and Behavior, 26*(4), 403-434.

Henry, G. T. & Debra J. Rog . (1998). A Realist Theory and Analysis of Utilization. *New Directions for Evaluation, 78*, 89-102.

Hodges, S. P., & Hernandez, M. (1999). How Organizational Culture Influences Outcome Information Utilization. *Evaluation and Program Planning, 22*, 183-197.

House, E. R. (1972). The Conscience of Educational Evaluation. *Teachers College Record, 73*(3), 405-414.

Julian, D. A. (1997). The Utilization of the Logic Model as a System Level Planning and Evaluation Device. *Evaluation and Program Planning, 20*(3), 21-257.

Papineau, D. & Kiely, M.C. (1992). Participatory Evaluation in a Community Organization: Fostering Stakeholder Empowerment and Utilization. *Evalution and Program Planning, 19*(1), 79-93.

Patton, M. Q. (1997). *Utilization-Focused Evaluation: The New Century Text..* Thousand Oaks, CA: Sage Publications.

Poulin, M. E., Harris, P.W., & Jones, P.R. (2000). The Significance of Definitions of Success in Program Evaluation. *Evaluation Review, 24*(5), 516-536.

Roberts-Gray, C., Steinfeld, S., & Bailey, W. (1999). Goal Setting and Progress Evaluation in Youth Empowerment Programs. *Evaluation and Program Planning, 22*, 21-30.

Smith, M. K. (1998). Empowerment Evaluation: Theoretical and Methodological Considerations. *Evaluation and Program Planning, 21*, 255-261.

Wholey, J. S. (1996). Formative and Summative Evaluation: Related Issues in Performance Measurement. *Evaluation Practice, 17*(2), 145-149.

Wholey, J. S. (1991). Using Evaluation to Improve Program Performance. *The Bureaucrat: The Journal for Public Managers, Summer,* (1991): 55-59.

Wholey, J. S. & Newcomer, K. (1997). Clarifying Goals, Reporting Results. *New Directions for Evaluation, 5,* 91-98.

Wholey, J. S., & Scanlon, J.W., Duffy, H.G., Fukumotu, J.S. & Vogt, L.M. (1970). *Federal Evaluation Policy Analyzing the Effects of Public Programs.*