University of Massachusetts Amherst

From the Selected Works of Sharon F. Rallis

2004

What is Program Evaluation? Generating Knowledge for Improvement

Sharon F. Rallis, *University of Massachusetts Amherst*

Kathleen A Bolland, *University of Alabama - Tuscaloosa*

Available at: https://works.bepress.com/sharon_rallis/7/

What is Program Evaluation? Generating Knowledge for Improvement*

SHARON F. RALLIS^a and KATHLEEN A. BOLLAND^b

^aUniversity of Massachusetts-Amherst, Hills South, Amherst, MA 01003, USA (E-mail: sharonrallis@earthlink.net); ^bUniversity of Alabama, Box 870314, Tuscaloosa, AL, 35487-0314, USA

Abstract. In this introductory article, we discuss the nature of Program Evaluation, describing the concepts that underlie our formal and informal evaluative efforts. Program Evaluation, like any deliberate inquiry process, is about learning. The process explicates program purposes, activities, and outcomes and generates knowledge about their merit and worth. This knowledge can inform planning and lead to program improvement. We present and discuss various definitions of Program Evaluation, focussing on its purposes and uses. We also provide an overview of the inquiry process, grounding the search for merit and worth in the American Evaluation Association's Guiding Principles for Evaluators. Because program evaluations are typically conducted to inform decision makers, we discuss aspects of professional practise that contribute to the use of an evaluation.

Introduction

Program evaluation, like any deliberate inquiry process, is about learning. The process explicates program purposes, activities, and outcomes. Evaluation goes beyond describing; it generates knowledge about the program. Framing the questions, rigorous research design, and analysis

* This chapter draws heavily on previous work by Sharon Rallis and Gretchen Rossman; see the following references. Rallis, S.F. and Rossman, G.B., "Mixed Methods in Evaluation Contexts: A Pragmatic framework", in A. Tashakkori and C. Teddlie (eds.), *Handbook of Mixed Methods in the Social and Behavioral Sciences* (Thousand Oaks: Sage, 2002). Rallis, S.F. and Rossman, G.B., "Communicating Quality and Qualities: The Role of the Evaluator as Critical Friend", in A.P. Benson, D.M. Hinn and C. Lloyd (eds.), *Visions of Quality: How Evaluators Define, Understand, and Represent Program Quality* (Oxford: JAI Press, 2001), pp. 107–120. Rallis, S.F. and Rossman, G.B., "Dialogue for Learning: Evaluator as Critical Friend", in R. Hopson and M.Q. Patton (eds.), *How and Why Language Matters in Evaluation, New Directions for Evaluation*, 86 (San Francisco: Jossey-Bass, 2000). Rossman, G.B. and Rallis, S.F., *Learning in the Field: An Introduction to Qualitative Research*, 2nd edition (Thousand Oaks: Sage, 2003). Rossman, G.B. and Rallis, S.F., "Evaluation as Learning: Critical Inquiry and Use as Action", in V.J. Caracelli and H. Preskill (eds.), *The Expanding Scope of Evaluation Use, New Directions in Evaluation* 88 (San Francisco: Jossey Bass, 2000).

and interpretation of data are at the heart of good program evaluation practice. Data become information that can be used for program improvement and ultimately, according to many evaluators, social betterment. What makes the inquiry evaluative is that the process focuses on the merit and worth of these purposes, activities, and outcomes.

As president and treasurer, respectively, of the American Evaluation Association in 2004, we informally evaluate our organization. Through review of bylaws, minutes, and budgets, by talking with a wide range of members, and in observations of meetings, we collect data about membership, governance, policies and procedures, programs, projects, and relationships. We question the relationships between our mission and purposes and our activities and accomplishments. Our analyses and interpretations of all these data reveal areas of strength and of need. This ongoing evaluation helps us to identify issues and activities to propose for board consideration.

We collect, analyze, and interpret data as a matter of course because we are program evaluators, and we recognize the importance of the evaluation process to any organization's growth. Informal program evaluation is not unique to those of us who call ourselves program evaluators, of course. The two of us may differ, though, because we have studied the more formal processes of evaluation and the theories, constructs, and assumptions that underlie evaluative efforts. This understanding informs our informal evaluation on an almost subliminal level. In this introductory article, we will discuss the nature of evaluation, describing our understanding that underlies our formal and informal evaluative efforts.

So, simply stated, program evaluation is an inquiry or data collection activity that yields information about the value of a program – information that is useful for those who want to make programmatic decisions based on what they have learned. It is not that simple, however; even the title of this chapter is likely to raise the hackles of some evaluators, for evaluation is characterized by disagreements as well as by consensus. What counts as data for evaluation? Is program improvement the only purpose for evaluation? Must the results contribute to social betterment? Is it nitpicking to distinguish data from information from knowledge? The very word evaluate is based on the word value and therefore suggests a judgment. But who makes the judgment – the evaluator or the person(s) who commissioned the evaluation? Each of these questions applies to evaluations and the evaluation process. The term is also used to label, well, what? Is evaluation a profession, a field, a discipline, a trans-discipline? Is it a subset of research or is research just one tool of evaluation? Is it a science or an art?

If there were no disagreements about evaluation, a simple dictionary definition – or at least one in an introductory textbook – would suffice. However, to provide an overview of evaluation, this introductory chapter will go beyond a simple definition and discuss these and other issues.

Definitions, purposes, and uses

Although there is disagreement, most definitions of program evaluation involve three key concepts: (a) systematic inquiry; (b) judgment of merit, worth, value, or significance; and (c) information for decision-making. A selection of definitions from some prominent evaluation theorists illustrates the similarities along with differences in emphasis. Taken from different decades, the definitions also illustrate that the basic meaning of the term has not changed over the years. Although evaluative activity surely goes back to the beginning of time, Ralph Tyler is often credited with being the founder of modern evaluation.

A quick look at Tyler's¹ approach to educational evaluation suggests an objectives-based model in which program goals and behavioral objectives are established, data are collected about student performance, and performance is compared to objectives. Program evaluation, then, is a process of determining how effective a program is in helping clients to achieve behavioral objectives. Although Tyler may be known for what seemingly is a very rational, managerial, objectives-based approach to evaluation, he stressed the importance of determining the needs of stakeholders and addressing those needs in evaluations, with an aim to program improvement².

Another founding voice in evaluation is Lee Cronbach, who advocated for the potential of evaluation to contribute to our knowledge of and ability to deal with social problems. His “ninety-five evaluation theses” in his classic 1980 proclamation, *Toward Reform of Program Evaluation*, include the following emphases on the learning purpose of evaluation³:

- The evaluator is an educator; his success is to be judged by what others learn;
- Program evaluation is a process by which society learns about itself;

¹ Tyler, R.W., *Basic Principles of Curriculum and Instruction* (Chicago: University of Chicago, 1949).

² Nowakowski, J.R., *An Interview with Ralph Tyler* (November 1981). Available at <http://www.wmich.edu/evalctr/pubs/ops/ops13.html> (consulted 9 July 2004). Tyler, R.W., *Basic Principles of Curriculum and Instruction* (Chicago: University of Chicago, 1949).

³ Cronbach, L. and Associates, *Toward Reform of Program Evaluation: Aims, Methods, and Institutional Arrangements* (San Francisco: Jossey-Bass, 1980), pp. 2-11.

- Program evaluation should contribute to enlightened discussion of alternative plans for social action.

Smith⁴ also argued that evaluation is a social service activity, stating that its inquiry-oriented nature was secondary. He explained that evaluation (a) deals primarily with questions of value; (b) has identifiable clients (e.g., decision makers; policy makers; public); (c) attends to problems and issues defined by clients rather than by discipline; (d) is based on public application of method; (e) is time and cost bound; and (f) must be assessed for usefulness as well as for accuracy.

Weiss's definition is both elegant and simple: "Evaluation is the *systematic assessment* of the *operation* and/or the *outcomes* of a program or policy, compared to a set of *explicit or implicit standards*, as a means of contributing to the *improvement* of the program or policy".⁵ The Joint Committee on Standards for Educational Evaluation described program evaluation as "the systematic investigation of the worth or merit of an object".⁶ Newcomer, Hatry, and Wholey remind us of the importance of quality in the evaluation itself, noting that evaluation "provides processes and tools that agencies of all kinds can apply to obtain *valid, reliable, and credible data* to address a variety of questions about the performance of public and nonprofit programs".⁷

Patton⁸ identifies three interrelated components: (1) program evaluation is the systematic collection of information about (2) a range of foci such as activities, characteristics, and outcomes (3) for the purpose of making judgments that may inform decisions and improve program effectiveness. Patton's emphasis on use corroborates the view of evaluation as knowledge generation and learning. Newcomer et al. also emphasize learning: "evaluation is accepted here as a valuable *learning strategy* to enhance knowledge about the logic of the underlying program, as well as the practical results of the program".⁹

⁴ Smith, N.L., *The Kinds of Knowledge Required for Discipline-Level Assessments of Evaluation Methodology* (Portland: NWREL, 1978).

⁵ Weiss, C.H., *Evaluation*, 2nd edition (Upper Saddle River, NJ: Prentice-Hall, 1998), p. 4.

⁶ Joint Committee on Standards for Educational Evaluation, *The Program Evaluation Standards: How to Assess Evaluations of Educational Programs*, 2nd edition (Thousand Oaks, CA: Sage, 1994), p. 3.

⁷ Newcomer, K.E., Harty H.P. and Wholey, J.S., "Meeting the Need for Practical Evaluation Approaches: An Introduction", in J.S. Wholey, H.P. Harty and K.E. Newcomer (eds.), *Handbook of Practical Program Evaluation*, 2nd edition (San Francisco: Jossey-Bass, 2004), p. xxxiii, emphasis added.

⁸ Patton, M.Q., *Utilization Focused Evaluation: The New Century Text*, 3rd edition (Thousand Oaks: Sage, 1997).

⁹ Ibid., p. xxxiv, emphasis added.

Evaluators offer multiple perspectives on application or use of evaluation information. Patton¹⁰ highlights the importance of *instrumental use*, that is, intended uses by intended users. Weiss¹¹ suggests less clear links between results and utilization; she describes *enlightenment use* during which findings provide a background of understandings and ideas that “creep”¹² into decision deliberations. In *persuasive use*¹³, evaluation findings are used to argue on behalf of a policy or program. Rossman and Rallis¹⁴ introduce *symbolic use* that serves to legitimize or crystallize beliefs or values related to a program and *emancipatory use* when program participants produce and use evaluation information to improve their own work and lives.

Evaluation as learning: the inquiry cycle

Use is inherent in the learning process. According to Rossman and Rallis¹⁵, learning is the process that transforms data into information that can be used for a variety of purposes:

People gather data (sensory building blocks such as numbers, sounds, words, movements) as naturally as they breathe. They filter these data through their own unique experiences and existing understandings and make judgments about the meaning of the data. Patterns emerge, and the data coalesce into information. When a person uses the information in some way (to act, to decide, to form a new idea), the product is knowledge, and learning has occurred.¹⁶

This learning or inquiry cycle forms the foundation of evaluation and knowledge generation. The process itself is a cycle of questioning:

¹⁰ Patton, M.Q., *Utilization Focuses Evaluation: The New Century Text*, 3rd edition (Thousand Oaks: Sage, 1997).

¹¹ Weiss, C.H., “Improving the Use of Evaluations: Whose Job is it Anyway?”, in A.J. Reynolds and H.J. Walberg (eds.), *Advances in Educational Productivity: Evaluation Research for Educational Productivity*, vol. 7. (Greenwich, CT: JAI Press, 1998), pp. 263-276.

¹² Weiss, C.H., “Knowledge creep and decision accretion”, *Knowledge: Creation, Diffusion, and Utilization*, 1(3) (1980): 381.

¹³ Leviton, L.C. and Hughes, E.F.X., “Research on the utilization of evaluation: A review and synthesis”, *Evaluation Review*, 5 (1981): 525-547.

¹⁴ Rossman, G.B. and Rallis, S.F., *Learning in the Field: An Introduction to Qualitative Research*, 2nd edition (Thousand Oaks: Sage, 2003).

¹⁵ Rossman, G.B. and Rallis, S.F., “Evaluation as Learning: Critical Inquiry and Use as Action”, in V.J. Caracelli and H. Preskill (eds.), *The Expanding Scope of Evaluation Use, New Directions in Evaluation 88* (San Francisco: Jossey Bass, 2000).

¹⁶ Rossman, G.B. and Rallis, S.F., “Evaluation as Learning Critical Inquiry and Use as Action”, in V.J. Caracelli and H. Preskill (eds.), *The Expanding Scope of Evaluation Use, New Directions in Evaluation 88* (San Francisco: Jossey Bass, 2000), p. 56.

evaluators and program leaders ask epistemological questions: *What do we know about the program? What do we need to know? How will we learn it? What do we do with this knowledge?* In essence, as they participate in the learning cycle, program leaders are evaluators as well.

The cycle begins with establishing goals and expectations for the program: *What do program leaders and staff choose to do with program resources? What do they hope will happen because of these choices? How will participation in or implementation of this program help achieve this vision? What will success look like? How will they know?* Dialogue around these questions is informed by the program theory implicit, if not explicit, in the minds of the program leaders and explicates the criteria for judging the program. We will talk more about program theory after we describe the other parts of the cycle.

Next, the process focuses the inquiry on what program stakeholders actually want to learn about the program. This phase of the inquiry explores intriguing puzzles or troubling issues that stakeholders want or need to address, as well as purported areas of strength and weakness. Stakeholders may also realize that the evaluation must go beyond goals and objectives to look at unintended consequences. The specific questions identified at this point will guide data collection as well as making meaning from the data.

Data collection follows, drawing both from sources that already exist and from new sources. Questions related to data collection include: *What form will the data/evidence take, and where can we find it? How will we collect it, and how will we organize it so that we may make sense of it? What values will the data represent? What data and what perspectives are missing?*

The next activity in the cycle is mindful analysis and interpretation of the data in light of the program's articulated values. The information is interpreted in light of the program's purposes. In short, evaluators question the data and assign meaning to them. Guided by the focus questions, they group the data, noting patterns and rules, articulating relationships, and categorizing. They are alert for unexpected outcomes and surprises. These activities serve to clarify, to correlate, and to judge. Evaluators ask: *Who actually participates? What do program participants report is actually happening? What are they doing? Who is benefiting? According to whom? Is this program activity related to another? Did this activity influence this result? What is working and why? Are the activities congruent with our values?*

In this learning and use-focused view of evaluation, dialogue naturally shifts to action and change, to application: *What practices should continue, and how can we strengthen them? What practices do we need*

to change? What supports and resources do we need to improve or alter our practices? The success or shortcomings of past practices have meaningful consequences as program leaders make decisions that can have an impact on future practices.¹⁷

The construct of program theory plays an integral role in the evaluation/inquiry cycle as a mechanism for understanding a program's goals and how the program is designed to achieve those goals. As Weiss explains, program theory is: "the set of beliefs that underlie action", "a set of hypotheses upon which people build their program plans", and "an explanation of the causal links that tie program inputs to expected program outputs".¹⁸ Bickman defines program theory as "a plausible and sensible model of how a program is supposed to work",¹⁹ while Wholey notes that program theory identifies "program resources, program activities, and intended program outcomes, and specifies a chain of causal assumptions linking program resources, activities, intermediate outcomes, and ultimate goals".²⁰ In simple terms, if program personnel perform *intervention X*, they expect *result Y*. Implicit throughout is the act of valuing or the search for merit and worth.

In search of merit and worth

Merit and worth are not synonymous, although both are related to defining quality. Some evaluations may judge the program's essential merit, goodness, or success; others assess the program's worth, often in light of the expenditure of public monies to support the program; and yet others describe its essential attributes or characteristics. Merit is defined in the Oxford English Dictionary as "a point of intrinsic quality; a commendable quality, an excellence, a good point".²¹ Scriven defines merit as "the 'intrinsic' value of evaluands, as opposed to extrinsic or system-related value/worth".²² In this light, merit suggests

¹⁷ Rossman, G.B. and Rallis, S.F., "Evaluation as Learning: Critical Inquiry and Use as Action", in V.J. Caracelli and H. Preskill (eds.), *The Expanding Scope of Evaluation Use, New Directions in Evaluation 88* (San Francisco: Jossey Bass, 2000).

¹⁸ Weiss, C.H., *Evaluation*, 2nd edition (Upper Saddle River, NJ: Prentice-Hall, 1998), p. 55.

¹⁹ Bickman, L. (ed.), "Using Program Theory in Evaluation", in *New Directions for Program Evaluation Series*, no. 33 (San Francisco: Jossey-Bass, 1987), p. 5.

²⁰ Wholey, J.S. (ed.), *Organizational Excellence: Stimulating Quality and Communicating Value* (Lexington, MA: Lexington Books, 1987), p. 78.

²¹ Brown, L. (ed.), *The New Shorter Oxford English Dictionary* (Oxford: Clarendon Press, 1993), p. 1748.

²² Scriven, M., *Evaluation Thesaurus*, 4th edition, (Thousand Oaks, CA: Sage, 1991), p. 227.

that the evaluator's role is to determine the goodness or excellence of a program or service with the implication that judgments are also made about its shortcomings or weaknesses. Merit questions about a program include: How is it working? What is working? What is not working? How can we improve the program? This judgment centers on the program or service itself and is typical of the determinations evaluators are contracted to provide.

Worth moves beyond the program itself, exploring the extrinsic value of the program or service, that is, how it is assessed and valued by others. For example, an evaluation considers the worth of an initiative when it looks at the expenditure of monies for the program: are the program activities and services valuable enough to invest funds? Scriven notes that worth "usually refers to the value to an institution or collective [program or service], by contrast with intrinsic value, value by professional standards, or value to an individual consumer (merit)".²³ He further notes that, "when people ask 'what something is worth', they are usually referring to *market value* and market value is a function of the market's (hypothesized) behavior toward it, not of the thing's intrinsic virtues or its virtue for an identifiable individual".²⁴ Evaluators express worth when they make judgments about the overall meaning of the program to a particular set of constituents or to society in general, and when they make recommendations about funding.

Both of these judgments about merit and worth depend on detailed descriptions of the characteristics and attributes of complex social programs. These attributes – or qualities – may be found in the "thick descriptions"²⁵ that allow the evaluator and program personnel to interpret activities and events. Judgments of merit or worth are drawn from these thick descriptions and interpretations. The evaluator's role is to describe the "attribute, a property, a special feature or characteristic"²⁶ of a program, providing the grounding for making judgments about merit and worth.²⁷

²³ Scriven, M., *Evaluation Thesaurus*, 4th edition, (Thousand Oaks, CA: Sage, 1991), p. 382.

²⁴ Scriven, M., *Evaluation Thesaurus*, 4th edition, (Thousand Oaks, CA: Sage, 1991), pp. 382–383, emphasis in original.

²⁵ Geertz, C., *The Interpretation of Culture: Selected Essays* (New York: Basis Books, 1973).

²⁶ Brown, L. (ed.), *The New Shorter Oxford English Dictionary* (Oxford: Clarendon Press, 1993), p. 2438.

²⁷ For a more complete discussion of quality, merit, and worth in evaluation, see Rallis, S.F. and Rossman, G.B., "Communicating Quality and Qualities: The Role of the Evaluator as Critical Friend", in A.P. Benson, D.M. Hinn and C. Lloyds (eds.), *Visions of Quality: How Evaluators Define, Understand, and Represent Program Quality* (Oxford: JAI Press, 2001), pp. 107–120.

Because of its emphasis on program quality – value and worth, evaluation as practiced has evolved into an often resented and feared activity necessary for program survival. Both evaluators and consumers might well remember Cronbach and Associates' Thesis #15 that evaluation is “better used to understand events and processes for the sake of guiding future activities,” rather than “looking back in order to assign praise or blame”.²⁸ Thus, use re-surfaces: evaluations that seek to discover and explicate quality and qualities of programs can serve to make a practical difference in the program because the resulting descriptions provide a basis for making informed judgments about merit or worth. To provide the description, facilitate the analysis, and avoid the “blame game” – and thus promote utilization,²⁹ we suggest a specific evaluator role: that of critical friend.³⁰ This construction can help to eliminate another problem that can arise: “Decision makers who acquiesce to the expertise of the evaluator may later find that they neither understand nor believe in evaluation data”.³¹

The evaluator's search for merit and worth takes an ethical stance, codified in the American Evaluation Association's Guiding Principles for Evaluators and covering the following issues:

1. Systematic Inquiry: Evaluators conduct systematic, data-based inquiries about whatever is being evaluated.
2. Competence: Evaluators provide competent performance to stakeholders.
3. Integrity/Honesty: Evaluators ensure the honesty and integrity of the entire evaluation process.
4. Respect for People: Evaluators respect the security, dignity and self-worth of the respondents, program participants, clients, and other stakeholders with whom they interact.
5. Responsibilities for General and Public Welfare: Evaluators articulate and take into account the diversity of interests and values that may be related to the general and public welfare.³²

²⁸ Cronbach, L. and Associates, *Toward Reform of Program Evaluation: Aims, Methods, and Institutional Arrangements* (San Francisco: Jossey-Bass, 1980), p. 4.

²⁹ Patton, M.Q., *Utilization Focused Evaluation: The New Century Text*, 3rd edition, (Thousand Oaks: Sage, 1997).

³⁰ Rallis, S.F. and Rossman, G.B., “Dialogue for Learning: Evaluator as Critical Friend”, in R. Hopson and M.Q. Patton (eds.), *How and Why Language Matters in Evaluation, New Directions for Evaluation* 86 (San Francisco: Jossey-Bass, 2000).

³¹ Patton, M.Q., *Practical Evaluation* (Beverly Hills: Sage, 1982), p.141.

³² American Evaluation Association. Available at <http://www.eval.org> (consulted 9 July 2004).

These principles guide the conduct of inquiry for evaluators and offer consumers of evaluation a framework of what to expect.

Conducting the inquiry

As a quintessentially applied field, evaluation must serve both the social justice purposes of the program and the information needs of those who have contracted for the evaluation. This demand for high-quality, useful information characterizes the evaluation enterprise and sets it apart from research conducted for more theoretical or “pure” purposes.³³ Because of this information demand, evaluators are generally pragmatists who understand that evaluation research is a series of decisions that are compromises based on logistics, feasibility, stakeholder interests, the value stance of the evaluator, time, and other resources.

Certain activities are integral to the evaluation process, regardless of the overall approach or particulars of method. An important first step is identification of the purpose of the particular evaluation at hand (for example, is the purpose to plan program improvement? Is it to determine if the program is reaching its target audience? Is it to decide whether the program can expand? Is it to gauge client satisfaction?) Articulating program purposes presupposes a prior step – identification of the stakeholders. Purposes may vary with stakeholder. At the same time, some purposes will suggest that certain stakeholders are more important than others. The classic “chicken-egg” question can be applied throughout the evaluation process. Rarely must one specific step come before another, and the actual sequence will be dictated by the compromises mentioned above. Once the program leaders and evaluators have begun to focus the evaluation by identifying purposes and stakeholders, they will develop evaluation questions. Developing evaluation questions from purposes is similar to developing objectives from goals; it is a process of more clearly specifying the evaluation purpose so that decisions about data sources and evaluative criteria, for example, can be made.

Another preliminary step often employed is an *evaluability assessment*.³⁴ Before embarking upon a program evaluation, it is important

³³ Patton, M.Q., *Qualitative Evaluation and Research Methods*, 3rd edition, (Thousand Oaks: Sage, 2002).

³⁴ Wholey, J.S., “Evaluability Assessment”, in J.S. Wholey, H.P. Hatry and K.E. Newcomer (eds.), *Handbook of Practical Program Evaluation*, 2nd edition, (San Francisco: Jossey-Bass, 2004), pp. 33-62.

to consider whether sufficient resources are available to conduct a useful, credible evaluation and whether the results can actually influence decisions. This assessment can be done after the evaluation questions have been drafted, which may, however, result in their modification.

Answering evaluation questions involves three basic activities: *description, comparison, and prediction*. To describe, evaluators ask: What are the characteristics of the program? How are they organized into patterns? Next, they ask: How do the patterns compare to the patterns of similar programs? How are they similar and how are they different? How does this program compare to other programs or to defined standards? If the comparison is positive, the program can be judged to have merit or worth, to be of value. Finally, to be useful, the evaluation must also provide information for predicting the continued success of the program, or better success if changes are made, or whether the program can be successful in other contexts. A useful evaluation identifies characteristics of the program that made a difference; that is, it reveals which activities or resources appear most related to program success or failure. Identifying these activities or resources enables predictions that can lead to recommendations for program stability, termination, or improvement.³⁵

To gather the data that can describe, compare, and predict, evaluators identify or develop data collection instruments or methods, identify data sources or samples, and develop plans for data analysis. They often use mixed-methods designs, although some will prefer a specifically quantitative or qualitative approach.

Description, comparison, and prediction contribute to making value judgments about a program or policy that inform decision-making and lead to recommendations for improvement. These three activities, while present in all social inquiry, are crucial to making the judgments required in evaluation and are complex inter-related cognitive activities that are seldom performed in a linear process. Usually, all three activities are related and used iteratively in an evaluation. We must describe and understand in order to compare and predict; similarly, comparisons and testing predictions yield richer descriptions and deeper understandings. At one time or another during an evaluation, one activity will receive a relatively stronger emphasis. In fact, most evaluations begin with descriptive data collection because they need to build shared

³⁵ For a more complete discussion of these activities, see Rallis, S.F. and Rossman, G.B., "Mixed Methods in Evaluation Context: A Pragmatic Framework", in A. Tashakkori and C. Teddlie (eds.), *Handbook of Mixed Methods in the Social and Behavioral Sciences* (Thousand Oaks: Sage, 2002).

knowledge about the program. Some evaluators begin by making explicit the logical models or theory base, but others build the logical model or generate the theory as they learn about the program. The model or theory then shapes both the design and the analysis as the process continues. Ultimately, the model becomes part of the prediction.

Conclusion

The goal in program evaluation, then, is to learn, in order to make judgments about the program that may lead to improvement. The judgments are evaluative, based on specified criteria, and must be assessed for usefulness. Grounded in an ethic that values competence, honesty, and integrity, that respects individual rights and dignity, and that recognizes diversity of interests and values, evaluation aims for social betterment.

Evaluators choose methods to collect and analyze data according to the questions being asked about the program. Given the program goals, objectives, and context, and given available resources, evaluators ask what are the best methods to help understand the program's merit, to make the most useful comparisons that reveal merit, and/or to predict worthy outcomes? Because evaluations have multiple purposes that result in learning and judging, many evaluators draw from the multiple methods for data collection and analysis that are available to them. The articles to follow will explicate approaches and methods currently in use in archival science and will provide examples of evaluation practice in libraries, in museums, and in records and/or information management services.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.