

Evaluating Food Stamp Nutrition Education: A View from the Field of Program Evaluation

Ellen Taylor-Powell, PhD

ABSTRACT

The purpose of this paper is to review current approaches and issues in the field of program evaluation that apply to the effort of improving the evaluation of Food Stamp Nutrition Education (FSNE). As nutrition educators respond to increasing demands for accountability and measuring results, there are challenges to consider, as well as opportunities to capture. This paper includes a focus on the internal and external functions of evaluation, with attention given to evaluation's internal use and value in improving practice. A list of 10 building blocks is offered as essential for planning useful and credible evaluation of FSNE.

KEY WORDS: evaluation, performance measurement, outcomes, FSNE, community-based program evaluation

(*J Nutr Educ Behav.* 2006;38:12-17.)

INTRODUCTION

As reported in this special section,^{1,2} an extensive national effort is underway to improve the evaluation of Food Stamp Nutrition Education (FSNE), an optional component of the United States Department of Agriculture (USDA) Food Stamp Program (FSP) that provides nutrition education to program participants via a federal-state partnership.¹ This effort is set within a performance-based management and accountability context that has gained momentum since passage of the Government Performance and Results Act (GPRA, or Results Act)³ in 1993 requiring all federal agencies to set goals, measure performance, and report accomplishments. "Outcomes" has become part of our everyday lexicon, and logic models are in widespread use, even standard agency practice, as a convenient way to depict the intended sequence of connections between program investments, operations and results.⁴⁻⁸ The results-oriented emphasis has created considerable activity as agencies and programs struggle to respond, and nutrition education is no exception. Early indications suggest that the effort to measure results pays off in improved programmatic and organizational performance.⁹ But the road is not easy, and the challenges are many. The purpose of this paper is to review experience from the field of program evaluation as it applies to improving FSNE evaluation. Challenges and opportunities will be discussed, as will the

internal and external functions of program evaluation with implications for improving nutrition education evaluation.

CHALLENGES AND REAL WORLD CAUTIONS

Fear and skepticism. For many, engaging in evaluation elicits fear and trepidation. Images emerge of programs being judged and targeted for reductions, or of complicated statistical analyses that require sophisticated technical expertise. Skepticism prevails about the role, ability, and value of evaluation to validly reflect one's work and its impact ("How can any data adequately describe this complex program?"). And there is skepticism about the freshness and force of the accountability agenda ("We've heard this before"). Demystifying and establishing evaluation's value is a challenge that must be overcome.

Confusing language. There also is all the evaluation jargon, terminology, variety of methods and approaches. Individuals come to evaluation from different backgrounds and training. Some may equate evaluation and research. Others may impose the scientific method as the sole paradigm in the search for *truth* and *proof*. There may be multiple interpretations of what evaluation means, its purpose and possibilities. In fact, much of what people call *program evaluation* may be more precisely labeled *performance measurement*.^{10,11} Performance measurement is "the ongoing monitoring and reporting of accomplishments, particularly progress towards preestablished goals."¹² It typically uses quantitative indicators as part of routine reporting systems to provide evidence of outputs and outcomes. It addresses *what* is occurring and achieved but does not address *how* and *why* programs perform as they do, relevant

Evaluation Specialist, University of Wisconsin-Extension

Address for correspondence: Ellen Taylor-Powell, PhD, University of Wisconsin-Extension, 432 N. Lake Street, Room 609, Madison, WI 53706; Tel: (608) 262-2169; Fax: (608) 262-9166; E-mail: ellen.taylor-powell@ces.uwex.edu
©2006 SOCIETY FOR NUTRITION EDUCATION
doi: 10.1016/j.jneb.2005.11.007

for program improvement decisions or attributing effects.¹¹ Program evaluation, on the other hand, is a broad concept that includes performance measurement, as well as other types of assessment activities to address the variety of questions raised about programs.¹³⁻¹⁹ For example, process evaluation answers questions about what is provided to whom; how programs operate; fidelity of implementation and why there are gaps between program plans and program implementation.^{20,21} Outcome evaluation answers questions about whether the program achieved the intended results, and who benefits and who does not.^{22,23} Impact evaluation answers questions about the net effect of programs, and whether and to what extent the program *caused* the observed changes compared to what would have happened without the program.^{16,24} Each type of evaluation has a specific purpose for an intended use, with questions it seeks to answer that determine design and data collection methods. Data that are relevant to support budgetary and accountability demands are quite different from data to effectively inform program improvement decisions. Untangling our language and understandings will help improve evaluation.

Multiple users. Useful evaluation starts with understanding who wants to know what, for what purpose or use.¹⁵ But in most program settings, there are multiple information users with differing information needs, values, priorities, and expectations. They may not know or be unable to articulate their information needs in a timely and concise manner, or their information needs change as the external environment changes. For FSNE, this complexity is even greater. Implementation involves multiple layers of users: Congress; USDA; implementing agencies within the USDA (Food and Nutrition Service (FNS) and the Cooperative State Research, Education, and Extension Service (CSREES)), with national, regional, and state administrations and program staffs within each agency; other partners (eg, state health departments) and stakeholders; professional colleagues; as well as the FSNE audience and the public. Understanding, coordinating, and prioritizing information needs across these diverse groups is understandably complex. The improved communications that result from engaging in performance measurement and evaluation may be the key benefit.²⁵

Program variability. Another significant challenge is imposed by the great variation in program contexts, content, intensity, audience (demographics and characteristics, including readiness to change), and implementation, including who is actually delivering the program. Administrative structures, resources, and guidelines at the national level vary. States and localities have substantial flexibility in program implementation. Even the term *program* may have no standard meaning. At the local level, it may be a 2-hour workshop or a 3-year,

comprehensive community initiative. The audience may be elderly food stamp recipients or a broad range of community residents and policy makers. Critical questions are likely to pertain to the level and quality of program implementation and what can be expected from different types of implementation.

What to measure. Given this program variability and multiple information users, deciding *what* to measure is another major challenge. The current emphasis is on outcomes—changes achieved from programmatic investments; for example, changes in dietary behavior as measured by preprogram and postprogram behaviors. But determining outcomes is difficult and arbitrary. There are likely to be multiple desired outcomes. Some are more immediate and can be readily associated with program effort, for example changes in knowledge and skills. The more impressive, community- or population-wide outcomes—increased food security, reduced obesity—may take many years to observe and are influenced by multiple factors, often outside the program's control. Program contexts vary, but without identifying common outcome measures, it is impossible to aggregate information and tell a larger story. On the other hand, creating a set of outcome measures to achieve uniformity of meaning may result in meaningless information²⁶ or even be counterproductive.²⁷ Also, in the effort to *measure* outcomes, there is the tendency to focus only on quantitative indicators of outcomes and on predescribed, *positive* outcomes, overlooking the opportunity to understand what led to those outcomes, what works or doesn't work for whom and why, or what unexpected results happen, positive or negative.

Rigor and practicality. As in all research and evaluation, tradeoffs must be made between rigor and practicality. Individuals trained in the medical or scientific paradigm seek objective, rigorous, science-based information. Experimental or quasi-experimental designs are the gold standard to show that the program *caused* the observed change. Yet such designs are expensive, time-consuming, and often inappropriate for community-based programs. Design decisions are likely to depend upon circumstances and resources, how data are to be used, and professional judgment.²⁸ Balancing the quest for rigor with real-world realities is likely to remain a challenge.

Data quality. Finally, but by no means least importantly, all programs face a major challenge in ensuring that data are credible and trustworthy. For FSNE, that challenge is compounded, given its multilevel structure and the many individuals involved in data collection and reporting. A study of 13 federal agencies found such problems as unclear definitions of data elements and inadequate training, resulting in different interpretations and variation in data recording and entry; use of different time frames (eg, varying fiscal

years) and different recording systems, making data aggregation difficult.²⁹ The challenge of data quality involves more than just the technical issue of what constitutes quality information; equally important is the organizational environment that encourages responsibility and facilitates commitment to quality data.

OPPORTUNITIES

Despite such challenges, there are unparalleled opportunities as nutrition education evaluation becomes integrated into routine organizational operations.

Size and prominence. First is the size and prominence of FSNE. *Improving the nation's nutrition and health* is one of USDA's five strategic goals for fiscal years 2002-2007.³⁰ Nutrition education in FSP has grown from 7 states in 1992 to 52 state agencies in fiscal year 2004 (50 States, the District of Columbia and Virgin Islands) with an approved federal budget of \$228 million.³¹ This visibility and size provide the incentive and resource base to create an evaluation system that can credibly tell the FSNE story and be a model for others.

Partnerships. Second is the opportunity to capitalize on and strengthen existing partnerships, as well as create new alliances. Collaboration is the *modus operandi* for effective and efficient programming in the 21st century, as well as a policy mandate.³² FSNE is breaking new ground in the cooperative development of outcome measures to assess dietary behavior, data collection systems, and reporting frameworks. Collaboration, however, is not easy or automatic. Experience shows that time and resources must be invested in the *process* of the partnership if desired programmatic outcomes are to be achieved.

Evaluation methods and resources. In the last 20 years, the field of evaluation has exploded, providing a myriad of approaches, methods and tools. Choice is driven by the question being addressed and the resources available, but the preeminence of one paradigm or approach no longer exists. There is opportunity to be innovative and creative in applying a mix of quantitative and qualitative methods, such as a compilation of purposefully selected case studies to provide evidence of attribution; focus group interviews; and personal stories to complement quantitative data. There are established principles of ethical behavior,³³ standards of practice,³⁴ and increasing numbers of trained and experienced evaluators whose involvement can benefit the performance measurement team.^{29,35}

Technology. The Internet and advances in information technologies make rapid communications, nationwide information sharing, and automated data collection and re-

porting possible. Those who collect and submit data can immediately see their input, as well as access aggregated data sets and create customized reports. This enables localized use, fundamental to staff commitment and enhanced accountability, and makes it possible to readily assess data accuracy. Likewise, the Internet makes it possible to share widely standardized resources and training to support more uniform data collection and reporting. The Community Nutrition Education web site (www1.uwex.edu/ces/lmcourse/) is a first-of-a-kind in this regard. Composed of two self-instructional, interactive learning modules, individuals first develop skills in logic modeling as a tool for enhancing program development and evaluation. Then they are introduced to a logic model framework for community nutrition education and a standardized reporting protocol that captures and communicates program performance across the entire nation.

Existing research and resources. Nutrition education benefits from an extensive field of basic and applied research that continues to expand. It can aid in the setting of realistic performance targets; help in educating providers of funding about what is feasible to achieve and measure; and be shared with staff so that best practices are employed. Existing evidence may be available that supports or verifies hypothesized causal connections or may be used, in conjunction with a program logic model, to justify proposed indicators of performance. Knowledgeable and experienced researchers might be contracted for specialized studies. Peer or expert panels can be invited to provide judgments and evidence of outcomes. There may also be existing databases and surveillance systems (such as the Behavioral Risk Factor Surveillance System³⁶) that can be used to monitor trends or provide complementary data.

HOW DOES EVALUATION SUPPORT GOOD NUTRITION EDUCATION PRACTICE? WHAT'S IN IT FOR YOU?

Experience shows that if evaluation is the result of ill-conceived mandates or is viewed only as creating reports for external reporting, program staff are likely to be uncommitted, possibly even resistant; data quality will be poor; reporting will be inconsistent; and there will be minimal use of data in program decision making or management. What matters to staff is evaluation's internal use and value in program improvement.^{15,37,38}

Internal value. What is the internal value of evaluation for nutrition educators?

- *Bragging rights are legitimized.* Evaluation provides the basis for "legitimately bragging" about your accomplishments.⁹
- *Purpose is clarified.* Focusing on outcomes, reporting strategic goals and performance data increases understanding

of purpose, heightens intentionality, and promotes communication within organizations.^{11,39}

- *Results are achieved.* While too narrow a focus on outcomes is problematic and the determination and definition of results need careful orchestration, measuring results produces results.^{9,40}
- *Planning and budgeting are improved.* Credible and quality information contributes to better planning, and improved decision making and resource allocations.
- *Practice improves.* Asking critical questions of programs, checking assumptions, operationalizing program concepts, monitoring differential changes and impacts triggers corrective action and leads to improved practice.³⁸
- *Professional growth is enhanced.* Engaging in evaluation creates important changes in the attitudes and behaviors of those involved above and beyond the findings or use of the evaluation; what Michael Patton coined as the “process use” of evaluation.¹⁵
- *Resources are generated.* Increasingly, funding decisions are dependent upon the submission of evaluation plans, logic models, and the reporting of results.
- *Scholarship is demonstrated.* Evaluation provides an opportunity to demonstrate scholarship, that which adds to the nutritional science and evaluation knowledge bases.
- *Recognition increases.* Evaluation provides the opportunities and data to recognize and celebrate individual, team, and program excellence.
- *Organizational learning occurs.* When evaluation is thoughtful and purposeful, it involves critical examination, reflection, discovery, learning, and corrective action, which are the foundation of a learning organization.

External value. The other category of benefits pertains to evaluation’s value in external communications and promotion:

- *Demonstrates accountability.* Providers of funding and the public see an organization fulfilling its stewardship role. Making a good-faith effort at outcome measurement is an initial focus of wise providers of funding²⁷ and can create credibility and attention even prior to full implementation.⁴¹
- *Generates resources.*
- *Extends visibility of agency and programs.*
- *Promotes agency and programs.*
- *Attracts volunteers, partners, and clients.*
- *Leverages resources.*
- *Spreads knowledge and best practices.*

Both internal and external functions are important, but caution is advised to ensure that the external function does not outweigh the internal function or become the driving force for evaluation.

TOP 10 EVALUATION BUILDING BLOCKS

Several ingredients are necessary for the internal and external value of evaluation to be realized. The following “top

10” are considered essential building blocks, if staff are to engage in useful and credible evaluation.

1. **Leadership.** Fundamental is leadership and commitment at every level in order to overcome skepticism, demonstrate value, ensure appropriate measures and data quality and integrate the principles of self-examination and improvement. Senior administrators and management must create a climate that values results, communicate the importance of evaluation for accountability and learning, make data quality and use of data organizational goals, and provide the necessary financial and technical support.²⁸ An influential champion is often key for building momentum and sustaining the evaluation vision.⁴²⁻⁴⁴
2. **Stakeholder engagement.** Engaging stakeholders from oversight bodies, clientele groups and programming units is commonly the initial step in evaluation planning. It takes time and resources but is important for building consensus and understanding, for developing appropriate and locally relevant measures, for developing feasible data management systems, and for creating buy-in. Integrating staff and encouraging their active involvement in evaluation design and implementation bears directly on subsequent data quality and use.
3. **Common language and understanding.** Language and understanding become particularly important when working across organizations, units and program sites, or disciplines and content areas. Experience in all sectors demonstrates the importance of effective horizontal and vertical communications. Inconsistent and unreliable data result when standards and common definitions are lacking.⁴⁵ At the level of data collection, exact terminology and definitional understanding of concepts, terms and procedures are even more important in order to ensure consistent data collection across sites, among different staff, and over time as staff turnover occurs.
4. **Start with what.** Evaluation is often equated with data collection and defined as a survey or quantitative and qualitative methods. Program managers and staff typically fall into discussions of *how*—measurement, methods, and instrumentation—before establishing a clear understanding and shared commitment of *what* is to be measured. Defining *what* to measure—the questions one intends to answer—reflects the *intended* use of the data and the *value priorities* of the involved stakeholders; it is an inherently political process. Logic model development with key stakeholders helps to differentiate outputs from outcomes, promotes understanding of the program’s theory of action, and helps in identifying *what* to measure, including the external factors that might affect progress and results. Then choices can be made about design, measures, and data collection methods.
5. **Integrate evaluation.** Evaluation is often an add-on, a separate function, commonly inserted at the end of a

program. Yet if evaluation is equated with asking critical questions about programs in order to have useful information for decision making, it is *part* of programming. Planning evaluation during program design and integrating evaluation into programming helps ensure the collection of essential data throughout the life of a program, is more cost effective, and builds quality assurance procedures into ongoing work to become a part of normal operations.

6. **Prioritized and streamlined data collection.** Assessing the value added of information as well as quality issues will help streamline and reduce data burden. While data quality standards are likely to vary because of differing professional practices, contexts, and intended use of data, choices need to be made about characteristics to emphasize and standards to use. All data have limitations. Acknowledging limitations and the role of external factors in observed outcomes builds credibility and enhances appropriate interpretation and use of data.
7. **Showcasing use.** The lack of awareness among program staff as to evaluation's value and use to them is one of the biggest barriers to improving evaluation. Reporting systems tend to be constructed to provide data for external reporting or public relations purposes. Using evaluation internally in policy, program, and resource allocation decision making is considered the best incentive for fostering results-oriented management and for improving data quality.⁴⁶
8. **Investments.** The types of changes necessary to create and maintain quality evaluation require investment—material and human—that demands political will from the top.¹¹ Training has proven to be an important tool in organizational change. Greater proficiency at all levels in strategic planning, evaluation planning, measurement, and using data leads to more accurate and useful information.^{29,39} Experience also demonstrates the limited value of one-shot or sporadic training and the value of customized training to meet differing needs and learning styles, complemented by mentoring and coaching. Continuous investment in training is vital. Another significant investment involves the building, upgrading, and maintenance of data information systems and technologies for collecting, analyzing, and communicating information. Successful, results-oriented organizations make significant investments in information management.⁴⁶
9. **Procedural, policy and structural changes.** Training staff and/or investing in the latest technologies and information management systems, however, is not sufficient. Equally important are accompanying procedural and structural changes that demonstrate the organization's commitment to learning from evaluation and improving data quality; for example, administrative structures that encourage the coordination of data collection across program areas and sites; inclusion of evaluation responsibilities in job descriptions and grant

applications; and valuing evaluation through incentive and reward systems.

10. **Practice, practice, practice.** Experience at the Center for Disease Control and Prevention shows that people and organizations learn evaluation by doing it.⁴² Other organizations find the same. Initial efforts may be less than perfect, but continued practice pays off. Resources, incentives, and structures that encourage staff to practice evaluation result in the collection and use of quality data, leading to enhanced programmatic and organizational performance.

IMPLICATIONS FOR RESEARCH AND PRACTICE

All federal programs are under pressure to measure performance and report results. Nutrition education is not alone or atypical. Many of the challenges faced by the decentralized nutrition education programs also are common to other community-based programs. Learning from and adding to the expanding theoretical and empirical research base of community-based program evaluation will benefit all. In this work, FSNE has some unique opportunities based on the size and importance of the FSP, the existing nutritional research base, and collaborative relationships among agency partners, as well as among researchers, practitioners, and evaluation professionals. Demonstrating the value of evaluation and educating nutrition educators and staff at all levels in how to use evaluation data will provide the momentum for improvement. Ongoing training and support will be important, as will be the institutional structures and investments that encourage individuals to practice and learn from evaluation.

ACKNOWLEDGMENTS

The author expresses deep appreciation to the Society of Nutrition Education and the Economic Research Service of USDA for the invitation to participate in the July 2003 SNE FSNE workshop, Advancing the Practice of Food Stamp Nutrition Education Evaluation: Art and Science, and then for their encouragement and assistance to turn her discussion remarks into this manuscript.

REFERENCES

1. Guthrie J, Stommes E, Voichek J. Evaluating Food Stamp Nutrition Education: Issues and Opportunities. *J Nutr Educ Behav.* 2006;38:6-11.
2. Townsend M. Evaluating Food Stamp Nutrition Education: Process for development and validation of evaluation measures. *J Nutr Educ Behav.* 2006;38:12-18.
3. Public Law 103-62. *Government Performance and Results Act of 1993*; Aug 3, 1993.
4. Wholey J. *Evaluation: Promise and performance.* Washington, DC: Urban Institute Press; 1979.
5. Wholey J. Evaluability assessment: Developing program theory. *New Directions for Program Evaluation.* 1987;33:77-92.

6. McLaughlin JA, Jordan GB. Logic models: A tool for telling your program's performance story. *Eval Prog Plann.* 1999;22:65-72.
7. Taylor-Powell E, Jones L, Henert, E. Enhancing Program Performance with Logic Models. Available at: <http://www1.uwex.edu/ces/lmcourse/>. Accessed December 2002.
8. Millar A, Simeone R, Carnevale J. Logic models: A systems tool for performance management. *Eval Prog Plann.* 2001;2:73-81.
9. Hakes JE. Can measuring results produce results: one manager's view. *Eval Prog Plann.* 2001;24:319-327.
10. US General Accounting Office. Program Evaluation: Agencies challenged by new demand for information on program results. Washington, DC: US General Accounting Office; April 1998. Publication GAO/GGD-98-53.
11. Newcomer KE. Using performance measurement to improve programs. *New Directions for Evaluation.* 1997;75:5-14.
12. United States General Accounting Office. Performance measurement and evaluation: Definitions and relationships. Washington, DC: US General Accounting Office; April 1998. Publication GAO/GGD-98-26.
13. Newcomer KE. Evaluating Public Programs. In: Perry JL, ed. *Handbook of Public Administration.* 2nd ed. San Francisco: Jossey-Bass Publishers; 1996.
14. Hatry HP. *Performance Measurement.* Washington, DC: The Urban Institute; 1999.
15. Patton MQ. *Utilization-Focused Evaluation.* 3rd ed. Thousand Oaks, CA: Sage Publications; 1997.
16. Rossi PH, Freeman HE. *Evaluation: A Systemic Approach.* 5th ed. Newbury Park, CA: Sage Publications; 1993.
17. Worthen BR, Sanders JR. *Educational Evaluation: Alternative Approaches and Practical Guidelines.* New York, NY: Longman; 1987.
18. Weiss C. *Evaluation.* 2nd ed. Upper Saddle River, NJ: Prentice-Hall; 1998.
19. Wholey JS, Hatry H, Newcomer K, *Handbook of Practical Program Evaluation.* Englewood Cliffs, NJ: Prentice-Hall; 1994.
20. Judd CM. Combining process and outcome evaluation. *New Directions for Program Evaluation.* 1987;35:23-41.
21. Scheirer MA. Process evaluation. In: Wholey J, Hatry H, Newcomer KE, eds. *Handbook of Practical Program Evaluation.* San Francisco: Jossey-Bass; 1994.
22. Reisman J, Clegg J. *Outcomes for Success! The Evaluation Forum: Organizational Research Services, Inc., and Clegg and Associates;* 1999.
23. W.K. Kellogg Foundation. *W.K. Kellogg Foundation Evaluation Handbook.* Battle Creek, MI: W.K. Kellogg Foundation; 1998.
24. Mohr LB. *Impact Analysis for Program Evaluation.* 2nd ed. Thousand Oaks, CA: Sage Publications; 1995.
25. Kettl DF. In: Newcomer KE. Using performance measurement to improve programs. *New Directions for Evaluation.* 1997;75:9.
26. Figueredo AJ, Sechrest L. Approaches used in conducting health outcomes and effectiveness research. *Eval Prog Plann.* 2001;24:41-59.
27. Plantz MC, Greenway MT, Hendricks M. Outcome measurement: Showing results in the nonprofit sector. *New Directions for Evaluation.* 1997;75:15-30.
28. Divorski S, Scheirer MA. Improving data quality for performance measures: Results from a GAO study of verification and validation. *Evaluation and Program Planning.* 2001;24:83-94.
29. Scheirer MA, Newcomer KE. Opportunities for program evaluators to facilitate performance-based management. *Eval Prog Plann.* 2001; 24:69.
30. United States Department of Agriculture. United States Department of Agriculture: Strategic Plan for FY 2002-2007. Available at: <http://www.usda.gov/ocfo/pm/sp2002.htm>. Accessed June 2004.
31. National Food Stamp Nutrition Education. Approved Federal Funds for Food Stamp Nutrition Education by Fiscal Year. Available at: <http://www.nal.usda.gov/foodstamp/FSNEFederalFundsbyFiscalYearpdf.pdf>. Accessed June 2004.
32. United States General Accounting Office. Nutrition Education: USDA provides services through multiple programs, but stronger linkages among efforts are needed. Washington, DC: US General Accounting Office; 2004. Publication GAO-04-528-204.
33. American Evaluation Association. Guiding Principles for Evaluators. 1994. Available at: <http://www.eval.org/EvaluationDocuments/aeaprin6.html>. Accessed April 2004.
34. Joint Committee on Standards for Educational Evaluation. *The Program Evaluation Standards.* Thousand Oaks, CA: Sage Publications. 1994. Available at: <http://www.eval.org/EvaluationDocuments/progreval.html>. Accessed April 2004.
35. Wholey JS, Newcomer KE. Clarifying goals, reporting results. *New Directions for Evaluation.* 1997;75:91-98.
36. CDC Web site. Behavioral Risk Factor Surveillance System. Available at: <http://www.cdc.gov/brfss/#questionnaire>. Accessed January 5, 2006.
37. Sanders JR. Mainstreaming Evaluation. *New Directions for Evaluation.* 2003;99:3-6.
38. Hatry HP, Morley S, Rossman SB, Wholey JS. *How Federal Programs Use Outcomes Information: Opportunities for Federal Managers.* Washington, DC: National Academy of Public Administration; 2003.
39. Urban Institute. How and why nonprofits use outcome information: Findings from a symposium. Washington, DC: The Urban Institute; 2002.
40. Osborne D, Gaebler T. *Reinventing Government: How the Entrepreneurial Spirit is Transforming the Public Sector.* New York, NY: Plume Book; 1993.
41. Taylor-Powell E. Bricks and mortar: Capacity building top to bottom and inside out. Presentation at: Annual American Evaluation Association Conference, Session #324; November 2003; Reno, NV.
42. Milstein B, Chapel T, Wetterhall S, Cotton D. Building capacity for program evaluation at the Centers for Disease Control and Prevention. *New Directions for Evaluation.* 2002;93:27-46.
43. Mackay, K. The World Bank's ECB experience. *New Directions for Evaluation.* 2002;93:81-101.
44. Bernstein DJ. Local government measurement use to focus on performance and results. *Eval Prog Plann.* 2001;24:95-101.
45. United States General Accounting Office. Managing for results: Challenges agencies face in producing credible performance information. Washington, DC: US General Accounting Office; 2000. Publication GAO/GGD-00-52.
46. United States General Accounting Office. Executive Guide: Effectively implementing the government performance and results act. Washington, DC: US General Accounting Office; June 1996. Publication GAO/GGD-96-118.