

A Synopsis of the American Journal of Evaluation: Volume 26 (3)

Bianca E. Montrosse

Claremont Graduate University

The current issue of the *American Journal of Evaluation* (AJE) is one of the more comprehensive issues, in that articles for seven of the ten AJE sections are included. What follows is a summary of the articles in the order in which they appear.

Articles

In “The Use of Multiple Evaluation Approaches in Program Evaluation,” Katrina Bledsoe and James Graham argue that the use of multiple evaluation approaches can lead to evaluations that are scientifically credible, valid, and useful. Bledsoe and Graham begin by describing the program and the purpose of their evaluation. In the introduction, they also briefly outline the four evaluation models (e.g., Empowerment, Theory-Driven, Consumer-Based, and Inclusive) that guided their evaluation efforts. Bledsoe and Graham begin their analysis by discussing how the utilization of these approaches lead to the development of program recommendations, highlighted how future evaluations could establish a continuous

evaluation cycle, and the identification of side effects. Further, they highlight several challenges that arose as a result of utilizing mixed approaches. Bledsoe and Graham conclude with implications for evaluation practice, practitioners, and theorists.

The second article, “Is Sustainability Possible? A Review and Commentary on Empirical Studies on Program Sustainability” by Mary Ann Scheirer, presents results from her synthesis of eighteen health related evaluations. The study examined the type, extent of, and factors contributing to program sustainability. Scheirer found support for three difference types of sustainability: Individual, Organizational, and Community. Despite data limitations, most reported the achievement of some type of sustainability. From those results, she identified five factors that contribute significantly to sustainability: (a) an emergent program design; (b) an effective program champion (e.g., executive director); (c) a congruent fit with organization’s philosophy and structure; (d) the extent to which staff and/or clients perceive program benefits; and (e) community support. Although her sample only includes healthcare evaluations, Scheirer argues that her findings are generalizable to other content areas.

Forum

In “Integrating Personnel Evaluation in the Planning and Evaluation of School Improvement Initiatives,” Anthony Normore argues for the use of personnel evaluation in school improvement initiative evaluations. In this article, personnel evaluation is not used in the traditional sense (i.e., evaluators conducting personnel evaluation). Rather, it is defined as the evaluator examining personnel evaluation procedures, and particularly its influence on school improvement initiatives.

Normore further asserts that personnel evaluation can provide integral and necessary information for conducting sound school improvement initiative evaluations for several reasons. For example, in the formative sense, it can serve as a change catalyst by improving teacher evaluation procedures which, in turn, creates a consistent improvement process. In the summative sense, it can shed light on the degree of congruence between evaluation practices and reforms. Personnel evaluation also has the potential to provide detail about teacher evaluations and learning, which research has found moderates the relationship between initiatives and their outcomes. Finally, it can impact reform implementation on the part of the teachers.

Ethical Challenges

This article marks the first piece published by the new Ethical Challenges section editor, Leslie J. Cooksy. In her opening statement, she expresses her desire to honor Michael Morris's (i.e., the previous section editor) contribution, while also introducing new formats for discussing ethics in evaluation.

In "The Complexity of the IRB Process: Some of the Things You Wanted to Know About IRBs but Were Afraid to Ask," Cooksy answers fifteen questions related to the IRB process with assistance from Charles Hoehne, Walton Francis, and Robin Miller. In the article, she describes the IRB as a board that is composed of representatives from various disciplines that seeks to ensure participants and their well being are protected. The questions Cooksy addresses can be divided into five major areas: (a) how evaluations are subject to IRBs, (b) IRB process and materials, (c) the distinction between privacy, confidentiality, and anonymity, (d) consent versus assent, and (e) the use of incentives. In discussing whether evaluations are

subject to IRB review, Cooksy states that although some evaluations fall under the exemption category, most require IRB approval, even in international settings. For evaluators who are not affiliated with a university, and thus, do not have access to an IRB board, Cooksy recommends that a member of the evaluation team speak to the project officer concerning IRB requirements. If an evaluation is subject to IRB review, then the team is required to submit a detailed explanation of the evaluation, all data collection instruments and supporting documents, and a copy of the grant (if applicable). In addition, some IRBs require documentation that all members of the team have met research training requirements. The length of time for approval varies from institution to institution for a variety of reasons (e.g., number and frequency of IRB board meetings, incomplete IRB applications, the need for an outside expert opinion).

One of the major areas of confusion for most evaluators is the difference between privacy, confidentiality, and anonymity. Privacy is defined as stakeholders having a voice in determining what, when, and how their information is shared with the evaluation team. Confidentiality guarantees that stakeholders cannot be identified by their responses. Evaluators can ensure that responses are kept confidential by utilizing codes, removing identifying information, shredding documents, storing information in locked cabinets, and password protecting data stored on computers. Anonymity can only be guaranteed when evaluation team members cannot identify respondents (e.g., when conducting random-digit dialing or public observation).

In most cases, content or assent forms are required. Consent is defined as someone of legal age (i.e., 18 years of age or older) agreeing to participate in data collection. Because those under the legal age may not necessarily understand all aspects of participation, assent is required. When parents give their permission, even though

they will not be participating, along with their children, it is defined as assent. Special assent rules apply to wards of the state.

Because there are no clear rules regarding the use of incentives, most IRBs create their own guidelines. Regardless of these rules, IRBs are required to ensure that the use of incentives does not coerce potential respondents.

Exemplars

This is the first interview Christina Christie conducts as Exemplar section editor. In the first article in this two-part series entitled “The Colorado Health Communities Initiative,” Ross Conner briefly describes the initiative, its principles and model, its communities, and its major outcomes. Briefly, the purpose of Connor’s evaluation was to examine program implementation across sites, identify short-term outcomes, and investigate longer term outcomes.

In “A Conversation with Ross Conner: The Colorado Trust Community-Based Collaborative Evaluation,” Christie asks Conner a series of questions related to the evaluation design and implementation. She concludes with a brief commentary that examines the relationship between Connor’s theoretical approach and his practical procedures.

The Historical Record

This entry marks the third installment of “The Oral History of Evaluation” project. This project “continues Jean King, Mel Mark, and Robin Miller’s effort to document the principle intellectual influences on individuals who were part of the

pivotal moments in the field of program evaluation” (The Oral History Project Team, 2005). This interview documents the evaluation journey of Michael Scriven.

Like many evaluators, Michael Scriven’s evaluation journey began informally. He embarked on his first evaluation in high school by entering an essay competition and arguing for the feasibility of developing a system for evaluating heroism. His dream of becoming a fighter pilot in the Royal Air Force (RAF) was ended when World War II ended. Thus, having taken numerous science and mathematics courses, which was required for entry into the RAF, he decided to attend the University of Melbourne where he majored in Science and Mathematics.

After completing his formal schooling, Scriven took a job at the University of Minnesota. It was there that he began to become interested more formally in evaluation and published his first evaluation article, “The Logic of Evaluation.” From there, Scriven went on to instruct at several national and international universities, including Swarthmore College, Indiana University, University of California at Berkley, University of San Francisco, University of Western Australia, Claremont Graduate University, Auckland University, and Western Michigan University. During his tenure at these universities, he published numerous articles and books, which helped shape the field of evaluation. As outlined in the interview, Scriven’s thinking and writing about evaluation has been influenced by a number of individuals within (e.g., Daniel Stufflebeam, Malcolm Provus) and outside (e.g., Theodore Sturgeon, A. E. von Vogt) the discipline. Scriven continues to be a dominant presence in evaluation theory and practice.

Method Notes

In “Concept Mapping as a Technique for Program Theory Development: An Illustration Using Family Support Programs” Scott Rosas argues that concept mapping can be utilized to conceptualize a program’s implicit theory in theory-driven evaluations. Rosas defines concept mapping as a “multi-step process that helps articulate and delineate concepts and their interrelationships through group processes (brainstorming, sorting, rating), multivariate statistical analyses (multidimensional scaling, hierarchical cluster analysis), and group interpretation of conceptual maps produced.” Through his illustration, Rosas asserts that concept mapping improves design sensitivity, program conceptualization, and stakeholder-evaluator relations. Rosas concludes with a brief discussion of the limitations and issues surrounding this technique, which include the distinction between concept maps and program theories, the interaction between those involved and results, brainstorming prompt effects, and researcher statistical knowledge and software availability.

Book Reviews

The section begins with a review by Michael Hendricks of “Foundations and Evaluation: Contexts and Practice” edited by Marc Braverman, Norman Constantine, and Jana Kay Slater. The purpose of the book as purported by Hendricks is to promote effective evaluation practice within foundation evaluations. This book is divided into three sections: (a) perspectives from foundations; (b) understanding foundations as a context for evaluation; and (c) building capacity for evaluation practice. After providing a brief overview of the three sections,

Hendricks begins to outline the utility of this book. He recommends this book to foundation evaluators because it is “informative, sobering, and motivating.” With the assistance of eight colleagues, Hendricks concludes by highlighting book limitations (e.g., book length, limited audience applicability, overlap across chapters, and a lack of analytical consistency within chapters).

Next, Chris Coryn conducts a review of Huey-Tsyh Chen’s new book “Practical Program Evaluation: Assessing and Improving Planning, Implementation, and Effectiveness.” According to Coryn, the intention of the book is to clarify evaluation concepts, including those used in practice. The book is divided into four sections: (a) introduction; (b) program evaluation to help stakeholders plan intervention programs; (c) evaluating implementation; and (d) program monitoring and outcome evaluation. Coryn begins his critique of the book by pointing out several limitations. He argues that the book presents program theory as being static, misuses or misrepresents terminology, utilizes examples that are not congruent with current practice, and only briefly addresses some important evaluation issues (e.g., context, side-effects). Despite these limitations, Coryn recommends this book to both novice and seasoned evaluators for several reasons. Most notably, he believes the strength of this book lies in the presentation of Chen’s evaluation taxonomy, the inclusion of evaluation approaches that answer the “why” and “how” questions in evaluation, and an illustration of how to assess causal linkages in outcome evaluations.

This section concludes with Michael Patton and Michael Scriven separately critiquing “Empowerment Evaluation Principles in Practice” edited by David Fetterman and Abraham Wandersman. Patton says that the purpose of this book is to distinguish empowerment evaluation (EE) from other closely related forms (e.g.,

participatory, inclusive feminist, etc.). In doing so, the book highlights and describes the 10 empowerment evaluation principles (EEP) through the use of case examples. Patton argues that the book assists in clarifying EE from other similar approaches. However, Patton argues that the book fails to address other important components of the theory, such as the explicit relationship between the EEP and self-determination and the need for more research on EE to document the impact and outcomes of the method.

Scriven, taking a much different approach, evaluates the book in terms of its validity, credibility, and ethicality. Because J. Bradley Cousins offers a critique of why EE cannot maintain validity in summative evaluations in the book, Scriven begins his argument with examining whether empowerment evaluation can be valid in formative evaluations. He maintains that EE is not valid in formative evaluations because it fails to circumvent self-serving bias, to professionally filter method or content errors, and to prevent evaluator bias as a result of interaction with program personnel. Because it is not valid in the formative or summative arenas, Scriven further argues that EE cannot be credible. And, thus, if EE is not valid or credible, then it cannot be ethical. Scriven also briefly evaluates EE as an acceptable method in ascriptive evaluations, which he defines as a category that “seeks only to ascribe the appropriate degree of merit, worth, or significance to the evaluand, not to aid a decision maker or program developer with their special tasks.” Scriven contends EE is even less suitable for ascriptive evaluations because they are even more intolerant of the aforementioned validity threats. Scriven concludes with a short paragraph recommending this book to novice and seasoned evaluators because “it is full of good things” (i.e., case studies, suggestions, and lines of thought worth considering).

In Response

This section includes four responses to the aforementioned Patton and Scriven book critiques. The first two responses are offered by David Fetterman, Abraham Wandersman, and Jessica Snell-Johns. The last two are final thoughts offered by Patton and Scriven.

In his response, Fetterman thanks reviewers for their comments, corrects misstatements about the book, and proposes future discussion topics for this exchange. In “Appreciation and Agreements” Fetterman thanks and acknowledges Patton, Scriven, and Cousins for their impact on his EE thinking and practice. In “Misstatements about the Book” Fetterman argues that both Patton and Scriven misunderstood the purpose of the book and case examples. Fetterman maintains that the intent of the book is to present the 10 empowerment evaluation principles and how they operate in practice. And thus, it should only be evaluated against that criterion. Further, although Scriven disagrees, Fetterman contends that self-evaluation has been and will continue to be an important “time-honored role.” In terms of the case examples, Fetterman argues that the book provides solid evidence concerning the methods impacts and outcomes. Fetterman concludes by stating that the future of the field rests on evaluation research, scholarly exchange, and a movement toward common understanding.

Wandersman and Snells-John, take a similar approach to responding to Patton and Scriven. The purpose of their response is to identify areas of agreement, further identify and clarify misstatements, and discuss the future of EE. In “Areas of Agreement,” Wandersman and Snells-John agree that more evaluation research on EE is needed, EE is not appropriate in all contexts, and that the distinction between

EE and other theories needs to be more explicit. In “Areas for Clarification and Dialogue,” they clarify and discuss several concerns. First, they disagree with the value and definition of self-evaluation presented by Scriven. They believe EE ensures accurate information is gathered during self-evaluations and assists in developing a positive feedback and action cycle. Second, Wandersman and Snells-John argue that there is not one agreed upon opinion regarding evaluator-client relationships and that all relationships (or lack of) are subject to bias. Third, they address the appropriateness of EE in summative evaluations. In constructing their argument, they point to a case example in chapter four in which EE lead to a summative decision. Wandersman and Snells-John further argue that the role of the empowerment evaluator is ongoing, critical, and dynamic, and thus, it is not amateur evaluation. They conclude with a brief discussion concerning the next phase of EE, which they perceive to be more evaluation research on the approach.

This section ends with two separate responses by Patton and Scriven in which they reassert and support their previously outlined critiques.