

American Journal of Evaluation

<http://aje.sagepub.com>

The Use of Multiple Evaluation Approaches in Program Evaluation

Katrina L. Bledsoe and James A. Graham
American Journal of Evaluation 2005; 26; 302
DOI: 10.1177/1098214005278749

The online version of this article can be found at:
<http://aje.sagepub.com/cgi/content/abstract/26/3/302>

Published by:



<http://www.sagepublications.com>

On behalf of:

[American Evaluation Association](#)

Additional services and information for *American Journal of Evaluation* can be found at:

Email Alerts: <http://aje.sagepub.com/cgi/alerts>

Subscriptions: <http://aje.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations <http://aje.sagepub.com/cgi/content/refs/26/3/302>

Articles should deal with topics applicable to the broad field of program evaluation. Articles may focus on evaluation methods, theory, practice, or findings. In all cases, implications for practicing evaluators should be clearly identified. Examples of contributions include, but are not limited to, reviews of new developments in evaluation, descriptions of a current evaluation study, critical reviews of some area of evaluation practice, and presentations of important new techniques. Manuscripts should follow APA format for references and style. Length per se is not a criterion in evaluating submissions.

The Use of Multiple Evaluation Approaches in Program Evaluation

Katrina L. Bledsoe

James A. Graham

The College of New Jersey

Abstract: The authors discuss the use of multiple evaluation approaches in conducting program evaluations. Specifically, they illustrate four evaluation approaches (theory-driven, consumer-based, empowerment, and inclusive evaluation) and briefly discuss a fifth (use-focused evaluation) as a side effect of the use of the others. The authors also address the usefulness of a multimethod research design with these approaches, especially in developing responsive evaluations (to both communities and organizations). Finally, the authors address some of the benefits and challenges of working with the four different evaluative approaches in a community program, and they discuss the value of using multiple approaches in the evaluation field.

Keywords: *community-based programs; evaluation; approaches; theory-driven; empowerment; consumer-based; inclusive; multimethod*

The question of what evaluative approach works best under which circumstance has been a constant in evaluation literature. For instance, Stufflebeam (2001) attempted to answer this question by providing a taxonomy and ranking of different approaches to determine the most useful (and not so useful) ones. However, we believe that evaluators are more likely to use the

Katrina L. Bledsoe, The College of New Jersey, Department of Psychology, Box 7718, Ewing, NJ 08618; phone: (609) 771-2582; e-mail: bledsoe@tcnj.edu.

Authors' Note: We would like to thank Barbara J. Lehman of the Department of Psychology at the University of California, Los Angeles; David P. Nalbone of the Department of Behavioral Sciences at Purdue University, Calumet; E. Jane Davidson of Davidson Consulting; Randy Rhoades, the editor; and the anonymous reviewers for their thoughtful comments. Additionally, we are especially grateful to Dolores Bryant, Charles Gershenson, and Donna Pressma for their support of this project.

American Journal of Evaluation, Vol. 26 No. 3, September 2005 302-319

DOI: 10.1177/1098214005278749

© 2005 American Evaluation Association

elements of multiple evaluation approaches (whether they are aware of it or not) when conducting evaluations. That is, evaluators are likely to incorporate components of several approaches to enhance the viability and fidelity of program evaluations. Such incorporation of multiple evaluation approaches can help in better understanding the needs of stakeholders and program recipients, producing better method designs, and yielding more accurate recommendations by which to enhance program development and change.

The use of multiple evaluation approaches is not a new phenomenon, of course; it is considered commonplace in practice (Chavis, 2004). Scriven (1997) advocated conducting evaluations that use a variety of components of different approaches. However, although the use of multiple approaches has received much attention in practice, it has received little attention in the formal literature, perhaps because the discussion and informal use is so commonplace (Chavis, 2004).

The aim of the current article is to discuss a program evaluation that used the components of multiple evaluation approaches to examine a community-based family literacy program servicing parents and young children aged 0 to 6 years. In identifying the evaluation approaches we used (and by extension the specific components of those approaches), we hope to help evaluators understand when it is best to use a certain approach. Our focus is on the use of multiple evaluation approaches, that is, the use of several theories of perspectives on evaluation and their accompanying components (e.g., empowerment evaluation, the development of organization self-determination; theory-driven evaluation, the development of program theory) rather than the use of multiple methods (i.e., qualitative and quantitative designs).

Examples of a multimethod approach are abundant in the evaluation literature; for instance, the use of multiple methods has been identified as helpful not only in working with consumers but also in informing program development (e.g., Michalski, Mishna, Worthington, & Cummings, 2003; Towns, Cole-Henderson, & Serpell, 2001). However, discussions of the use of multiple evaluation approaches (which often include multiple methods) in conducting program evaluation are far fewer. Christie and Alkin (2003) demonstrated that using a theory-driven framework within a use-focused evaluation further increased the likelihood of use in an outreach program. Similarly, Shaw and Replogle (1996) suggested that a comprehensive evaluation framework that considered multiple approaches was needed to better evaluate school-linked services for literacy programs. The work we present uses multiple evaluation approaches (namely, specific components) and methodologies with a community-based social services organization that is implementing a community-based family literacy program in an ethnically diverse, low-income area in Trenton, New Jersey. Specifically, we focus on how using several evaluation approaches, and by extension a variety of methodologies, led to an evaluation that was considered by the organization as having the following components: (a) scientific credibility (i.e., we made use of basic social science methodology, and our concepts and methodologies were grounded in scholarly literature), (b) validity (we accurately portrayed the program and community), and (c) utility to the organization (the organization agreed to implement most of the recommended program suggestions).

We recognize that there are several components of each evaluation approach that are considered cornerstones to that particular approach. However, the focus of the current discussion is on the specific components we used from each approach to conduct a viable and feasible program evaluation of the literacy program.

Aims and Structure of the Article

As a backdrop to the issue of using multiple approaches in program evaluation, we illustrate four evaluation approaches (i.e., theory-driven, consumer-oriented, empowerment, and inclusive evaluation) and briefly discuss a fifth (use-focused evaluation), which we believe was

achieved because of the use of the others. We begin with a general background of the purpose of the evaluation and the program logic used by both the program developers and the evaluators. Next, we explore our use of several key components from each of the approaches and the manner in which they were used. In that context, we also discuss some of the types of measures used and resulting data. We also address some of the challenges and limitations of using the approaches. Finally, we report some of the programmatic recommendations we made as a result of using multiple approaches and discuss the implications for evaluation practitioners and theorists.

Background of an Urban, Community-Based Family Literacy Program: Fun With Books

Purpose of the Evaluation

The evaluation of the Fun With Books (FWB) program in Trenton, New Jersey, began at the request of the organization's CEO and the director of the program. Both stakeholders had noted that after 4 years of operation, the program appeared to be one the most successful ventures of the 40 programs offered by the organization. However, no formal evaluation had been conducted to confirm the achievement of the program's purported goals. The organization's primary focus was on the welfare and well-being of children, their families, and by extension the organization's commitment to the No Child Left Behind (NCLB) Act, a federal government law designed to ensure the academic success of children in classroom settings. The organization's interest was not only to provide services to an ethnically and socioeconomically urban community but also to make use of the literature from the fields of early childhood education and developmental psychology and to help support its belief in NCLB (CHS Organization, 2002).

Description of the FWB Program

FWB is an interactive family literacy program that uses children's literature and music to support the development of preliteracy and school readiness skills in young children. FWB has been in existence for over 4 years, as part of a larger umbrella program called Family and Children Early Education Services (FACES). The FACES program focused on four distinct but interrelated agendas: health and behavioral health, parent education, child care, and family literacy.

The health and behavioral program focuses on providing education and access to health services; the parent education program articulates the provision of parenting skills; the child care component is designed to provide reasonable and quality child care to families who need it; finally, the family literacy portion encourages in-home reading between parents and children and promotes school readiness. This latter piece is considered the signature part of the program, especially because this was considered by the organization to be the most scientifically based of the four components (CHS Organization, 2002).

The 6-week program was offered to families in Trenton and surrounding communities who had at least one child between the ages of birth and 6 years. Older siblings who attended also participated in the program. The reasoning for this diversity in age was the developers' desire to further support their belief that parent-child bonding provides the opportunity to encourage cognitive stimulation through reading and interactive activities into the school-age years.

Program activities consisted of parents and staff members reviewing the book of the week and exploring the book's theme. Additionally, structured art, craft, and music activities, in

which parents engaged with their young children, were included. The activity and music portions of the program were designed to highlight the theme of the book chosen for the week. At the end of each session, parents were encouraged to read the story with their children in the home environment. Such activity has been demonstrated not only to increase the likelihood of parent-child bonding but also to foster school preparedness and readiness (e.g., Morisset, 1997). With every week's attendance, each family was given the book to add to their home library. Each week, the program focused on various themes that directly related to an area of literacy development (e.g., colors, food, sounds, shapes, numbers, and animals). The final session of the program served to connect all of the themes discussed throughout the 6-week session.

Program goals were focused on promoting literacy and teaching valuable parenting skills. The program sought to encourage parents to read to their children at home throughout the week, to enhance parent-child interaction, and to provide activities that support cognitive development and preliteracy in young children.

The conceptual logic model in Figure 1 illustrates the program developer's beliefs that the literacy program should support the development of preliteracy through cognitive stimulation (e.g., reading to a child actively stimulates the brain), foster nurturing interactions between parents and children (e.g., encouraging positive physical interactions, which also serve to stimulate the brain), and reduce parents' stress levels (thereby allowing them the capacity to gain better parenting and reading skills). By accomplishing these frontline goals, the program seeks to increase the capacity, motivation, and opportunity for children to learn. The strategies lead to the short-term goal of promoting in-home reading between parents and children. It is believed that in-home reading leads to the long-term goal of school readiness.

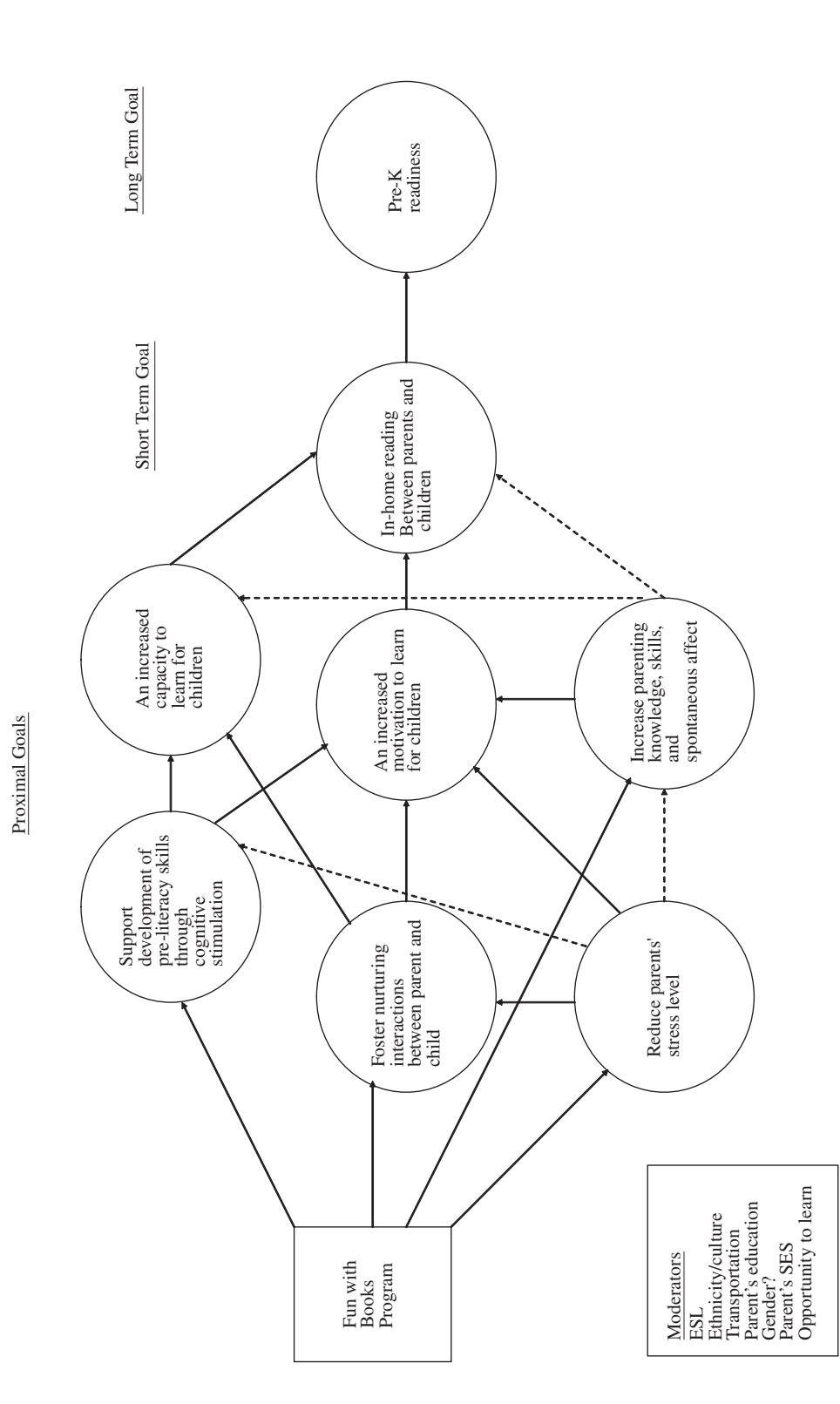
Context of the Evaluation

Although no formal evaluation work had been completed in FWB's initial 4 years of operation, some process data, such as number of consumers attending the program and population demographics, had been documented through the overarching FACES program. However, those data described consumers served rather than effects on consumers. The challenge was to generate data that could yield both process and outcome information. An additional challenge was to accurately measure concepts from the conceptual model that did not appear to be easily quantified. For example, when asked, many of the stakeholders were at a loss as to illustrate the concept of "cognitive stimulation." Yet because this concept was supported in developmental literature, the desire to measure it was strong.

After discussing the program with the FACES director, and keeping in mind the designer's desire to make use of the developmental psychology and educational literature in the program, the evaluation team worked closely with stakeholders to develop the logic model (Figure 1), evaluation work plan, and research and evaluation questions.

The stakeholders' belief that offering activities that promoted cognitive stimulation (e.g., brain activity) and nurturing interactions (e.g., hugging and kissing) between parents and children caused an increase in in-home reading between parents and children was not conducive to measurement; we were unsure of how this occurred and, if it did occur, how to accurately measure it given the limited context (money and community setting). Thus, the focus of the evaluation was on identifying measurable developmental variables related to positive family functioning, fostering in-home reading, prekindergarten school readiness, and overall program effectiveness. Because both the program director and the CEO used science-based programs as a benchmark for FWB, this also meant that the organization felt that it would need science-based evidence for its program. However, we fully understood that the community context in which the organization and program operated (e.g., highly diverse, low socioeconomic status)

Figure 1
Program Model



was not infrastructurally sound to support traditional methodological approaches such as a true experimental design; the evaluation required methodological, conceptual, and theoretical flexibility. Therefore, the evaluation team chose to use the components of multiple evaluation approaches that would allow such flexibility.

Evaluation Approaches Used

We chose approaches that would be responsive to the needs of the community, consumers served by the organization, and the organization itself. The goals of the program were based on the perception that the intervention strived to be a model program for family literacy in New Jersey and the United States and would be representative of the NCLB Act. The agenda was one in which the evaluation team and stakeholders recognized the benefits to both applied and scholarly work. A multifaceted approach to evaluation would allow the disciplines of psychology and education to benefit from the findings of this project while providing the Trenton community with valuable information.

The evaluation team decided to use approaches consisting of the empowerment, theory-driven, consumer-oriented, and inclusive evaluation approaches. Each was chosen taking into consideration the diverse context of the organization and the consumers it served, as well as the socioeconomic status of the community members and the community itself. A fifth approach, use-focused evaluation (Patton, 1997), was identified in postevaluation activities, and the evaluation team later identified this as a side effect (Scriven, 1991) of the combined use of the other approaches (as discussed later in this article).

Empowerment Evaluation

Empowerment evaluation, a form of participatory evaluation (Fetterman, 1996), incorporates the perspectives of all those with vested interest in a program (stakeholders), usually those who are invested in program development and the distribution of services. In short, empowerment evaluation allows for increased involvement of program stakeholders and allows them to identify and define those needs (and in what manner to evaluate those needs) that are most important to a program.

Theory-Driven Evaluation

The second approach used was theory-driven evaluation (see, e.g., Chen, 1990; Weiss, 1997). Although there are many variations of the definition of theory-driven evaluation (e.g., Gargani, 2003), we defined theory-driven evaluation as using a synthesis of both stakeholder program logic and social science theory to define what a program does, in what manner, and how much of an effect each goal and objective can have on the outcome. In our design, the best-case scenario for theory-driven evaluation was to enable the evaluator to ascertain actual causal mechanisms (or at the very least, predictors) of the program strategies and link those to changes in program participants.

Table 1
Approaches and Components Used

Approach	General Purpose	Components Used
Empowerment	<ul style="list-style-type: none"> • To provide a sense of ownership in program design, methodology, and evaluation 	<ul style="list-style-type: none"> • Goal setting • Strategy development
Theory driven	<ul style="list-style-type: none"> • To provide a science-based foundation for the program to rest • To provide feasible measures • To help the organization formally articulate its own program theory 	<ul style="list-style-type: none"> • Conceptual/logic modeling • Development of program theory
Consumer based	<ul style="list-style-type: none"> • To encourage stakeholders to focus on the needs of consumers, as well as the community • To encourage the consideration of the wide range of needs 	<ul style="list-style-type: none"> • Searches of literature/files on organization and consumers • Exploration of consumers served and informal discussion of consumers' needs of the organization
Inclusive	<ul style="list-style-type: none"> • To encourage the organization to consider the dynamic and diverse cultural context of the Trenton community • To encourage diverse perspectives on parenting and education • To encourage consumer-based programming 	<ul style="list-style-type: none"> • Using the context of the community as the main foundation in which the program exists • Asking those from community settings to be active partners at the design and evaluation stages • Understanding the reality in which consumers live and interpreting through those lenses

Consumer-Based Evaluation

We also used a consumer-based approach (Scriven, as cited in Fitzpatrick, Sanders, & Worthen, 2004), concerned with incorporating and designing evaluation procedures and questions, with the needs and opinions of those would-be recipients of program services in mind. In this case, we wanted to ensure that the FWB program participants would be used in designing the methodology and in providing critical information as to how well the program was accomplishing its intended goals and objectives.

Inclusive Evaluation

We felt that an inclusive approach (Mertens, 2003) was also a major perspective used. Inclusive evaluation is meant not only to include stakeholders who have traditionally been recognized, such as funders, administrators, staff members, and participants, but also to seek out accurate and credible representation of members of groups who have been traditionally excluded from or misrepresented in the evaluation process (Mertens, 2003). In the case of FWB, we wanted to design and conduct an evaluation sensitive to the needs and cultural nuances of the community it serves, predominately a community of color.

The evaluation team approached the evaluation with the understanding that each evaluation approach has its distinct foundations and strategies. As in a Venn diagram, we acknowledged that approaches have parts of mutual exclusivity yet do coincide. In some cases, using components of one approach could not occur without the use of another. For instance, we could not engage in logic modeling without using a self-determining strategy. As shown in Table 1, we decided what components of each approach we wished to use.

In the following section, we provide a discussion of the strategies used from each approach and the types of data and results that were generated.

Empowerment Approach

The empowerment evaluation approach involved the inclusion of organizational stakeholders in the development of long-term, user-friendly program evaluation systems. We felt strongly that we should leave the organization with the ability to design and manage its own program-monitoring evaluation. We felt that the strategy of goal setting (in this case goal identity) would help the organization have some sense of self-determination in the program evaluation. Such self-determination allowed the organization to make its own discoveries about the program, its needs, and its consumers (Fetterman, 1996). We met regularly with stakeholders and attended key events hosted by the organization. We felt that this process was helpful; from our vantage point, we were able to gain information on the community context, evaluation data, strengths, limitations, and next steps for program improvement and to share this feedback with stakeholders in weekly meetings as they redesigned their organization's programs. Additionally, in trying to encourage a self-determinative and self-evaluative attitude, we suggested that the organization participate in interviewing staff members to capture qualitatively the context under which the program was functioning.

The evaluation team worked with the stakeholders to design 20 semistructured questions to solicit staff members' opinions concerning the program's goals and participants. The protocol was adapted from interview questions from the Head Start FACES project (Family and Children Early Education Services, 1997a, 1997b) and was designed to help the evaluation team gain an understanding of staff members' education and training, values and attitudes, and general program buy-in. Additionally, the questions served to extract information to help us understand the organizational structure under which the staff operated. We found this especially useful in helping the organization determine the realistic expectations of what it could accomplish given its staffing and monetary resources.

Theory-Driven Approach

Prior to the evaluation team's arrival, the organization had employed a data analyst and manager to provide basic descriptive statistical analysis of the number and percentage of consumers served for each program and, to a significantly lesser degree, consumer satisfaction with services provided. Thus, the organization recognized the need to be able to establish and own its feedback mechanism (Donaldson & Gooler, 2002). However, there was little organizational knowledge of how to institute this feedback mechanism, so the development of effective but understandable process and outcome measurement instruments was of high priority. Additionally, we focused on trying to develop a framework for data collection that the organization and program could restructure as needed. We met weekly to establish realistic measures (those that consumers would be able to both understand and relate to as well as measures that would be easy to consistently distribute and analyze) and to develop a database of information that the organization could control and maintain after the conclusion of the external evaluation.

The use of program concept models was also helpful. One of the driving forces behind the FWB program was the designers' desire to rely on research on literacy, early childhood education, and child development. Thus, a major concern for the program was establishing a clear rationale for why the program should exist and what techniques and strategies were being used to accomplish the goals and objectives. The evaluation team used a theory-driven approach because we wanted to synthesize theories of the designers with theories from the social sciences. The designers believed that FWB was based on scientific theories of family literacy programming and developmental psychology; we responded to this belief by attempting to model many of the measures and data analyses in the literature from these fields.

For instance, researchers of literacy and literacy programs have speculated that the reasons why many of the nation's children are not adequately prepared for school success are quite complex (e.g., Gadsden, Brooks, & Jackson, 1999). A major contributing factor may be the social and economic difficulties faced by parents, such as disrupted family structures and low socioeconomic status. We also considered the use of literature on emergent literacy (precursory knowledge acquired prior to conventional literacy instruction) and that documenting successful literacy programs. Using this literature was vital in helping us choose questions that would give consumers the opportunity to provide vital and useful information with which to measure the program.

We used established, reliable measures documented by the Head Start FACES program. Stakeholder and team meetings produced several drafts of (a) the program logic and (b) the scientific underpinnings of that logic. Thus, the evaluation team was able to develop measures to assess variables such as literacy, parenting skills, and parenting stress. Additionally, we worked to obtain behavioral data through the use of observational measures (e.g., the Adult/Child Interactive Reading Inventory; DeBruin-Parecki, 1999) to corroborate self-report information provided by program participants. Sample measures included but were not limited to children's interactions with books (e.g., how many times a child held a book and pretended to read) and parents (e.g., how many times a parent made physical contact with a child in a positive manner).

In consciously undertaking a theory-driven approach, we wanted to encourage the organization to specifically tailor its questions and hence the evaluation itself to measure actual participant change and program outcomes.

Consumer-Based Approach

A consumer-based evaluation approach is focused on designing and incorporating evaluation questions and procedures with the needs and opinions of those consumers who are receiving services and by extension would be most able to determine how well a program is meeting those needs (Scriven, 1991). The organization's stakeholders provided information on the types of questions that were most likely to be relevant to the community in which consumers resided. In some cases, the organization actually solicited the participation of past program recipients in initial discussions. In addition, several staff members (some of whom had been consumers of services through the organization) also participated in the program and provided needed feedback on survey instruments.

We also addressed community literacy and language barriers by developing measures that could easily be read by the program consumers themselves or by the evaluation team. For instance, we used some of our initial surveys as pilot data to help indicate the types of questions consumers would find understandable. These pilot data served as the foundation on which other evaluative measures were based and provided community members with a voice in the evaluative process, albeit indirectly.

Inclusive Approach

The evaluation team spent some time learning about the community in which the program is administered. The Trenton area is diverse (approximately 52% Black and 32% Latino/Hispanic) and is a lower-middle-class-to-poor area. Almost 25% of the economic resources of the city are devoted to the social services industry (U.S. Census Bureau, 2000). The area also has a large immigrant population from areas such as Latin America (e.g., Columbia), the Caribbean (e.g., Jamaica), and Eastern Europe (e.g., Poland). The evaluation team recognized the

diversity not only in socioeconomic status and education but also in cultural and parenting practices. For example, we worked to help program staff members understand that parents from diverse communities, specifically from international communities, may have different perspectives on parenting. Finally, the use of an inclusive approach encouraged the development and use of qualitative measures, such as interviews, to provide rich data about the organizational and community context.

We also conducted earnest discussions that focused on race, ethnicity, immigrant status, and socioeconomic status. Although the organization fully acknowledged that it catered to a diverse community, it was less aware of how that diversity would influence the types of experiences consumers had in the program and the types of feedback generated. Our use of an inclusive approach continually kept the aforementioned conversations at the fore, and organization members actively sought to consider culture in final evaluative discussions, reports, and future programs.

The Use of Multiple Evaluation Approaches in Developing Programmatic Recommendations

Programmatic recommendations were specifically geared toward changes, fine tuning, and program development. All approaches encourage the use of formative evaluation from different perspectives.

Identify Program Theory

We found the theory-driven approach most useful in discussions of scientifically based strategies. Specifically, in using logic models, we were able to clearly pinpoint, at least in a prescriptive manner, what the program purports to do. Donaldson and Gooler (2002) noted that this very act can sometimes serve to provide needed information, even in the absence of concrete data. In this case, we worked to identify mediating factors (i.e., seeking to support development of preliteracy through cognitive stimulation, fostering nurturing interactions between parents and children, and reducing parents' stress levels) that were presumed to lead to or predict in-home family reading. In addition, we were able to identify those factors that could affect the strength and magnitude of the program (moderators), such as parents' education levels and the ability to transport family members to the program.

The identification of these mediators and moderators can be helpful in pinpointing the exact paths along which a program works (Donaldson, 2001). For instance, one of the major discoveries that came out of the evaluation was that in many ways, the organization was still in a formative stage. Although the program had been in operation for almost 4 years prior to the evaluation, staff members and consumers were still unable to articulate the true purpose of the program. Additionally, in using program theory, we were able to help the program designers understand the difference between causality (does the program cause parents to engage in cognitive stimulation, which then causes an increased capacity to learn, which then causes in-home reading?) and correlation (is cognitive stimulation related to a capacity to learn that in turn is related to in-home reading?).

Increase Organizational Learning

An empowerment approach was useful in helping the organization come to its own realization (rather than the evaluation team's) concerning the lack of understanding of the program's

goals and objectives. For instance, in trying to redefine the program and prepare for future evaluation, program stakeholders began to publicly acknowledge that the staff as a whole was indecisive about whether the program was designed for families to “have a good time and relax” or for developing “prekindergarten readiness in children.” Through qualitative interviews, organizational leaders also discovered that the program changes each time it is administered, which likely results in different outcomes with each administration.

Increase Cultural Sensitivity

An inclusive approach encouraged the high-level stakeholders (i.e., the CEO and program director) to include the perspectives of those who administered or used the program. Although the organization had sometimes hired some of the consumers of its services in lower level administrative positions within the organization, current and previous consumers were not always asked to provide meaningful feedback on the program.

A discovery about the program volunteers led to a recommendation to provide volunteers with more in-depth information concerning the program and its goals. Because of the active and sometimes physical nature of the program, the organization had been training volunteers to work with participants on such aspects as how to deal with distracted children. However, overall, volunteers had even less knowledge about the program’s goals and the organizational community context than the staff members and consumers. For example, there was a wide disparity in income between the volunteers and participants (volunteers were recruited from a financial investment company in an upper-middle-class township, whereas participants were from lower-middle-class to working-class townships). As a result, there was little understanding as to what barriers the families faced prior to arriving to the program (e.g., a lack of finances, transportation issues).

Cultural and socioeconomic status differences also caused a divide. The participants were primarily Black and Latino and lower class, whereas the volunteers were White and upper middle class. These cultural and socioeconomic status differences often extended into the volunteers’ preconceived ideas of what was considered appropriate parenting and what was not. Using an inclusive approach, which focused on bringing to light some of the participants’ cultural values that were inherently different from those of the program staff members, the evaluation team was able to help the organization and staff members understand how perceptions based on cultural differences could affect the implementation of the program.

Extend the Length of the Program

Another recommendation was that the organization extend the length of the administration of the program to consumers. At the time this article was written, FWB was a 6-week program in which attendees participated once per week. Both staff members and consumers indicated through interviews and anecdotal responses that the program should be longer. Psychological research has established that it takes approximately 6 weeks for a change in behavior to become habituated (e.g., Sarafino, 2002). Specifically, just when the “habit” of reading became implanted in the family structure, and staff members and families had bonded (in many cases, staff members reported that informal social work was performed during the sessions), FWB ended. The evaluation team recommended that the program be extended to 8 or 10 weeks. The extension would not only assist in sustaining the reading habit but also allow for constructive interaction, such as dealing with a family’s economic problems, between FWB attendees and the staff members and volunteers.

Formally Test Variables

We also recommended that FWB become more formalized in testing and identifying its consumers. We suggested the use of school readiness measures. Stakeholders purported that the program ultimately aids in preliteracy and school readiness, and the evaluation team felt it important that those assumptions be grounded in concrete scientific evidence, such as linkages with test scores from psychological and school inventories. Specifically, we suggested that the literature on emergent literacy might be helpful. Emergent literacy is dependent on the social and cultural contexts in which children interact. Vygotsky (1986) posited that positive social interactions with an adult within the cultural contexts of society are fundamental to healthy social, cognitive, and emotional development. According to Rogoff (1990), skill development requires the interaction of two parties, a teacher (typically an adult) and a learner (typically a child), which is an “apprenticeship-type relationship” that involves “guided participation” during daily experiences. Children learn via daily interactions with their parents and other adults during play, teaching, and other routine situations. We felt that this literature would help the organization achieve its abstract goals in a scientific manner. Other program evaluations of literacy programs had sought to do this; we found evidence that the Even Start Literacy Program (Nistler & Maiers, 2000) and other similar programs have the potential to reduce dropout rates. Moreover, it was found that a joint home-school family literacy program was established to help parents understand the important role that they play in their children’s literacy growth.

Identify the Target Population

Finally, the organization’s program staff members and FWB’s volunteers as well as the evaluation team felt that the organization’s identification of its target population was inaccurate. FWB serviced children aged 0 to 6 years. Through the use of a theory-driven approach, the program staff members came to understand why children of different age groups exhibited different behaviors. For instance, children at different ages respond differently to reading as a result of differences in comprehension. Traditional developmental literature has established that children between the ages of 0 and 6 years are at dramatically different stages of cognitive understanding (e.g., Piaget & Erickson, cited in Myers, 2002). Thus, the program has different meanings for different consumers (i.e., from ages 0 to 3 vs. 3 to 6 years) and therefore different outcomes for various children.

The Use of Multiple Evaluation Approaches in Determining Recommendations for Future Evaluations

Our experience using multiple evaluation approaches indicated that this can also be helpful in determining recommendations for future evaluations. At times, evaluators may feel as if their use of a particular approach is limited by the methodology used (e.g., quantitative or qualitative methods). However, we believe that using multiple evaluation approaches encouraged the development of recommendations for future evaluation and the types of methodologies that would be most useful in outcome evaluation. The recommendations we discuss here are those that specifically encouraged establishing some kind of learning feedback mechanism and/or provided an opportunity to make an evaluative conclusion (e.g., was the goal accomplished or not?).

Include All Stakeholders

The evaluation team proposed that the FWB program consider supporting an ongoing evaluation program to provide a mechanism by which it would receive continuous program feedback. This mechanism could be used to assist the FWB team to continually monitor the effectiveness of the program, as well as to anticipate and respond to the changing needs of the community it serves. We suggested that all stakeholders be involved in the planning and evaluation processes, specifically past attendees of the program.

Develop an Evaluation Team

One of the priorities of the evaluation team was to help the organization foster a sense of ownership in the FWB evaluation, and the use of empowerment evaluation was helpful with this. Such ownership provides a strong foundation for organizational learning and for continuous quality improvement (e.g., Davidson, 2003). We recommended that the organization assemble a formal evaluation team that would be primarily responsible for developing and updating strategies for program improvement as well as program evaluation. This team would be responsible for conducting process and outcome evaluation on a yearly basis to provide both qualitative and quantitative data to longitudinally track changes and trends in the program and the community. Being able to track changes and trends would allow the program administrators and staff to continually monitor the consumers it serves and the cost-effectiveness of strategies and to provide realistic process and outcome measures to the organization's board and potential funders. Such collaboration serves to empower those who both evaluate and administer the program (Fetterman, 1996).

Benchmark Similar Programs

We also suggested that the organization benchmark other family literacy programs that might help it maintain and improve the FWB program. The theory-driven approach was useful; we worked with the organization to help define the program's strategies, goals, and objectives. Through this process, we made use of scholarly literature as well as active programs in the other regional areas. We believed that this approach would allow for the measurement of the program in terms of effectiveness, cost-effectiveness, and efficiency.

The Use of Multiple Evaluation Approaches in Raising Questions

We found that using multiple evaluation approaches served to identify side effects of the program, which we identify and discuss here, as questions for consideration. In using inclusive, empowerment, consumer-oriented, and theory-driven perspectives, we found that the organization was amenable to conducting reflective self-evaluation and engaging in honest organizational learning. Specifically, the overarching discovery was that the organization realized that its level of communication needed to be improved between administration, staff members, volunteers, and consumers. This self-discovery led to meaningful and candid discussions after data collection and analysis, with the organization generating a list of questions that program administrators and staff members needed to address.

Program Dosage

One of the issues that arose in postevaluative discussions was that of program implementation (or program dosage). The program director acknowledged that the program had been administered in several forms over the past 4 years, but the staff members were unsure which strategies worked best at achieving the overarching goals. For example, was the focus on providing literacy skills or parenting skills or on addressing social work issues, or all three? Could each strategy be administered effectively without compromising the others? Using a theory-driven approach led the organization to embrace a more scientific perspective in determining the effectiveness and perhaps causality of the program.

Lipsey (1990) has often discussed the importance of design sensitivity and its relationship to program dosage. In the case of FWB, a pertinent question is how often families should be engaged in the program. The organization acknowledged that there was no limitation to how many times a family could attend FWB. Thus, the organization wondered if the program provided diminishing returns; perhaps some strategies were useful only up to a certain point. Thus, a resulting question was, "Should the program be prepared to address changes in parents' skill sets for reading with their children?"

Program Effects on Participating Fathers

Another side effect discovered was the effect of the program on the participating fathers. Through an inclusive yet consumer-oriented evaluative approach, the organization found that fathers were active participants in the program. This was inconsistent with a significant portion of the developmental research, which stresses the bond between mothers and children (e.g., Goodman, 2002). The question was then, "What can the program do to encourage paternal participation in what has been informally defined as 'maternal territory'?"

Cultural Competency of Volunteers

The organization noted that all program staff members had backgrounds in family services, but volunteers had significantly more variability in experience, with many having no background or experience with family services or the population of people who received those services. The inclusive approach helped the organization place emphasis on understanding the cultural and social context of the participants and the volunteers. Many of the volunteers who worked with the FWB program were unlikely to have worked previously with the populations they encountered. We suggested that cultural competency training might be helpful to volunteers and staff members. Such training might be helpful in providing awareness to the unique perspectives of the consumers and in dispelling stereotypes that could alter the impact of the program on the families.

Benefits of the Program

Finally, the organization began to seriously question the true benefits of the program. Although the data the evaluation team collected seemingly provided empirical evidence that the program was effective in the short term, the organization was still left with the unanswered question of whether involvement in the program causes in-home reading and by extension prekindergarten readiness. Chen and Rossi (1987), Lipsey (1990), and more recently Davidson (2000) have described the precarious nature of ascertaining causality in program evaluation.

Thus, because of the lack of ability to establish strong causal links and the inability of both the organization and the evaluation team to measure them, it was difficult to say whether the purported long-term goals were accomplished.

Challenges to Using Multiple Evaluation Approaches

We have presented a strong case as to why the use of multiple approaches can be useful in program evaluation. However, using such a strategy is not without its challenges. One challenge is to be able to know which component of an approach will work best with others. This is not always knowable in advance, and sometimes, strategies are in direct conflict with one another. For instance, although empowerment strategies provide stakeholders and consumers with mechanisms for self-determination, many stakeholders and consumers do not have the background or organizational resources to conduct sound program evaluations that will truly determine the merit, worth, and significance of a program (Scriven, 1991).

An additional challenge is that of having lay stakeholders understand well enough the evaluation approaches that can be used, so that they can participate in choosing the components that should be used to conduct the evaluation. This can prove problematic if an evaluator makes decisions that may be best for the evaluation plan but are then vetoed by stakeholders with limited knowledge. For example, when using a theory-driven approach, we still encountered difficulties. Although stakeholders had a reasonable understanding of the literacy literature, they had limited understanding of how to accurately identify and measure the purported program components. Thus, it was difficult to persuade stakeholders that the use of experimental or quasi-experimental designs was not feasible in the current evaluation context. Occasionally, we were limited by resources (e.g., access to funding); mostly, we were limited by lack of understanding.

Although we discussed our preferred approaches with stakeholders early in the evaluation process, stakeholder understanding of this early discussion was, not surprisingly, limited. As a result, implementing specific components of each evaluation approach was somewhat difficult. For instance, although stakeholders understood the use of conceptual models, they were less perceptive in understanding that each component should be measurable; for example, the concept of "capacity to learn" is quite abstract. This led, in some cases, to compromised data collection (e.g., settling for small numbers, with missing data on key variables), especially data that might support a causal relationship between aspects such as cognitive stimulation and prekindergarten readiness.

Our use of empowerment strategies was often at odds with more empirically oriented strategies, such as theory-driven evaluation. For instance, although we collaborated on the development of surveys and observational measures, stakeholders did not understand validity and reliability and could not differentiate between satisfaction, process, and outcome measures. Our use of the inclusive approach was also not without its problems. For example, although the organization embraced a diversity-focused mission and spoke a lot about diversity, it often did not acknowledge the value of cultural and socioeconomic contexts. This topic sometimes raised defensiveness and resistance. When the evaluation team suggested that cultural divides might be the cause of some of the misconceptions of the volunteers or program staff members, some upper level stakeholders minimized the issue by saying that the group was bonded by "love." We also found it difficult to execute all the strategies in an unbiased and objective manner. For instance, our understanding of the program sometimes biased the manner in which we chose to analyze results. In some cases, stakeholders, and at times the evaluation team, became so attached to the goals identified during the evaluation that it was not until the postevaluative discussion that we were able to clearly identify side effects that could have been measured in the

course of the evaluation (e.g., the effect of the program on the participating fathers). Scriven (1991) noted that such collaboration between evaluators and stakeholders can bias the evaluation, evaluative judgments, and recommendations.

Finally, although the value of using multiple evaluation approaches is extolled by evaluators, not all are adequately trained in correctly using the approaches. This can be problematic if evaluators incorrectly use approaches, thus compromising the fidelity of the evaluation and possibly providing invalid results. Despite these challenges, we felt that our active use of strategies from multiple approaches was generally successful. In particular, we believed that an unexpected side effect of the use of the four approaches was a fifth one, that of use-focused evaluation (Patton, 1997). On the basis of the recommendations and questions for consideration, program modifications were made to staff and volunteer training, the curriculum, cultural contextual issues, administration, and program quality assurance. Program staff members agreed to implement changes that had been presented to them as feasible and reasonably easy to implement. Thus, the use of the information and recommendations provided was perhaps the most successful venture of the evaluation itself.

Implications

Evaluation Practice

The earlier question of what evaluative approach works best under which circumstance should also be closely linked with the issue questions, "What is the likelihood that the use of multiple evaluation approaches will increase use?" and "How would the findings be received, and in what manner would they be used, if any?" Multiple evaluation approaches helped inspire the kind of use that the evaluation team witnessed. An empowerment approach allowed the organization to take ownership of its own learning, while the inclusive and consumer-oriented approaches encouraged the organization to seriously consider the program's effectiveness in terms of the community's actual context rather than from the organization's perceived community context (e.g., a lack of ability to read, poor parenting skills due to low income status). Finally, a theory-driven approach encouraged the organization to use information, literature, and methodology that would enable the discernment of causality (or at the very least predictability) and provide science-based conclusions in accordance with the NCLB Act.

Practitioners and Theorists

The primary impetus for the current article is the realization that use of evaluation approaches is not neatly packaged into one particular approach. Much has been written about the theory-driven approach, the empowerment approach, and the inclusive approach, among others. Yet it seems rare that only one specific approach is used (Christie, 2003). Although we acknowledge that one approach might take the lead, it is more likely that several approaches may be at work, used at opportune times to generate specific outcomes (Christie, 2003).

By delineating the use of collaborative approaches in program evaluation, we were able to generate more comprehensive findings and by extension a more comprehensive program evaluation. Specifically, we were able to gather the in-depth information needed while helping increase organizational learning capacity and at the same time maintaining scientific objectivity, independence, and external evaluator credibility.

Perhaps it is best to view this from a critical multiplist perspective (Cook, 1985). Each approach has strengths and limitations, and using multiple approaches can serve to complement

the evaluation and enhance the information generated. Ultimately, we see this as a beginning dialogue to help establish a more collaborative relationship between evaluation approaches (see SenGupta, 2004).

The current article serves as an exemplar of how the use of multiple approaches in a program evaluation can generate a more comprehensive evaluation, generate informed recommendations, and increase the use of those recommendations for underserved communities. We also hope theorists will further reflect on the informal use of multiple approaches in evaluation. This may help minimize, if not resolve, the ongoing debate over the utility of different evaluation approaches and methodologies in evaluation practice.

References

- Chavis, D. (2004). Looking the enemy in the eye: Gazing into the mirror of evaluation practice. *The Evaluation Exchange*, 9, 8-9.
- Chen, H.-T. (1990). *Theory-driven evaluation*. Newbury Park, CA: Sage.
- Chen, H.-T., & Rossi, P. (1987). The theory-driven approach to validity. *Evaluation and Program Planning*, 10(1), 95-103.
- Christie, C. A. (2003). What guides evaluation? A study of how evaluation practice maps onto evaluation theory. In C. A. Christie (Ed.), *New directions for evaluation: The practice-theory relationship in evaluation* (Vol. 97, pp. 7-36). San Francisco: Jossey-Bass.
- Christie, C. A., & Alkin, M. C. (2003). The user-oriented evaluator's role in formulating a program theory: Using a theory-driven approach. *American Journal of Evaluation*, 24, 373-386.
- CHS Organization. (2002). *Stress variables scale*. Trenton, NJ: Author.
- Cook, T. D. (1985). Post-positivist critical multiplism. In R. L. Shotland & M. M. Mark (Eds.), *Social science and social policy* (pp. 21-62). Beverly Hills, CA: Sage.
- Davidson, E. J. (2000). Ascertaining causality in theory-based evaluation. In P. J. Rogers, A. J. Petrosino, T. A. Huebner, & T. Hacsí (Eds.), *New directions for evaluation: Challenges and opportunities for program theory evaluation* (Vol. 87, pp. 17-26). San Francisco: Jossey-Bass.
- Davidson, E. J. (2003). Linking organizational learning to the bottom line: Methodological issues, challenges, and suggestions. *The Psychologist-Manager Journal*, 6, 54-67.
- DeBruin-Parecki, A. (1999). *Assessing adult/child storybook reading practices* (CIERA Report No. 2-004). Ann Arbor: University of Michigan, School of Education, Center for the Improvement of Early Reading Achievement.
- Donaldson, S. I. (2001). Mediator and moderator analysis in program development. In S. Sussman (Ed.), *Handbook of program development for health behavior research* (pp. 470-496). Newbury Park, CA: Sage.
- Donaldson, S. I., & Gooler, L. E. (2002). Theory-driven evaluation of the Work and Health Initiative: A focus on winning new jobs. *American Journal of Evaluation*, 23, 341-346.
- Family and Children Early Education Services. (1997a). *Head Start family and child experiences survey: Parent interview*. Washington, DC: Head Start Program.
- Family and Children Early Education Services. (1997b). *Head Start family and child experiences survey: Staff interview*. Washington, DC: Head Start Program.
- Fetterman, D. M. (1996). Empowerment evaluation: An introduction to theory and practice. In D. M. Fetterman, S. J. Kaftarian, & A. Wandersman (Eds.), *Empowerment evaluation: Knowledge and tools for self-assessment and accountability* (pp. 3-49). Thousand Oaks, CA: Sage.
- Fitzpatrick, J. L., Sanders, J. R., & Worthen, B. R. (2004). *Program evaluation: Alternative approaches and practical guidelines* (3rd ed.). Reading, MA: Addison-Wesley.
- Gadsden, V. L., Brooks, W., & Jackson, J. K. (1999). Promoting literacy and the concept of possibility. In H. J. Walberg & J. Herbert (Eds.), *Promoting positive outcomes* (pp. 253-295). Washington, DC: Child Welfare League of America, Inc.
- Gargani, J. (2003, November). *The history of theory-based evaluations 1909-2003*. Paper presented at the annual conference of the American Evaluation Association, Reno, NV.
- Goodman, G. D. (2002). *The internal world and attachment*. Hillsdale, NJ: Analytic Press.
- Lipsey, M. W. (1990). *Design sensitivity*. Thousand Oakes, CA: Sage.
- Mertens, D. L. (2003). The inclusive view of evaluation: Visions for the new millennium. In S. I. Donaldson & M. Scriven (Eds.), *Evaluating social programs and problems: Visions for the new millennium* (pp. 91-108). Mahwah, NJ: Lawrence Erlbaum.

- Michalski, J. H., Mishna, F., Worthington, C., & Cummings, R. (2003). A multi-method impact evaluation of a therapeutic summer camp program. *Child and Adolescent Social Work Journal, 20*, 53-76.
- Morisset, C. E. (1997). "You know, I think my boy's gonna be a reader!" *Supporting school readiness through community-based intervention* (Report No. 37). Washington, DC: Office of Educational Research and Improvement.
- Myers, D. G. (2002). *Psychology* (6th ed.). New York: Worth.
- Nistler, R. J., & Maiers, A., (2000). Stopping the silence: Hearing parents' voices in an urban first-grade family literacy program. *The Reading Teacher, 53*, 670-680.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text*. Thousand Oaks, CA: Sage.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York: Oxford University Press.
- Sarafino, E. P. (2002). *Health psychology: Biopsychosocial interactions* (4th ed.). New York: John Wiley.
- Scriven, M. (1991). *Evaluation thesaurus*. Thousand Oaks, CA: Sage.
- Scriven, M. (1997). Minimalist theory: The least theory that practice requires. *American Journal of Evaluation, 19*, 575-604.
- SenGupta, S. (2004). On methodology: Rip Van Evaluation and the great paradigm war. *The Evaluation Exchange, 9*, 3-4.
- Shaw, K. M., & Replogle, E. (1996). Challenges in evaluating school linked services: Toward a more comprehensive evaluation framework. *Evaluation Review, 20*, 424-469.
- Stufflebeam, D. (2001). Evaluation models [Entire issue]. *New Directions for Evaluation, 89*.
- Towns, D. P., Cole-Henderson, B., & Serpell, Z. (2001). The journey to urban school success: Going the extra mile. *Journal of Negro Education, 70*, 4-18.
- U.S. Census Bureau. (2000). *Trenton city, New Jersey statistics and demographics*. Retrieved from <http://trentonnj.areaconnect.com/statistics.htm>
- Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: MIT Press.
- Weiss, C. H. (1997). How can theory-based evaluation make greater headway? *Evaluation Review, 21*, 501-524.