
Nonprofit Organizations and Outcome Measurement: From Tracking Program Activities to Focusing on Frontline Work

American Journal of Evaluation
33(3) 431-447
© The Author(s) 2012
Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1098214012440496
<http://aje.sagepub.com>



Lehn M. Benjamin¹

Abstract

Why do we continue to see evidence that nonprofit staff feel like outcome measurement is missing important aspects of their work? Based on an analysis of over 1,000 pages of material in 10 outcome measurement guides and a focused literature review of frontline work in three types of nonprofit organizations, this article shows that existing outcome measurement frameworks focus on how staff implement programs rather than how staff work with clients. Outcome measurement guides direct nonprofits to track program activities completed and the outcomes resulting from those program activities. In contrast, the accounts of frontline work in nonprofits show that nonprofit staff start by building a relationship with the person they are serving and then adjusting programs and services to better meet the needs and goals of this individual. Consequently, outcome measurement may go some distance in helping us understand nonprofit performance but may also mischaracterize nonprofit performance.

Keywords

nonprofit, outcome measurement, frontline work, accountability

It has been 15 years since the United Way first published its outcome measurement guide for nonprofit organizations. Over that time period, it has sold more than 170,000 copies and many other guidebooks and tools have been developed to help nonprofits measure outcomes. Perhaps not surprisingly, recent national surveys of nonprofits show that measuring outcomes is the most common purpose for undertaking evaluative activities by nonprofits and half of nonprofits have developed logic models or theories of change (<http://www.innonet.org>). Despite these efforts, nonprofit staff express a persistent concern that outcome measurement misses important aspects of their work (e.g., Carman, 2007, p. 68; Hwang & Powell, 2009, p. 294).

¹ Department of Public and International Affairs, George Mason University, Arlington, VA, USA

Corresponding Author:

Lehn M. Benjamin, Department of Public and International Affairs, George Mason University, 217 Truland Building 4D5, 3330 N. Washington Boulevard, Arlington, VA 22201, USA
Email: lbenjami@gmu.edu

To elucidate what might be missing and to inform efforts to better measure nonprofit outcomes, this article compares the content of 10 outcome measurement guides targeted at nonprofits with empirical studies and firsthand accounts of what nonprofit staff do as they work with those they serve. The purpose of this comparison was to examine the extent to which prominent outcome measurement frameworks actually measure what nonprofit staff do to support positive outcomes for their target populations. The analysis shows that existing outcome measurement frameworks focus on how nonprofit staff implement programs rather than focusing on how nonprofit staff work with clients. More specifically, current outcome measurement frameworks direct nonprofits to track program activities completed (e.g., parenting education workshops completed) and the outcomes resulting from those program activities (e.g., parents support children with school work, delinquency rates decline). In contrast, accounts of frontline work show that nonprofit staff start by building a relationship with the person in front of them. This relational work is necessary for staff to adjust programs and strategies to better meet the goals and needs of the individual; to determine possible courses of action in light of the individual's circumstances and capacity; and to mobilize outside resources in support of the individual's goal. Comparing the guides and the frontline work reveals how existing outcome measurement frameworks do not adequately capture the relational work staff do even though this relational work is critical for the ability of nonprofits to support positive outcomes for those they serve. Consequently, while current outcome measurement efforts may go some distance in helping us understand nonprofit performance, this analysis suggests that these efforts risk mischaracterizing nonprofit performance.

This article is organized as follows: The next section reviews the existing literature on nonprofits and outcome measurement and describes how the analysis presented here contributes to this literature. I then examine frontline work by reviewing empirical studies and firsthand accounts of practice in three areas: human services, neighborhood development, and grassroots organizing. These three areas were chosen because they capture different ways that nonprofit staff work with their target population to address poverty and distress: human service staff focus on change at the individual or family level; neighborhood development staff focus more on change at the community level; and grassroots organizing staff focus on change at the policy level. While not representative of all nonprofits, this diversity helped to illuminate whether the aspects of nonprofit work not well captured in current outcome measurement frameworks were confined to a specific subset of nonprofits. I present an analytic framework, based on this review, for characterizing the frontline workers' activities. The fourth section presents the analysis of the 10 outcome measurement guides focusing specifically on what nonprofits are told to measure about their work. These guides, totaling over 1,000 pages of text, were identified through a systematic scan of organizational websites of funders, management support organizations, and professional associations. The final section compares what nonprofits are told to measure in the guides to the analytical framework presented in the third section and discusses the risks of focusing narrowly on program activities completed and the outcomes resulting from those program activities. I outline areas for further research in the final section.

Outcome Measurement in the Nonprofit Sector

Outcome measurement involves the ongoing process of defining, monitoring, and using performance indicators to improve organizational effectiveness and efficiency (Poister, 2003). Outcome measurement started getting a lot of attention by nonprofit staff during the 1990s in response to funder requirements for measurable results (see <http://www.geofunders.org>), although the concepts and ideas have deeper roots in evaluation (Shadish, Cook, & Leviton, 1991; Wholey et al., 1986). At that time, outcome measurement was seen as distinct from past evaluative efforts in a few important respects: (a) it was a tool to help nonprofits move past tracking activities to assessing how their target population was better-off as a result of the services provided, (b) it put evaluation more firmly in

the hands of nonprofits, and (c) it emphasized improving practice over determining a definitive causal relationship between programs and measurable results (New England Nonprofit Quarterly, 1998).

Over the course of 20 years consultants, funders, national associations, and management support organizations have developed guidebooks and tools to help nonprofits develop outcome measurement systems for their organizations as well as shared standardized outcome measurement systems (<http://www.fsg.org/impactareas/overview.aspx>). With all this attention to measurable outcomes, it is not surprising that national surveys of nonprofits show that measuring outcomes is the most common purpose for undertaking evaluative activities by nonprofits (Fine, Thayer, & Coghlan, 2000; <http://www.innonet.org>).

A number of studies have examined the adoption and implementation of outcome measurement by nonprofits. Some of this work has documented the adoption of outcome measurement by nonprofits in response to funder requirements (Benjamin, 2008; Campbell, 2002; Christensen & Ebrahim, 2006; Cutt & Murray, 2000; Ebrahim, 2005; Hwang & Powell, 2009; Julian, 2001). Other studies have looked at nonprofit organizations' efforts to measure and manage for outcomes, highlighting successful practices (Hendricks, Plantz & Pritchard, 2008; Houchin & Nicholson, 2002; Morley, Vinson, & Hatry, 2001; Sawhill & Williamson, 2001). Researchers have documented the prevalence of evaluative activities among nonprofit organizations in specific locations (Carman, 2007; Carman & Fredericks, 2008; Eckerd & Moulton, 2011; Fine et al., 2000; LeRoux & Wright, 2010; Zimmermann & Stevens, 2006).

Yet, across these studies, we see evidence that nonprofit staff believe outcome measurement is missing valuable aspects of their work and consequently failing to capture important differences they are making in the lives of those they serve (e.g., Benjamin, 2004; Carman 2007; Carman & Fredericks, 2008; Hwang & Powell, 2009). As one practitioner put it, "There are some things that just don't seem to be an easy fit for the model . . . So that's why part of the concern is that if you strictly follow the outcome measurement model you may not be adequately representing your agency. Yeah, you'll have the things down there but there may be a whole list of other things that may be most important that are not being included" (Director of Community Center, personal communication, August 15, 2002).

Existing literature suggests four possible reasons for the mismatch between outcome measurement frameworks and the work that staff do to ensure positive outcomes for those they serve. First, outcome measurement has become tightly coupled with funding requirements, reinforcing the tendency to focus data collection on what is necessary for reporting to donors rather than what is necessary for internal development and learning (Christensen & Ebrahim, 2006; Cutt & Murray, 2000; Ebrahim, 2005). Second, performance measurement frameworks do not currently measure the important role nonprofits play in engaging citizens in the collective effort to address social problems (Knutsen & Brower, 2010; Smith, 2010). Third, as a standardized framework, outcome measurement conflicts with the experimental nature of nonprofit work (Hwang & Powell, 2009). Fourth, nonprofits do not have the capacity to adequately measure outcomes (Botcheva, White, & Huffman, 2002; Carman & Fredericks, 2008).

While this literature sheds some light on why nonprofit staff believe that outcome measurement frameworks are missing valuable aspects of their work, there has been no systematic examination of how nonprofits are actually guided to measure their performance or how these outcome measurement frameworks align with what nonprofits do to achieve positive outcomes for those they serve. Comparing what nonprofits are directed to measure with what staff actually do to achieve outcomes for their clients is important for two reasons. First, recent national survey results show that 79% of nonprofits are measuring their own performance without the help of a professional evaluator, suggesting that nonprofits are relying on other resources, including "how to" guidebooks and tools, to help them measure their performance. And even though no systematic survey has been done to determine what outcome measurement guides nonprofits are using, half of nonprofits surveyed in a

national study reported having developed a logic model/theory of change (<http://www.innonet.org>) and records from publishers show that significant numbers of well-known guidebooks have been distributed. For example, the *United Way's Outcome Measurement Guide* has sold more than 170,000 copies, while the *Kellogg Foundation's Logic Model Guide* has a distribution of 20,000, and *The Rensselaerville Institute's Outcomes Funding Framework* is in its fourth printing (Publishers, personal communication, April 2011).

Second, if nonprofit staff believe outcome measurement is missing important aspects of their work, then looking more closely at this work is a logical step. Existing research has focused on the nonprofit organization as a whole (e.g., documenting the evaluative activities undertaken by the nonprofit; the effect of outcome measurement requirements on the nonprofit). But any effort by a nonprofit to achieve better outcomes must be translated into the frontline encounter (Brodkin, 2008). Only by looking at what staff actually do as they work with users to achieve positive outcomes can we understand what may be missing in existing outcome measurement frameworks. As one nonprofit practitioner explained, "I think [outcome measurement] is a worthwhile thing. I am not complaining about the time it takes; it is a tool that needs to be continuously refined and made better, but it needs to get closer to the actual activity" (Interview cited in Hwang & Powell, 2009, p. 294). This article compares what frontline staff actually do to support positive outcomes for those they serve with the content of 10 outcome measurement guides to elucidate what aspects of nonprofit work may be not well captured by existing outcome measurement frameworks.

Frontline Work in Nonprofits

Frontline work is the direct work staff and volunteers do with users—clients, residents, constituents, or peers—to change some circumstance and improve their lives. This section focuses specifically on nonprofits that provide human services, nonprofits that redevelop neighborhoods, and nonprofits that engage in grassroots organizing to address poverty and distress. Broadly speaking, human service work involves building skills and increasing knowledge to encourage individuals and families to adopt practices or behaviors that can lead to a better life. Grassroots organizing work builds leadership within marginalized communities so that these communities are better able to take action that changes institutions and policies that contribute to their distress. Neighborhood development work involves revitalizing the physical, economic, and social infrastructure of neighborhoods.

Reviewing the literature in these three areas sheds light on the work nonprofit staff do as they seek to serve marginalized communities. These three areas were chosen because they capture different ways that nonprofit staff work with their target population to address poverty and distress: human service nonprofits work more at an individual level, community development nonprofits work at a neighborhood level, and grassroots organizing nonprofits work to change policy. While not representative of all nonprofits, this diversity helps illuminate whether the aspects of nonprofit work not well captured in current outcome measurement frameworks is confined to a specific subset of nonprofits. However, the conclusions drawn here cannot be generalized to other types of nonprofits.

While recognizing the diversity of this work, the literature suggests that frontline staff in these nonprofits engage in four common relational practices—listening, naming, challenging, and linking—as they adjust programs or services to meet the specific needs and goals of users; work with users to determine the best course of action in light of the user's particular circumstances; and mobilize other resources outside of the program or organization to ensure positive outcomes for users. Examples are given for each of the four tasks below and Table 1 summarizes this work. The conceptualization of this work requires further empirical probing and this is addressed in the final section.

Table 1. Nonprofit Frontline Work

Relational work	Examples from human service	Examples from grassroots organizing	Examples from neighborhood development
Listening	Listen to clients for evidence of willingness to change	Listen to constituents for evidence of commitment and enthusiasm	Listen to residents to identify pressing concerns in the neighborhood
Naming	Work with clients to define individual service plan	Work with constituents to define a collective account of common experiences	Work with residents to set the agenda for neighborhood development
Challenging	Challenge clients to make behavioral changes while recognizing clients as authorities on their problem.	Challenge constituents to take on new leadership roles while providing the support they need to be successful	Challenge residents to take charge of their neighborhood while recognizing the limits of the organization to address resident concerns
Linking	Connect clients with other community supports	Connect constituents with others facing similar concerns	Connect residents with resource holders outside the community

Listening

Listening to users is central to the work of frontline staff. Depending on the organizational mission, nonprofit staff listen for certain experiences, capacity, emotions, and context. Listening not only involves attending to what someone says, often with a specific programmatic or organizational filter, but also asking skillful questions to convey interest and care and demonstrate attention. For example, delivering human services requires staff to listen to users in a way that shows respect and an openness and receptivity to client concerns. All of this is necessary to build the personalized relationships necessary for “client cooperation” (Clary, 1987; Handler, 1992). Listening in human service organizations may be shaped by assumptions staff have about the social worth of the person, attribution of responsibility for the client’s problem, amenability of the client to change, desired end results, and the view of the client as an object or a subject (Hasenfeld, 2000, 2010). These assumptions can be challenged or reaffirmed depending on how staff listen to clients, including how they respond to what they hear; whether they acknowledge what users say, and how they ask questions to understand the context of users’ lives (<http://www.buildingmovementproject.org>).

Similarly, for nonprofits that engage in grassroots organizing, listening is a central mode of practice (Gecan 2004; Warren 2001). Organizers learn that “[Organizing is] the totally unglamorous repetitive process of listening to individuals one on one. You listen to each person very carefully. You hear what people say. You look for the emotions behind their thoughts. You look for the values that shape their thoughts” (Eichler, 2007, p. 43). Listening is far from a homogenous practice for organizers. Different organizing traditions advise organizers to listen for specific information when interacting with community members. For example, organizers working from a power-based tradition listen for signs of potential leadership, or “someone with enthusiasm, energy and commitment,” (Smock, 2004, p. 39). Organizers working from a women’s organizing tradition listen for what they can learn from those most reticent to speak, asking open-ended questions as a way of encouraging the development of voice (Belenky, Bond, & Weinstock, 1997; Stall & Stoecker, 1998).

Naming

Naming refers to how staff work with their target populations to define problems and determine possible courses of action. The way nonprofit staff work with their target populations to develop an account of the problem that spurs specific action varies across nonprofits. For example, for

grassroots organizers naming has been an explicit social change strategy: naming is used to translate a common set of individual experiences into a collective account that creates a common bond, reduces isolation, and provides a basis for collective action (Chetkovich & Kunreuther, 2006; Evans & Boyte, 1992; Mansbridge & Morris, 2001; Smock, 2004). In her review of major organizing models, Smock (2004) notes that diagnosing and framing the community's problems is a central task for organizers: "[they] provid[e] a framework that residents can use to reflect on and make sense of their experiences . . . by helping residents to identify causal explanations for their grievances, the organizations shape the strategies that residents adopt for addressing these problems" (pp. 6–7).

For frontline staff in neighborhood development organizations, naming captures a core dilemma of practice. Neighborhood development organizations, like community development corporations, grew out of the frustration of inner-city leaders with policies and programs that were imposed from the outside and the belief that strategies to address the problems of the inner city should be controlled and defined by the community (Halpern, 1995). Consequently, neighborhood development organizations are supposed to be resident led but housing development goals require partnerships with lenders and others that can limit the extent to which residents can actually influence the nonprofit organization's agenda (Silverman, 2005). Consequently, staff walk a tightrope between encouraging residents to "name" or identify neighborhood problems that should guide the organization's work and the concerns of external funders. Not knowing how to handle this process can result in the organization losing the trust and faith of residents and any claims of "representing the neighborhood" made by the nonprofit.

Challenging

Challenging refers to the work nonprofit staff do to determine how and when to nudge someone to take an action necessary to produce desired changes. Challenging is a term that connotes both struggle and victory (<http://www.thefreedictionary.com/challenging>). Challenging requires a level of attention by frontline staff to an individual's readiness to take risks by trying something new or letting go of something old: whether it is encouraging a woman to leave an abusive relationship; encouraging a resident to organize a neighborhood watch, even though this could mean reporting on residents they know; or encouraging a new leader to facilitate a meeting with public officials. Not challenging adeptly can set someone up for failure or leave them stuck in circumstances that are detrimental. Challenging also includes the work required to reflect and respond when users question staff suggestions. Not responding adeptly to challenges from users can mean a breach in trust, loss of legitimacy, or a missed opportunity to improve nonprofit practice.

For grassroots organizers, building leadership requires understanding an individual's "path to power" (Laura Barrett, Gamaliel Foundation, personal communication, August 25, 2011). "Developing leaders doesn't just happen by default, it takes an intensive process of recruitment, training, mentoring and support . . . organizers challenge new recruits to take on increasingly difficult roles until they are confidently engaged in levels of leadership that they never may have thought possible . . . throughout that process the organizers work to ensure that the leaders are successful in their new roles" (Smock, 2004, p. 38). Challenging emerging leaders to take on new responsibilities too soon can set them up for failure while failing to challenge emerging leaders can result in an organizer getting out in front, taking on tasks or roles that should be done by emerging leaders (Eichler, 2007; Gittell & Vidal, 1998). Similarly, human service staff are in the business of changing people: "the daily tasks focus on providing treatment, education or socialization with the intent of altering the physical, psychological, social or cultural attributes of their clients" (Sandfort, 2003, p. 607). Here, frontline staff balance a tension between acceptance and change. On one hand, staff should accept users as they are, without moral judgment, and as important experts in their own lives. On the other hand, staff bring professional expertise to bear on user's problems and are in the position of trying to support users to change (Berlin, 2005).

Linking

Linking refers to the work nonprofits do to connect those they serve to other individual and institutional actors that can support change. Human service workers are encouraged to increase client power by, among other things, “linking clients to supportive social networks that can lend them resources and reduce dependence on the agency” (Hasenfeld, 1992b, p. 270). In her work on empowering human services, Naomi Gottlieb (1992) notes that those organizations where the staff recognized and valued the power of a user’s peers, created conditions of greater autonomy and self-direction among users. The degree to which human service workers build relationships among users and other social supports depends on the type of service delivered and the professional norms of the organization.

In neighborhood development, staff link residents both with other residents, usually along neighborhood or block lines and link the neighborhood and its residents with institutions outside neighborhood boundaries. Neighborhood development staff may organize block parties or neighborhood socials for residents to get to know one another. They also work to build relationships between outside resource holders (employers, banks, commercial establishments) that might invest in the neighborhood. For example, Kretzmann and McKnight (1993), the originators of asset-based community development, explain that “one of the central challenges for asset based community developers is to constantly build and rebuild relationships between and among local residents, local association and local institutions” (p. 9).

Summary

Accounts of frontline work in nonprofits show that staff do more than complete a set of program activities, they do relational work. This relational work is critical for supporting positive outcomes for those they serve, both those intended by the program and other outcomes desired by users. This relational work is necessary to adjust programs and services to meet the particular goals and needs of users; to determine possible courses of action, given the user’s situation, and to mobilize outside resources all in an effort to support positive outcomes for users.

Outcome Measurement: Measuring What Matters About the Work

Over the past 20 years, many tools and guidebooks have been developed to help nonprofits measure outcomes. These guides are a source of normative messages about what nonprofit organizations should measure in order to document their effectiveness. These tools and guidebooks are not simply sources of technical resources on how to measure outcomes; they are carriers of normative ideas about practice (Powell & DiMaggio, 1991; Prior, 2003; Scott, 1995). Outcome measurement guides reflect dominant conceptions of nonprofit performance as well as direct nonprofits in the process of measuring their performance.

Selection and Analysis

The 10 guides selected for this analysis were identified through a scan of organizational websites conducted between August 2006 and March 2007. The organizational websites included funders from the Neighborhood Funders Group, local and national professional associations of nonprofits, management support organizations including those that focused on evaluation (e.g., Harvard Family Research Project, Innonet, TCC Group) as well as nonprofit professional publications (e.g., *Chronicle of Philanthropy*, *Nonprofit Quarterly*, *Nonprofit Times*). The assumption was that if nonprofits were interested in or required to develop an outcome measurement system, they would be guided by organizations that would have some expertise or authority on outcome measurement in their

environment. In total, close to 500 websites were scanned using key words: accountability, performance, evaluation, and outcome.

Four publications—the Harvard Family Research Project newsletter, a publication by The Rensselaerville Institute, a report by NeighborWorks, and a guide for adopting Results-Oriented Management and Accountability (ROMA)—reviewed outcome measurement resources. Of the 19 resources listed in these publications, 10 of them were guides that appeared on more than one list and were targeted to nonprofits addressing a range of social problems (in contrast to resources targeted to government or nonprofits addressing a specific issue like substance abuse). These guides were chosen for analysis (see Appendix A). These 10 guides are not inclusive of all available outcome measurement material out there nor are these guides representative. This would assume that there is a population of outcome measurement material that can be readily identified and selected. However, the 10 guides collectively provide a reasonable proxy of how nonprofits are directed to measure their performance.

Once the guides were identified, the texts were then converted into Microsoft Word and uploaded into ATLAS.ti. The coding involved three analytical steps. First, the content was sorted into three categories using the following questions: (a) what are nonprofits told to measure? (b) how are they told to measure it? and (c) how are they guided to use this information? The findings reported here focus on the first question. Second, emergent themes were identified. For example, what were nonprofits told to measure were coded into subcodes: outcomes and process. In the next iteration, this coding scheme was refined. For example, “outcomes” were further coded into types of outcomes: knowledge, conditions, behavior, and so on. Finally, in addition to combing through the texts multiple times to further refine the codes, a search and an automatic code function in ATLAS.ti was used to identify all the instances of customer, client, beneficiary, participant, constituent, resident, community in order to detect any unanticipated themes or messages about the target population. In addition, all instances of terms such as outcome (to examine how the term was defined and the direction given to nonprofits about identifying their outcomes) and accountability (to examine how the term was used and in relationship to what set of stakeholders) were examined. In what follows, three measurement directives (in the form of advice provided by each guide) are presented with supporting evidence from the texts.

General Description of the Guides

The 10 guides shared a common purpose: encouraging and supporting nonprofits in the measurement of outcomes and using these data to improve effectiveness in addressing social problems. The 10 guides explain the rationale for measuring outcomes, including how measuring outcomes is distinct from reporting on activities. The guides then detailed how nonprofits should develop their outcome measurement system, including how to identify outcomes and other performance measures. This included positing questions and using models or templates. Despite the similarities, each of the guides had their unique emphasis. For example, Results-Based Accountability (RBA) emphasized the distinction between population and program accountability. Targeting Outcomes of Programs (TOPS) provided an integrated framework for program planning and evaluation. The Balanced Scorecard (BSC), originally developed by Kaplan and Norton for firms, looked at the internal drivers of external performance (internal process, employee learning, and growth). The Success Measures (SM) guide offered a roadmap for a participatory approach to evaluation and performance measurement and The Rensselaerville Institute’s Outcomes Funding Framework (TRI) suggested tracking milestones customers achieve in order to manage for performance. Two of the guides targeted a sub-field of nonprofits (e.g., SM is targeted to community development organizations; ROMA is targeted to community action agencies) and addressed issues and concerns reflective of that field. The TRI guide targeted funders as well as nonprofit organizations.

Measurement directives about process. Nine of the ten guides pointed to process measures as insufficient for assessing nonprofit performance (see Table 2). In most guides, the “process” of working with users was defined as completing program-related activities (e.g., workshops delivered, counseling session completed, and students served). For example, one guide explained “Output measures [are the] results generated from the use of program inputs . . . These metrics track the number of people served, services provided, or units produced by a program or service. They may sometimes be referred to as activity measures” (BSC).

While most of the guides equated process with the completion of program-related activities, the advice in two guides suggested that the process cannot be equated to the completion of program-related activities. One guide stated: “It does not seem right to think of clients, workers, supplies and office space as inputs to the service sausage machine, producing outputs of cured, served, or fixed clients. The implication that there is a mechanistic relationship between inputting staff resources and outputting customer benefits seems absurd, if not insulting to [service providers] . . . working with people is simply more complicated than working with machines” (RBA). Outcomes Engineering (OE) noted: “For transformation-focused work . . . there are emerging opportunities that must be seized then-and-there or are lost . . . The responsiveness of the initiative in seizing these opportunities, as well as the quality of exchange between the parties, are the true measures of an initiatives performance.” Other guides similarly pointed to the need to measure not just the activities completed but how users felt about the quality of the interaction; for example, the guides suggested asking users about the pleasantness or friendliness of the service experience, counselor compassion, and dignity of treatment (BSC, Kellogg Foundation [KF], RBA, United Way of America [UWA]).

Measurement directives about responsiveness. The guides also directed nonprofits to measure responsiveness to users, as an indicator of good practice but how responsive was conceptualized varied across the guides and at times seemed to conflict with other directives suggesting nonprofits measure whether the program was implemented as intended. For example, responsiveness was conceptualized as (a) an ongoing process of adjusting to user’s needs (OE); (b) timely service (UWA, Urban Institute [UI]); and (c) as a reflection of the interests and priorities of those served (SM). At the same time, three of the guides implied that nonprofits measure whether programs were implemented as intended (KF, RBA, TOPS). For example, one guide noted that “most programs are designed to provide certain services in a certain way. Process evaluation can test if the program is operating true to design” (RBA). It was unclear how nonprofits were supposed to balance being responsive to users with program fidelity or to what extent frontline staff should adjust what they do in response to those they serve and when such adjustment may compromise desired outcomes.

Measurement directives about outcomes. All the guides offered distinct tools to help nonprofit staff identify meaningful outcome measures. Yet, outcomes were conceptualized differently across the guides. Five guides defined outcomes as changes in the target population as a result of participating in program activities (KF, ROMA, TRI, UWA, UI). Here, the program was an intervention to which the client or user responded and the outcomes were those that followed the program logic. In contrast, two of the guides defined outcomes as changes desired by the target population to which the program contributes (SM, OE). Here, users were the primary focus and the program or services were tools to support users in achieving desired outcomes. One guide also suggested nonprofits measure how staff and the organization changed as a result of the work with users, not just how the users change as a result of receiving services or participating in the program (OE).

These three sets of directives—directives about process, directives about responsiveness, and directives about outcomes—illustrate how current outcome measurement frameworks define nonprofit performance in terms of program activities completed and changes resulting from those program activities. The focus on program activities and program outcomes in existing outcome

Table 2. Measurement Directives About Process

Guide	Text from guides
Balanced scorecard	Simply counting people served or dollars spent won't cut it in today's environment. You need to demonstrate advancement on the high-level, mission-based objectives that your constituents are requiring you to provide.
Kellogg Foundation Logic model	Some grantees think activities are ends unto themselves. They report the numbers of participants they reach or the numbers of training sessions held as though they were results . . . Conducting an activity is <i>not</i> the same as achieving results from the accomplishment of that activity
Outcomes engineering	Today, most stakeholders insist that initiatives document and somehow quantify the sustained changes occurring in the lives of those they reach. They want evidence of the worth of the initiatives they support in outcome units and not in process terms
Results-based accountability	In the change-agent model, the "number of clients served" is not an end product. Serving clients is a means to a change in customer or social conditions, the true end or purpose of the work
Results-oriented management and accountability	Whether or not you think it is a good idea, the new emphasis on describing what you do in terms of results and outcomes is here to stay . . . So ROMA is an effort to shift our attention away from the inputs and operations of daily activity to the benefits that are produced for the participants and the community as a whole.
Success measures	The benefits from a program are not the same thing as program activities. Benefits are the positive changes we want to achieve in our community as a result of our activity. Activities are building houses, running training programs or youth programs. The changes make up the benefits picture, and are the purpose of our activities
Outcomes funding framework	A problem stemming from the endless nature of processes is the difficulty in asking for outcomes. It is always possible to join the procedural march. It is much harder to step out of the parade to gauge just how far you have come, or when you will or should stop. Indeed, the process mindset asks us to focus not on outcomes but on activities. It is what we do that counts
United way	Both types of measurement are important for managers, but in past decades the focus has been on the outputs. The new, common sense approach is concerned with provision of benefits to the public
Urban Institute key steps	Outputs are expected to lead to desired outcomes, but by themselves do not tell anything about the outcomes

measurement guidebooks is perhaps not surprising. Evaluation emerged as a distinct field during the War on Poverty. At the time, social scientists were called on to evaluate the effectiveness of major federal policy initiatives. Early evaluators were focused on methodological innovation and theories of value and knowledge (Shadish et al., 1991). Later evaluators started to offer more developed ideas of social programs, drawing on implementation studies coming out of public administration (Chen, 1990; Lipsky, 2010; Pressman & Wildavsky, 1973). Since that time evaluation scholars have called for richer theories of practice (Chen 1990; Schwandt, 2005; Shaw & Faulkner, 2006; Urban & Trochim, 2009; VanderPlaat, 1995). Despite these calls, this analysis shows that outcome measurement models still conceptualize nonprofit performance in terms of program implementation: the completion of program activities and the outcomes resulting from these activities.

While assessing the outcomes of program activities is critical, the accounts of frontline work shows that program activities are only one set of activities that staff undertake to support positive outcomes for those they serve. For example, outcome measurement frameworks suggest that

nonprofits measure programmatic activities like the number of parenting education classes completed in a student delinquency prevention program and the results of those programmatic activities such as (i) parents complete program; (ii) parents provide more school encouragement to their children' (iii) children have better attendance at school, fewer behavioral problems in school and improved grades. But nonprofit staff do more than just complete parenting education classes; nonprofit staff will spend time getting to know the parents and the particular issues they face with their children. These relationships will help the staff adjust and at times change the workshop content and/or delivery to better meet the particular needs of parents. These adjustments may spur changes in program or organizational goals and inform the advocacy work of nonprofits. At the same time, staff may spend quite a bit of time working with individual parents to determine the steps they are able to make at home to support their children in school. This means that users may have different short-term outcomes as they address their child's challenges in school. Staff may also link parents to outside resources that can address the general well-being of the family, which may be preventing the parents from fully participating in the program. All this relational work is critical in enabling parents to access the program, encouraging them to take positive action with their children, and ensuring that changes they make are more sustainable. Indeed, the differences among the guides also suggest this possibility. Consequently, these efforts risk mischaracterizing nonprofit performance.

Beyond Program to Frontline Work

This analysis suggests that developing outcome measurement frameworks that are informed by all the work frontline staff do is essential if outcome measurement is going to help nonprofits assess their contribution to the positive changes in users' lives and identify areas for improvement. Three initial observations follow from the comparison of how nonprofit work is conceptualized in the outcome measurement guides versus in the descriptions of frontline work in empirical studies and first-hand accounts. First, the relational work of frontline staff is underconceptualized in the guides. Although the guides suggest that nonprofits measure service quality, including staff pleasantness and empathy, the accounts of frontline practice suggest that the quality of the interaction is not an addendum to program activities but that it anchors and organizes these activities. It is hard to imagine how staff would be able to work with residents to address neighborhood safety, provide small business development workshops to budding entrepreneurs, or counsel victims of domestic violence without developing a relationship with those they serve. The depth and breadth of this relational work varies: This work may be short or long term, more intensive around one issue or more comprehensive in nature, take place at a single point of service or at different levels within the organization (residents sit on the governing board, recipients receive program services, clients volunteer to help other clients). The limited attention to this work in the outcome measurement material leaves little room for understanding how and when this work contributes to long-term outcomes, when this work may be enabling poorly designed or underresourced programs to show evidence of success, or when this work may be counterproductive to changes desired by users.

Second, the analysis of the guides reveals a varied conception of how much staff should adjust what they do to meet the specific needs of those they serve. The guides direct nonprofits to measure responsiveness to users but responsiveness is conceptualized in different ways across the guides and on the whole the individual guides do not offer a clear conceptualization of this adjustment work. The accounts of frontline practice suggest that adjustment work may be more of a continuum, where some nonprofit agendas are adjusted based on input from users while others make adjustments within the confines of the program. Regardless, without recognizing the way staff adjust program and services based on user experiences and goals, the outcome measurement material cannot shed light on the extent to which staff make changes, how often this occurs, what this suggests about

program assumptions or policy prescriptions, or when such adjustments is the result is a lack of clear direction given to frontline staff.

Finally, the analysis of the guides shows that by and large outcomes are defined as the result of program activities. This certainly was not the case across all of the outcome measurement material but the image of the client, constituent, or resident as a receiver and reactor to programs was certainly evidenced as was the conception of outcomes as program outcomes. In contrast, frontline accounts suggest that staff start with the person they are serving, their particular concerns and goals, and then work to organize program and nonprogram resources to support positive change for users. When users are portrayed as responding to program activities in ways that correspond to program logic, users seem more like a means to achieving program outcomes, rather than the program activities supporting the users in achieving desired outcomes.

This analysis shows that equating nonprofit performance with the completion of program activities and changes resulting from those program activities risks mischaracterizing nonprofit performance. For example, not recognizing the work staff do to build relationships with clients means we do not understand how programs require different degrees and kinds of relational work or how relational work affects programmatic effectiveness. Not considering the work staff do to adjust programs or services to meet client needs means we miss valuable information about the limits of the program and the possible mismatch between interventions and the target population. Not registering how staff work with users to determine possible courses of action means that we miss understanding how failure to achieve programmatic outcomes by clients may reflect real progress on the part of a client who makes a different set of choices that do not conform to program logic. Not registering how staff work with users to determine possible courses of action means that we also miss understanding how much ownership users have in the changes they make and thus how sustainable measurable outcomes may be. Finally, by not considering the outside resources staff mobilize for users, we fail to understand how much nonprofit success is dependent on larger systems of support.

Pointing to the limits of assessing nonprofit performance by tracking program activities completed and the changes in the nonprofit target population as a result of their participation in program activities is consistent with the conclusions drawn by others who note the diverse roles nonprofits play in addressing social problems (Quarter & Richmond, 2001; Smith, 2010). The analysis presented here suggests that this limitation may not simply be one of inadequate attention to the different roles nonprofits play—social service delivery versus citizen engagement and community participation—but of the conceptions of nonprofit work more generally. Even when nonprofits are primarily involved in social service delivery, staff do build relationships with those they serve; adjust programs and services to meet the particular goals and needs of their target population; work with those they serve to determine the best course of action; and mobilize outside resources all in an effort to support positive outcomes.

While additional research is necessary to make specific recommendations for revising existing outcome measurement frameworks and/or developing alternative frameworks, evaluators could take a few steps to ensure outcome measurement better captures nonprofit performance. First, evaluators might develop guidance to help nonprofits measure the activities that extend beyond the program but that are critical to see positive outcomes from any program intervention. Toward this end, evaluators might consider anchoring the understanding of nonprofit outcomes in the full scope of frontline work. If the frontline encounter becomes the focus, then the program becomes something that structures but does not determine this encounter. For example, evaluators working with nonprofit organizations could start to incorporate relational work in their efforts to help nonprofits measure performance. For some nonprofits, relationship building is significant and for others it is less so. Understanding the importance of the relationship for the work, evaluators could help nonprofits incorporate ways to capture this work in their performance measurement systems. Evaluators could also examine the extent to which frontline staff adjust the program or service to meet the needs and

goals of the client. Assuming that these changes are not simply the result of learning how to implement a new program, evaluators could give greater attention to responsiveness and to how staff adjust the program or service to help users achieve positive outcomes.

This analysis also raises several questions that require additional research by evaluators. First, are the common tasks identified here—listening, naming, challenging, and linking—useful concepts for anchoring efforts to understand nonprofit relational work across diverse organizational settings? Some more specialized outcome measurement models actually capture these tasks. For example, the Women's Funding Network developed a social change vector for its grantees that included listening as one of the four dimensions of practice (along with collaboration, empowerment, action, and outcomes). The SM model reviewed here includes a section on the linking work that nonprofit staff do, that is, community-building outcomes like building social relationships and networks. Evaluators could also draw on the extensive literature on frontline practice to shed additional light on this work and start to address questions such as: To what extent do nonprofits define their goals and strategies in response to those they serve? Do different social problems or nonprofit initiatives require greater responsiveness than others?

Second, building on the point raised earlier, evaluators could look to existing literature on relational work to inform additional research and inquiry in this area. Existing research demonstrates the importance of this relationship work for effectiveness. For example, empirical studies on relational work show how the quality of the relationships staff build with those they serve makes the achievement of task work easier (see Lann-Wolcott, Medvene, & Williams, 2010; Noddings, 2003). Research on emotional labor describes how workers must control their feelings and nonverbal expressions in order to have the desired outcome (see Guy, Newman, & Mastracci, 2008; Hochschild, 1983; Karabanow, 1999). We could ask questions such as: How important is this relationship work for this task work? How much time do staff spend in simply making a connection to a client who is socially disconnected before even considering any kind of support service? What types of relationships work against change? What types support change?

Third, evaluators could examine how the adoption of current outcome measurement frameworks has shaped the work frontline staff do with those they seek to serve. Since outcome measurement is intended to improve the lives of those served, ultimately those changes must filter down to the frontline encounter (e.g., See Sandfort, 2000). How do existing frameworks, which do not fully capture what frontline staff do, influence this work? Is outcome measurement simply irrelevant for frontline staff, something that they work with or against? And with what consequence? In addition to shifting the focus of outcome measurement from the program to the frontline work, other lines of research would help us better understand the adoption and implementation of outcome measurement in the nonprofit sector. For example, we do not have any systematic data on what tools and resources nonprofits are using to develop their outcome measurement systems nor do we know how they are adapting their systems over time.

Conclusion

This article considered the extent to which existing outcome measurement frameworks are aligned with what nonprofit staff do to ensure positive outcomes for those they serve. The analysis showed that existing outcome measurement frameworks focus primarily on program activities completed and the changes in the users as a result of those program activities. This does not fully capture the work of frontline staff and thus risks mischaracterizing nonprofit performance. The article suggested that understanding nonprofit performance requires reanchoring outcome measurement in a fuller conception of nonprofit work. Drawing attention to the neglected facets of nonprofit work in current outcome measurement frameworks does not detract from their utility in measuring program outcomes. However, by adopting a more robust conception of nonprofit relational work, we might

support more effective encounters between staff and constituents that are essential for achieving outcomes desired by users.

Appendix A

Outcome Measurement Guides and Tools Material Reviewed

Balanced Scorecard (BSC): Step by Step for Governments and Nonprofits by Paul R. Niven, 2003. Hoboken, NJ: John Wiley.

Logic Model Development Guide: Using Logic Models to Bring Together Planning, Evaluation, and Action by the W.K. Kellogg Foundation, 2000, One East Michigan Avenue East, Battle Creek, Michigan 49017-4058, www.wkkf.org.

Outcomes Engineering (OE). OE Toolbox by Barry Kibel, 2000, Pacific Institute for Research and Evaluation, PO Box 444158, Tucson AZ 85733.

Results-Based Accountability (RBA). *Trying Hard is Not Good Enough: How to Produce Measurable Improvements for Customers and Communities* by Mark Friedman, 2005, Victoria, BC, Canada: Trafford.

Results-Oriented Management Accountability (ROMA). 22 Steps to ROMA Implementation: Peeling the Onion, by Jim Masters, January 2000, Center for Community Futures, 6621 Elverton Dr Oakland, CA 94611, www.cencomfut.com/ROMA.rtf.

Success Measures (SM) Guidebook, by Development Leadership Network in Partnership with the McAuley Institute, Version 1 1999. 685 Centre Street Boston MA 02130.

Outcome Funding: A New Approach to Targeted Grantmaking. By Harold Williams, Arthur Y. Webb, and William J. Phillips, 1990, 4th Edition. Rensselaerville, NY: The Rensselaerville Institute (TRI).

United Way of America (UWA). 1996. Outcome Measurement Training Kit. A Practical approach. Alexandria, VA.

Excellence in Nonprofit Leadership and Management Enrichment Series Outcome Measurement II: Practical Data Collection, Analysis and Reporting Methods Produced by The Learning Institute for Nonprofit Organizations., Cosponsored by UWA. Participant Guide: Videotape Learning Module.

Urban Institute Outcome Measurement Training Kit. Key Steps in Outcome Management. 2003. Series on Outcome Management for Nonprofit Organizations. Washington DC: The Urban Institute.

Acknowledgments

The author would like to thank Thomasina Borkman, Laurie Goldman, Leslie Goodyear, Jennifer Greene, Mike Hendricks, all of whom read a version of this article and were generous with their time and suggestions. The errors and omissions remain the sole responsibility of the author. The author would also like to thank the Editor and three anonymous reviewers for their feedback.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported in part by a seed grant from George Mason University.

References

Belenky, M. F., Bond, L. A., & Weinstock, J. S. (1997). *A tradition that has no name: Nurturing the development of people, families and communities*. New York, NY: Basic Books.

- Benjamin, L. M. (2004). Community development financial institutions: Current issues and future prospects. *Journal of Urban Affairs, 26*, 177–195.
- Benjamin, L. M. (2008). Bearing more risk for results: Performance measurement and nonprofit relational work. *Administration & Society, 39*, 959–983.
- Berlin, S. B. (2005). The value of acceptance in social work direct practice: A historical and contemporary view. *Social Service Review, 79*, 482–510.
- Botcheva, L., White, C. R., & Huffman, L. C. (2002). Learning culture and outcomes management practices in community agencies. *American Journal of Evaluation, 23*, 421–434.
- Bozzo, S. L. (2000). Evaluation resources for nonprofit organizations: Usefulness and applicability. *Nonprofit & Management Leadership, 10*, 463–472.
- Brodtkin, E. (2008). Accountability in street level organizations. *International Journal of Public Administration, 31*, 317–336.
- Campbell, D. (2002). Outcomes assessment and the paradox of nonprofit accountability. *Nonprofit Management & Leadership, 12*, 243–258.
- Carman, J. G. (2007). Evaluation practice among community-based agencies: Research into the reality. *American Journal of Evaluation, 28*, 60–75.
- Carman, J. G., & Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. In J. G. Carman & K. A. Fredericks (Eds.), *Nonprofits and evaluation. New directions for evaluation (119)*; pp. 51–72). San Francisco, CA: Jossey-Bass.
- Chen, H. T. (1990). *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Chetkovich, C., & Kunreuther, F. (2006). *From the ground up: Grassroots organizations making social change*. Ithaca, NY: Cornell University Press.
- Christensen, R., & Ebrahim, A. (2006). How does accountability affect mission: The case of a nonprofit serving immigrants and refugees. *Nonprofit Management & Leadership, 17*, 195–209.
- Clary, E. G. (1987). Social support as a unifying concept in voluntary action. *Nonprofit and Voluntary Sector Quarterly, 16*, 58–68.
- Cutt, J., & Murray, V. (2000). *Accountability and effectiveness evaluation in nonprofit organizations*. London, England: Routledge.
- Ebrahim, A. (2005). Accountability myopia: Losing sight of organizational learning. *Nonprofit and Voluntary Sector Quarterly, 34*, 56–87.
- Eckerd, A., & Moulton, S. (2011). Heterogeneous roles and heterogeneous practices: Understanding the adoption and uses of nonprofit performance evaluations. *American Journal of Evaluation, 32*, 98–117.
- Eichler, M. (2007). *Consensus organizing: Building communities of mutual self-interest*. London, England: Sage.
- Evans, S. M., & Boyte, H. C. (1992). *Free spaces: The sources of democratic change in America*. New York, NY: Harper and Rowe.
- Fine, A. H., Thayer, C. E., & Coghlan, A. (2000). Program evaluation practice in the nonprofit sector. *Nonprofit Management & Leadership, 10*, 331–339.
- Fisher, R. (1994). *Let the people decide: Neighborhood organizing in America*. New York, NY: Twayne.
- Frumkin, P. (2001). *Managing for outcomes: Milestone contracting in Oklahoma*. Arlington, VA: The Price WaterhouseCoopers Endowment for the Business of Government.
- Gecan, M. (2004). *Going public: An organizers guide to citizen action*. Boston, MA: Beacon Press.
- Gittell, R. J., & Vidal, A. (1998). *Community organizing: Building social capital as a development strategy*. Thousand Oaks, CA: Sage.
- Gottlieb, N. (1992). Empowerment, political analyses, and services for women. In Y. Hasenfeld (Ed.), *Human services as complex organizations* (pp. 301–319). Newbury Park, CA: Sage.
- Guy, M. E., Newman, M. A., & Mastracci, S. H. (2008). *Emotional labor: Putting the service in public service*. Armonk, NY: M.E. Sharpe.
- Halpern, R. (1995). *Rebuilding the inner city: A history of neighborhood initiatives to address poverty in the United States*. New York, NY: Columbia University Press.

- Handler, J. F. (1992). Dependency and discretion. In Y. Hasenfeld (Ed.), *Human services as complex organizations* (pp. 276–297). Newbury Park, CA: Sage.
- Hasenfeld, Y. (Ed.). (1992a). *Human services as complex organizations*. Newbury Park, CA: Sage.
- Hasenfeld, Y. (1992b). The nature of human service organizations. In Y. Hasenfeld (Ed.), *Human services as complex organizations* (pp. 3–23). Newbury Park, CA: Sage.
- Hasenfeld, Y. (2000). Organizational forms as moral practices: The case of welfare departments. *Social Service Review*, 74, 329–351.
- Hasenfeld, Y. (Ed.). (2010). *Human services as complex organizations* (2nd ed.). Newbury Park, CA: Sage.
- Hendricks, M., Plantz, M. C., & Pritchard, K. J. (2008). Measuring outcomes of United Way -funded programs: Expectations and reality. In J. G. Carman & K. A. Fredericks (Eds.), *Nonprofits and evaluation. New Directions for Evaluation*, 119, 13–35.
- Hochschild, A. (1983). *The managed heart: Commercialization of human feeling*. Berkeley: University of California Press.
- Houchin, S., & Nicholson, H. J. (2002). Holding ourselves accountable: Managing by outcomes in girls incorporated. *Nonprofit and Voluntary Sector Quarterly*, 31, 271–277.
- Hwang, H., & Powell, W. W. (2009). The rationalization of charity: The influences of professionalism in the nonprofit sector. *Administrative Science Quarterly*, 54, 268–298.
- Julian, D. A. (2001). A case study of the implementation of outcomes-based funding within a local United way system: Some implications for practicing community psychology. *American Journal of Community Psychology*, 29, 851–874.
- Karabanow, J. (1999). When caring is not enough: An exploration of emotional labor at a Canadian street kid shelter. *Social Service Review*, 73(3), 340–357.
- Knutsen, W. L., & Brower, R. S. (2010). Managing expressive and instrumental accountabilities in nonprofit and voluntary organizations: A qualitative investigation. *Nonprofit and Voluntary Sector Quarterly*, 39, 588–610.
- Kretzmann, J. P., & McKnight, J. L. (1993). *Building communities from the inside out: A path toward finding and mobilizing a community's assets*. Evanston, IL: Institute for Policy Research, Northwestern University.
- Lann-Wolcott, H., Medvene, L. J., & Williams, K. (2010). Measuring the person-centeredness of caregivers working with nursing home residents with dementia. *Behavior Therapy*, 42, 89–99.
- LeRoux, K., & Wright, N. (2010). Does performance measurement improve strategic decision-making? Findings from a national survey of nonprofit social service agencies. *Nonprofit and Voluntary Sector Quarterly*, 39, 571–558.
- Lipsky, M. (2010). *Street level bureaucracy: Dilemmas of the individual in public service*. New York, NY: Russell Sage Foundation.
- Mansbridge, J., & Morris, A. D. (Eds.). (2001). *Oppositional consciousness: The subjective roots of social protest*. Chicago, IL: University of Chicago Press.
- Morley, E., Vinson, E., & Hatry, H. P. (2001). *Outcome measurement in nonprofit organizations: Current practices and recommendations*. Waldorf, MD: Independent Sector.
- New England Nonprofit Quarterly. (1998). [Special Issue]. Boston, MA: Third Sector New England.
- Noddings, N. (2003). Is teaching a practice? *Journal of Philosophy of Education*, 37, 241–251.
- Poister, T. (2003). *Measuring performance in public and nonprofit organizations*. San Francisco, CA: Jossey-Bass.
- Powell, W. W., & DiMaggio, P. (Eds.). (1991). *The new institutionalism in organizational analysis*. Chicago, IL: University of Chicago Press.
- Pressman, J. L., & Wildavsky, A. B. (1973). *Implementation: how great expectations in Washington are dashed in Oakland: Or, Why it's amazing that Federal programs work at all, this being a saga of the economic development administration as told by two sympathetic observers who seek to build morals on a foundation of ruined hopes*. Berkeley: University of California Berkeley Press.
- Prior, L. (2003). *Using documents in social research*. Thousand Oaks, CA: Sage.

- Quarter, J., & Richmond, B. J. (2001). Accounting for social value in nonprofits and for profits. *Nonprofit Management & Leadership, 12*, 75–85.
- Sandfort, J. R. (2000). Moving beyond discretion and outcomes: Examining public management from the front lines of the welfare system. *Journal of Public Administration Research and Theory, 10*, 729–756.
- Sandfort, J. R. (2003). Exploring the structuration of technology with human service organizations. *Administration & Society, 34*, 605–631.
- Sawhill, J. C., & Williamson, D. (2001). Mission impossible? Measuring success in nonprofit organizations. *Nonprofit Management & Leadership, 11*, 371–386.
- Schwandt, T. A. (2005). The centrality of practice to evaluation. *American Journal of Evaluation, 26*, 95–105.
- Scott, W. R. (1995). *Institutions and organizations*. Thousand Oaks, CA: Sage.
- Shadish, W. R., Cook, T. D., & Leviton, L. C. (1991). *Foundations of program evaluation: Theories of practice*. Newbury Park, CA: Sage.
- Shaw, I., & Faulkner, A. (2006). Practitioner evaluation at work. *American Journal of Evaluation, 27*, 44–63.
- Silverman, R. M. (2005). Caught in the middle: Community development corporations and the conflict between grassroots and instrumental forms of citizen participation. *Journal of Community Development Society, 36*, 35–51.
- Smith, S. R. (2010). Nonprofits and public administration: Reconciling performance management and citizen engagement. *The American Review of Public Administration, 40*(2), 129–152.
- Smock, K. (2004). *Democracy in action: Community organizing and urban change*. New York, NY: Columbia University Press.
- Stall, S., & Stoecker, R. (1998). Community organizing or organizing community? Gender and the crafts of empowerment. *Gender & Society, 12*, 729–756.
- Urban, J. B., & Trochim, W. (2009). The role of evaluation in research-practice integration: Working toward the golden spike. *American Journal of Evaluation, 30*, 538–553.
- VanderPlaat, M. (1995). Beyond technique: Evaluating for empowerment. *Evaluation: International Journal of Theory, Research and Practice, 1*, 81–96.
- Warren, M. R. (2001). *Dry bones rattling: Community building to revitalize American democracy*. Princeton, NJ: Princeton University Press.
- Wholey, J. S., Abramson, M. A., & Bellavita, B. (1986). *Performance and credibility: developing excellence in public and nonprofit organizations*. Lexington, MA: D.C Heath.
- Zimmermann, J. A. M., & Stevens, B. W. (2006). The use of performance measurement in South Carolina nonprofits. *Nonprofit Management & Leadership, 16*, 315–327.