

A GROUNDED MODEL FOR MORE ACCURATE AND USEFUL EVALUATION OF CONTRACTED SOCIAL SERVICES

Christopher Horne, is an Associate Professor of Public Administration at University of Tennessee at Chattanooga as well as an independent program evaluation consultant.



To better understand the increasingly common evaluation context of nonprofit social service programs provided under government contract, 30 in-depth interviews with a broad range of government and nonprofit administrators were analyzed, following the grounded theory approach, to identify factors that affect respondents' program outcomes. Some of the most important factors are specific to contracting and would not be captured in conventional evaluations focused narrowly on programs. The most important of these factors can be categorized as components of either the formal or emergent government-nonprofit relationship. If program evaluations are to contribute to the improvement and accountability of contracted social services, evaluation program models should broaden to include these key factors, which may otherwise be overlooked. This paper presents one such model and accompanying recommendations, grounded in data that administrators and evaluators may use to better pursue program improvement, accountability, and social betterment goals in evaluations of contracted social service programs.

Contracting with nonprofit organizations has become the preferred delivery method for government-funded social services (Austin, 2003; Smith and Lipsky, 1995), and scholars and practitioners have come to recognize that managing such cross-sector contract arrangements requires a different skill set than that required for management within a single hierarchical organization (Brown and Potoski, 2003; McGuire, 2006; Smith, 2008; Stoker, 2006; Van Slyke, 2002; Young, 1999). This paper considers how social service contracting affects program evaluation and reporting specifically. In-depth interviews with a broad range of government and nonprofit social service contract administrators were conducted to explore three interrelated questions: (1) What characteristics of the contracting relationship affect program outputs and outcomes? (2) How would ignoring these characteristics distort program evaluation? (3) How would ignoring these characteristics impede use of program evaluation for program improvement?

Employing a grounded theory approach, analysis of the interviews revealed that many factors that strongly shape program outputs and outcomes are specific to purchase-of-service contracting and would not typically be captured in the conventional program logic models frequently used to design program evaluations. Some of the most important factors are external to the program itself, but are, instead, components of either the formal or emergent government-nonprofit relationship. If program evaluation is to contribute to

improvement of contracted social services, we should broaden program models to include these key factors, which may otherwise be overlooked. The remainder of this paper develops one such model accompanied by a set of recommendations, both grounded in data that may be used to better pursue the goal of improving contracted social service programs through program evaluation.

RESEARCH DESIGN

This research followed a grounded theory approach (Charmaz, 2006; Locke, 2001; Strauss and Corbin, 1990) to develop a model depicting the lessons articulated—explicitly and implicitly—by expert public and nonprofit managers. In-depth interviews, lasting an average of 45 minutes, were conducted with 30 managers, 15 from nonprofit organizations and 15 from government agencies. A purposive network sampling strategy focused the study on lessons learned by practitioners who have developed expertise in managing government-nonprofit contracts through direct experience. Participants were selected based on the authors' personal knowledge of the managers' expertise, their recommendations, and the recommendations of trusted, knowledgeable colleagues.

The first 20 interviews were conducted in the same mid-sized Southern United States city. By first concentrating on one city and its government-nonprofit social services contracting network, the initial findings that emerged reflected the views of government and nonprofit managers on both "sides" of multiple government-nonprofit contracts. During the course of these initial interviews, analysis of data gathered during earlier interviews was used to develop questions

and probes for later interviews relevant to the same implementation context and, often, with reference to the same contract relationships.

Ten additional interviews were conducted with managers from all regions of the United States. This second wave of data collection shifted from exploratory to confirmatory interviews, yielding findings very similar to the first wave, attesting to the robustness of the lessons learned and strengthening confidence in their generalizability outside of the original geographic area.

The interviews pursued two main lines of questioning. Following questions about their background and their organizations, the managers were asked to describe examples of particularly successful and particularly *unsuccessful* government-nonprofit social service contracts. The participants were prompted to explain why they selected their examples, to define success, and to identify factors that contributed to or detracted from successful contract outcomes. Next, the managers were asked to describe what they wish more managers in the "other" sector (public or nonprofit) better understood about their own sector, how their own formal education prepared them to manage government-nonprofit collaboration, and what they believe students of administration ought to learn in preparation for government-nonprofit collaboration. These questions elicited the managers' tacit theories—those theories of practice developed through experience but rarely formally stated—used to explain positive and negative outcomes in government-nonprofit social services contracting. This somewhat "backdoor" approach to learning about program evaluation in the

government-nonprofit contracting context was taken instead of asking about program evaluation more directly to avoid biasing interviewees toward "program-evaluation-by-the-textbook" responses, though follow-up questions were used to steer interviews toward the topic of program evaluation.

The first 15 interviews were open coded, with these codes then organized into a provisional coding scheme. To increase the reliability of the data analysis, a second trained coder duplicated this process, and, after discussing the codes resulting from our individual efforts, we developed a coding scheme that we agreed captured the breadth of meaningful data in the interviews. Employing the grounded theory practice of "constant comparison," this coding scheme was modified as needed to accommodate additional themes emerging from the

remaining transcripts, and the resultant set of codes was applied across the entire dataset by both coders individually and then jointly with discussion to resolve any discrepancies. This final set of codes, with further validation drawn from previous research (Strauss and Corbin, 1990, p. 48 – 55), form the basis of the findings and recommendations presented in the following sections.

FINDINGS

Contract Managers' Expanded Logic Models

Analysis of the interview data culminated in a grounded model of the government-nonprofit contracting relationship, presented in Figure 1.

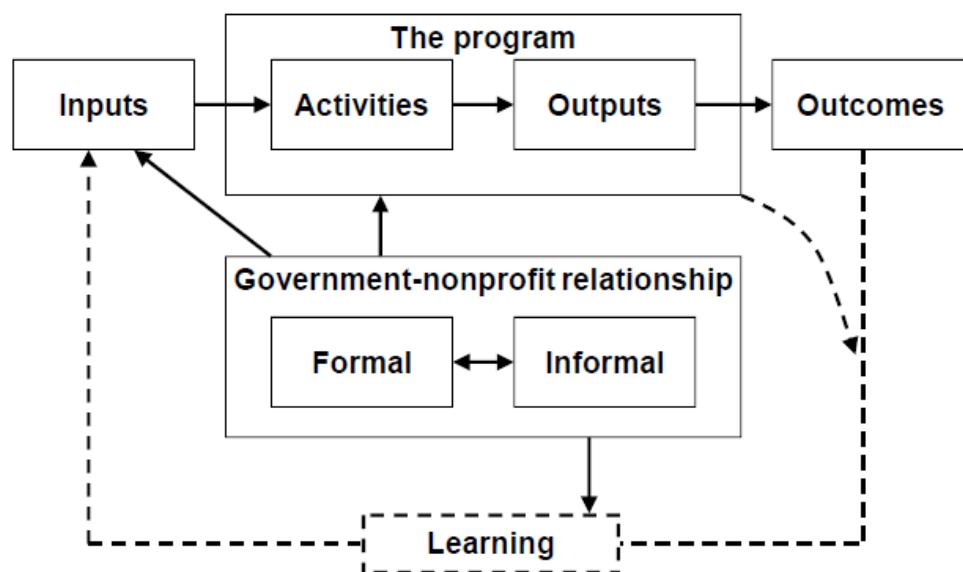


Figure 1: A grounded model of government-nonprofit social service contracting

The upper portion of the model depicts a standard logic model to represent the contracted social service program using the inputs → activities → outputs → outcomes sequence familiar to most social service practitioners and analysts. (This particular formulation of a program logic model has been popularized by United Way.) Program logic models with these components are a common starting point for developing program evaluation plans and specific performance measures (McLaughlin and Jordan, 2004; Poister, 2003; United Way, 1996), and rightly so; a well-developed logic model provides a straightforward guide for systematically identifying key performance measures and the theory of change underlying the program. The key contribution of this paper is placing the standard program logic model in the context of the larger government-nonprofit relationship, depicted by the remainder of the diagram.

Building the model began at an obvious, uncomplicated starting point: The written contract itself and the broader contracting process strongly influence available inputs, the range of permissible activities and outputs, and, in turn, program outcomes. This may be an obvious statement, but the neglect of this tenet in designing program evaluations is a source of frustration, particularly for nonprofit contract managers. In explaining successes and, especially, failures, nonprofit managers are adamant that their contracted programs be viewed in light of the formal structures established both in their contracts and through the contracting process. Three nonprofit managers' examples convey typical frustrations:

... because of the structural changes in the way we get referrals from the state, we went through about a year and a half of not getting any referrals, and ... we were supposed to be serving people referred by the state. ... They were blaming us for not getting enough referrals ... but you guys are the ones that are supposed to give us the referrals. What more can we do?

[Government agencies] keep adding on these little tidbits of things that just make it more difficult to do your job when your purpose is supposed to be helping clients.

[Government funding] has continued to get worse, to decrease and decrease and decrease over time, every year You can only do that for so long without quality being affected.

Performance-based contracts, specifically, were featured in stories of both positive and negative effects of contract arrangements on program outcomes. Some performance-based contracts were structured to reward nonprofits for client outcomes aligned with the organizations' missions while giving the nonprofits significant latitude to tailor services to individual clients and to engage in experimentation in how services were delivered. These contracts allowed the nonprofit organizations to realize their potential strengths, such as innovation and personalized service. As one nonprofit manager described it,

... they are not breathing down our necks about how we do it. We're getting paid based on results. That's the best.

Other performance-based contracts though, stifled innovation by specifying and reimbursing for a narrow range of service delivery methods, a known pitfall in

performance-based contracting (Behn and Kant, 1999). One nonprofit manager offered a poignant example of contract-dictated performance measures working against program goals:

So, if you can move more kids into adoption quicker, you get a little financial reward ... kids get moved too fast ... there's lawsuits pending ... because the kids got moved too early and they were abused further The commissioner of [the state] wants to reduce the number of kids in care, in state custody. We want to keep kids safe. It sounds like the same thing. It's not.

These quotes exemplify a recurring theme of the interviews: attributing program outcomes solely to the program itself is often inaccurate because it ignores the powerful effects of aspects of the government-nonprofit relationship outside the program (see also Ansell and Gash, 2007; Bryson et al., 2006; Stoker, 2006). Such aspects identified by the managers include requirements to "innovate," data collection requirements introduced in the course of contract monitoring, staff training requirements, stagnant funding levels in multi-year contracts, shifting client eligibility rules and changes in characteristics of clients referred by government to nonprofit service providers.

Informal Relationships Help and Hinder the Use of Program Evaluation

A focus on the mechanical details of contracts is notably lacking from the managers' explanations of why contracting succeeds or fails and their advice for future managers. Instead, much of their explanation and advice emphasizes the importance of informal relationships—the patterns of interpersonal interactions

governed by unwritten rules and unofficial roles that emerge in the context of the formal contracting arrangement. The managers commonly described and assessed informal relationships in terms of the degree of open communication and trust. When explaining successful contracting experiences, several managers traced long histories of shared experiences with their cross-sector counterparts, building mutual regard and trust. Participants were adamant about the necessity of positive informal relationships:

Collaboration is about relationship building. It's being open, transparent and communicative in appropriate ways, constantly, continually I ultimately rely on my relationship building to really get through.

A strong theme emerged to suggest *why* informal relationships figure so significantly in participants' stories and advice: Organizational and systemic learning are facilitated by, hindered by and filtered through the informal aspects of government-nonprofit relationships. Fiol and Lyles offer a simple definition of *organizational learning*: "the process of improving actions through better knowledge and understanding" (1985, p. 803). Building this "better knowledge and understanding" is, of course, a primary goal of program evaluation. The concept of *systemic learning* is appended to describe the type of learning that requires sharing of information throughout the collaborative system—here, between the government agency and nonprofit organization—and has the potential to improve the overall collaboration, beyond improvements isolated to a single organization. Based on participants' examples, organizational and systemic learning is conceptualized as

following reflection on the experience of engaging in program activities and producing program outputs and upon observing programs' outcomes (hence the two starting points for the learning feedback loop in Figure 1).

Instances of organizational and systemic learning that led to improvements in the structure of the government-nonprofit relationship inevitably occurred in the context of informal relationships characterized by trust and open communication. One nonprofit manager described the learning that takes place in an informal setting hosted by a government agency:

... they have a luncheon, and they invite this kind of ever-growing group together to share ideas and information and say "hey, we've been doing this," and "hey, here's this training opportunity."

Another nonprofit manager described how having a positive informal relationship with her government counterparts "makes all the difference," again, providing the context for systemic learning:

We had a question about a grant that we are working on right now, and the executive director picked up and dialed the cell phone number for the lady at the state level. And our other programs? There's no way we would have picked up and dialed a cell phone!

Most examples, however, explained the lack of—or even distortion of—organizational and systemic learning as due to distrust and the lack of regular, informal communication. The instances of distrust hindering learning mostly follow the pattern predicted by principal-agent theory, with the agent—here, the nonprofit managers—withholding or

distorting information from the principal—the public manager (Van Slyke, 2006). Two nonprofit managers told similar stories:

We are really, really careful of our relationships with our state contractors. They are, you know, they're our bread and butter. There's a lot of things that we don't talk to them about, such as, do you realize this program ain't doing what it's supposed to be doing We don't want to say that to them. For one thing ... for some of them, it's a very personal thing. We don't have the kind of relationship that would allow us to be that frank

I don't know if we feel empowered to go to [the state agency] and say ... what we're doing here is not working, it's not the best program for the population we are trying to serve, we need to look at something else. But ... it's not in our best interest to rattle that cage, because we have people, good employees that have been working with us for many years ...

In these and several additional examples, the nonprofit managers indicated their distrust of their government counterparts and, frequently, they described their communication with them as limited, strictly formal, or sporadic. In some instances, the only communication occurred through written performance reports and annual contract renewal processes. Perhaps, the public managers in these cases assume that the "invisible hand" of competitive market forces obviates the need for ongoing communication and learning (Behn and Kant, 1999). In any case, in the absence of trusting, open interpersonal relationships across the contract, some nonprofit managers resort to manipulation and distortion of evaluative data. Nonprofit contract managers reported withholding

client cases from "the denominator," strategically assigning cases to client groups retrospectively, referring hard-to-serve clients to other organizations just before performance measures would be registered and other similar tactics in an effort to bolster evaluation results. In every case, the nonprofit managers' defense rested on the lack of trust they had in their government counterparts, whom they suspected of watching for every opportunity to cut program funding, coupled with a commitment to keep their programs running for the sake of their clients and employees. Ironically, the government respondents, with one exception, did not describe program evaluation and reporting as tools for mechanistically allocating and withholding funding, but for program improvement and accountability to the public. By their failure to engender trust among contractors, though, the government contract managers unwittingly introduce systematic distortion and manipulation of performance measures, precluding full use of program evaluation for program improvement.

RECOMMENDATIONS FOR EVALUATING CONTRACTED SOCIAL SERVICE PROGRAMS

The effects of the formal and informal government-nonprofit relationship on the outputs and outcomes of contracted social services programs suggest at least five recommendations for improving program evaluation in that context. The first recommendation is foundational to the recommendations to follow: Developing evaluation plans for government-nonprofit social service contracts should begin with a model of the entire government-nonprofit social service delivery system (such as that presented in Figure 1), not only the more limited model of the program itself. By

expanding the scope of the program model, program evaluations would still capture the typical measures of program activities, outputs, and outcomes, but they would also capture system-level factors that affect program outputs and outcomes. The managers participating in this study might recommend measuring, for example, timeliness in reimbursement for services, changes in the characteristics of referred clients, adherence to contract monitoring plans and cost of program evaluation and reporting itself.

Second, the recognition that output and outcome measures themselves affect how programs operate and what outcomes are pursued suggests that nonprofit managers—those most directly involved in service delivery—should be involved in the development of such measures. A recurring theme in the interviews was that predefined performance measures were measuring the "wrong thing," eliciting one of two responses: Delivery of services considered less than ideal for clients to adhere to the "wrong" measures or delivery of the services considered best for clients followed by "gaming" the data collection and reporting process. All of the nonprofit managers and several of the government managers agree that seeking consensus on output and outcome measures (and, by the same token, consensus on intended outputs and outcomes) at the onset of a contract arrangement would go a long way toward preventing inconsistency between program activities and evaluation reports, paving the way for program evaluation that is actually useful for program improvement and accountability.

Third, evaluators and government and nonprofit contract managers need to clearly

articulate and understand the intended uses of program evaluation. In several instances, nonprofit managers were fearful that weak performance would automatically result in funding cuts while their government counterparts expressed a desire to use evaluation for improving programs delivered in long-term partnerships with nonprofit organizations. Whether or not these intentions reflect reality is uncertain, but making a clear, formal, explicit commitment to how evaluation results would be used would alleviate pressure on nonprofit managers to distort performance reporting and, again, enable the use of program evaluation for program improvement.

These first three recommendations can only be successful if pursued along with the fourth: Government and nonprofit managers must develop positive informal relationships characterized by trust and open communication. One of the clearest findings of this study is that the program evaluation fidelity and the likelihood that program evaluation can lead to program improvement are heavily contingent on having open, trusting relationships among government and nonprofit actors in the service delivery system. Developing such relationships does not require the intervention of a group therapist. Evaluators and contract managers may draw more practical application from another theme recurrent throughout the interviews: Positive informal relationships—those that facilitate consequential organizational and systemic learning—between government and nonprofit managers usually emerge from formal structures of the government-nonprofit relationship outside the contract itself. In describing the value of monthly meetings with the contracting government agency, a

nonprofit manager provided quite an explicit example of this theme:

Well, the structure just allows us to kind of get to meet once a month in an informal setting and bring different issues to the table ... so we actually have a relationship ... with different people representing the government. I would say that's an important factor.

Other public and nonprofit managers described relationships, often in very amiable terms that developed in task forces, board meetings, technical assistance seminars and learning circles. These positive informal relationships, then, did not develop spontaneously in informal settings, but in the context of formal structures built around (or, in some cases, without any reference to) the actual contract. These formal structures provide the time and space necessary for informal relationships to emerge. In turn, emergent, positive informal relationships provide the trust and open communication necessary to identify and address problems with formal collaborative arrangements (as suggested by the double arrows between "formal" and "informal" in Figure 1). Similarly, Ansell and Gash (2007) found opportunities for "face-to-face dialogue" to be necessary preconditions for successful collaboration, stressing their value for "breaking down stereotypes and other barriers to communication" (p. 558). Ring and Van de Ven (1994) and Van Slyke (2006) add that these structured interactions that provide opportunities for building trust must be sustained throughout the duration of the collaborative arrangement. In view of these lessons, government and nonprofit contract managers should seek opportunities to develop lasting, formal structures surrounding the government-nonprofit contract to encourage the emergence of

sustained positive informal relationships across sectors.

Fifth, the lessons learned by contract managers (and their frustrations) point to the need for a fundamental shift in how we understand the role of program evaluation in promoting accountability in government-nonprofit contracting. The dominant mode of program evaluation emphasizes accountability of the *nonprofit organization* to the *government agency*, frequently resulting in an “us-versus-them” mentality, gaming of data collection systems, and managers’ inability to use program evaluation for program improvement (whether they are aware of it or not). Notably, this conceptualization of the accountability function of program evaluation does not typify contract arrangements identified by respondents as most successful for clients. Instead, the accountability function in successful contract arrangements is conceptualized as promoting accountability of the *government-nonprofit social service delivery system* to the *public*. These contract managers spoke of being good stewards of public resources and maintaining a constant focus on program outcomes for clients. The adoption of this perspective—from a narrow focus on the program itself to a broader focus on the government-nonprofit service delivery system—by evaluators and government and nonprofit contract managers seems to be a prerequisite for using program evaluation as a tool for program improvement rather than self-defense.

CONCLUSION

This research represents an effort to identify factors that affect the validity and utility of program evaluation in government-nonprofit social services contracting. That the formal structure of the government-nonprofit service delivery arrangement affects program inputs, activities, outputs and outcomes is not surprising. Neither is it entirely surprising that the quality of the informal relationships among government and nonprofit managers facilitates or hinders meaningful learning and program improvement. What is more surprising is how little these principles appear to have pervaded program evaluation practice. Additional research is needed to determine whether the findings presented here generalize across social services contracting, other contracting domains and other forms of cross-sector collaboration, but these exploratory findings suggest routes for crafting program evaluation plans in contracting arrangements to better promote accountability, program improvement, and, ultimately, the social betterment goals of social service programs.



REFERENCES

- Ansell, C., & Gash, A. (2007). Collaborative governance in theory and practice. *Journal of Public Affairs Research and Theory*, 18, 543-571.
- Austin, M. J. (2003). The changing relationship between nonprofit organizations and public social service agencies in the era of welfare reform. *Nonprofit and Voluntary Sector Quarterly*, 32(1), 97-114.
- Behn, R. D., & Kant, P. A. (1999). Strategies for avoiding the pitfalls of performance contracting. *Public Productivity and Management Review*, 22(4), 470-489.
- Brown, T. L., & Potoski, M. (2003). Contract-management capacity in municipal and county governments. *Public Administration Review*, 63(2), 153-164.
- Bryson, J. M., Crosby, B. C., & Stone, M. M. (2006). The design and implementation of cross-sector collaborations: Propositions from the literature. *Public Administration Review*, 66 (December, special issue), 44-55.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. London: Sage.
- Fiol, C. M., & Lyles, M. A. (1985). Organizational learning. *The Academy of Management Review*, 10(4), 803-813.
- Locke, K. (2001). *Grounded theory in management research*. London: Sage.
- McGuire, M. (2006). Collaborative public management: Assessing what we know and how we know it. *Public Administration Review*, 66 (December, special issue), 33-43.
- McLaughlin, J. A., & Jordan, G. B. (2004). Using logic models. In Wholey, J. S., Hatry, H. P., and Newcomer, K. E. (Eds.), *Handbook of practical program evaluation*, pp. 7-32. San Francisco: Jossey-Bass.
- Poister, T. H. (2003). *Measuring performance in public and nonprofit organizations*. San Francisco: Jossey-Bass.
- Ring, P. S., & Van de Ven, A. H. (1994). Developmental processes of cooperative interorganizational relationships. *Academy of Management Review*, 19(1), 90-118.
- Smith, S. R. (2008). The increased complexity of public services: Curricular implications for schools of public affairs. *Journal of Public Affairs Education*, 14(2), 115-128.
- Smith, S. R., and Lipsky, M. (1995). *Nonprofits for hire: The welfare state in the age of contracting*. Cambridge: Harvard University Press.
- Stoker, G. (2006). Public value management: A new narrative for networked governance? *American Review of Public Administration*, 36(1), 41-57.
- Strauss, A. & Corbin, J. (1990). *Basics of qualitative research*. Newbury Park, California: Sage.
- United Way. (1996). *Measuring program outcomes: A practical approach*. Alexandria, Virginia: United Way.
- Van Slyke, D. M. (2002). The public management challenges of contracting with nonprofits for social services. *International Journal of Public Administration*, 25(4), 489-517.
- Van Slyke, D. M. (2006). Agents or stewards: Using theory to understand the government-nonprofit social service contracting relationship. *Journal of Public Administration Research and Theory*, 17, 157-187.
- Young, D. R. (1999). Nonprofit management studies in the United States: Current developments and future prospects. *Journal of Public Affairs Education*, 5(1), 13-23.