



AENSI Journals

Advances in Environmental Biology

ISSN-1995-0756 EISSN-1998-1066

Journal home page: <http://www.aensiweb.com/AEB/>



Survey in Concept, Challenges and Evaluation of Organizational Performance

Ali Raeespoor, Mehri Tavakoli, Hojjat Darabi Arkhlo, Ghalandar Hassanpur, Forough Sadr and Maryam Rostami

Faculty of Human Science, Department of Management, Yasouj Branch, Islamic Azad University, Yasouj, Iran.

ARTICLE INFO

Article history:

Received 12 October 2014

Received in revised form 26 December 2014

Accepted 1 January 2015

Available online 17 February 2015

Keywords:

ABSTRACT

This paper reviews the underlying theoretical bases for the evaluation of organizational performance. It then examines representative samples of empirical research into actual evaluation practices in a variety of nonprofits. Some of the most popular tools and systems for evaluation currently recommended by consultants and others are then reviewed. Looking at this prescriptive literature, it is shown that, by and large, it takes little account of the findings of empirical research and, as a result, its approaches may often prove ineffective. An alternative that attempts to integrate the research findings with practical tools that have value for practitioners is then suggested.

© 2015 AENSI Publisher All rights reserved.

To Cite This Article: Ali Raeespoor, Mehri Tavakoli, Hojjat Darabi Arkhlo, Ghalandar Hassanpur, Forough Sadr and Maryam Rostami., Survey in Concept, Challenges and Evaluation of Organizational Performance. *Adv. Environ. Biol.*, 9(2), 803-810, 2015

INTRODUCTION

It is a perplexing, but not uncommon, phenomenon in the world of nonprofit organization studies how little connection there is between the work of those who offer advice on how organizations in this sector might become more effective and that of those who carry out formally designed empirical research into how these organizations actually behave. Nowhere is this gap between “how to” and “what is” more apparent than in the field of performance assessment and evaluation.

In this field, it appears that three phenomena have been proceeding apace over the past 10 or more years:

1. There has been a steady growth in the rhetoric of what might be called the “accountability movement”—those who urge more and better accountability by organizations (business, government and nonprofit) to their “communities” or “stakeholders”. A major part of accountability is demonstrating that the organization has had a positive impact on those they seek to serve in a cost-efficient manner. Thus we have seen the emergence of “sector-bashing” books like that of Walter Stewart on the one hand and, on the other, positively oriented responses such as the Panel on Accountability and Governance in the Voluntary Sector. Both end up calling for more and better evaluation of nonprofit organizational performance. Similarly, there has been a growth of “watchdog” organizations purporting to offer information to potential charitable donors about inefficient nonprofits who waste money. (See for example, the Wise Giving Alliance, the American Institute for Philanthropy, The Charities Review Council of Minnesota and The Canadian Council of Christian Charities).
2. In response to the rhetoric, there has been a steady growth in the number of tools and “how to do it” books and articles aimed at improving evaluation. These are discussed below.
3. Finally, in a seemingly parallel universe, a relatively small number of researchers have set themselves the task of finding out about how much and what kind of evaluation has actually been going on, what happens during this evaluation process and what is done with the information once it is obtained.

This paper will first briefly describe some of the empirical research, then some of the prescriptive tools. The latter will be discussed in terms of the findings from the research so the gaps and contradictions between the two are revealed. In conclusion, the implications of this analysis for both prescriptive and research-oriented approaches to the field will be noted.

Corresponding Author: Ali Raeespoor, Faculty of Human Science, Department of Management, Yasouj Branch, Islamic Azad University, Yasouj, Iran.
E-mail: ALIRAESI106@Yahoo.Com

What Is Organizational Performance Evaluation?:

Evaluation is the process of gathering information on the results of past activities for the purpose of making decisions about them. Organizational Performance Evaluation occurs when this process is applied to assessing the state of the organization as whole. Typically, this refers to how well it is achieving its stated mission and involves looking at its goal attainment as well as how efficiently it has performed. (This is different from Program Evaluation in which the focus is on one specific part of the organization.)

Evaluation can occur in a formal, systematic way through the application of a professionally designed evaluation program or it can be carried out with varying degrees of informality ranging from gathering a few reports to completely impressionistic estimates about how things have been going.

Recent Research Into Real-Life Evaluation Experiences:

It would be a pleasure to report that, in the past 10 years, a number of large-scale studies of nonprofit organizational evaluation activities have been carried out. Studies which would provide data from many different kinds of nonprofits with a wide range of missions that would tell us how their performance was measured, what responses occurred during and after these assessments and, most importantly, the nature and extent of any improvements that occurred as a result of these evaluations. Unfortunately, these kinds of large-sample, comparative, longitudinal studies do not exist as yet, though one recently completed study comes close (see below).

What has appeared instead is a series of smaller studies, usually of one or a few organizations, often based on case study methodology. This makes it difficult to generalize across the whole nonprofit sector but, when most report similar findings, it is tempting to start drawing some tentative conclusions about what is likely to happen when organizational evaluation occurs under various circumstances.

The one large-scale study of evaluation practices was carried out in Canada in 2002 [14]. This research was based on a stratified random sample of 1,965 voluntary sector organizations of all types and sizes from across the country. In addition, 322 funders of these organizations were surveyed. They included various government bodies, federated funding organizations, private foundations and others. This conference will be hearing about this research in detail hence it will not be discussed at length here. Suffice it to say that, among other things, it reported that, though many nonprofits claim they carry out evaluation activities, only 18% attempt to assess the performance of the organization as a whole. Of the evaluations that are carried out, much of it fails to focus on outcome measurement or mistake output measures for outcome measures. Most funders would like to see evaluation carried out on the outcomes of their grants but over half provide no support to their fundees to help them carry it out and over half admit they don't make much use of the evaluation information they do receive. Finally, most nonprofits feel they don't have the time, staff or expertise to carry out evaluations well.

Turning now to the more in-depth case studies of the past 10 years, they can be roughly categorized in terms of those which have tried to follow the entire evaluation process through all its stages from conception to implementation, those that have focused primarily on the design and results-interpretation phase and those that have focused primarily on the results-application phase.

Overall process studies:

When researchers attempt to observe the behavior of evaluators, evaluatees and other interested stakeholders over time, the results all seem to support the "social constructionist" model advanced by Herman and Renz, [15]. This theory posits that evaluation is an essentially subjective process that requires the parties to negotiate an interpretation of every facet of it.

Time and again in this research, we see the parties bringing their own attitudes, perceptions, values and agendas to the process and engaging in some form of negotiation of their differences at each stage of it. For example, Herman and Renz [16] conducted one of the more sophisticated studies in that it looked at 44 nonprofit organizations in a single community over two points in time (1993 and 1999). At both times they found that key stakeholders such as funders or client groups all had opinions as to how the studied nonprofits were doing but all had differing criteria for judging and utilized different "data" to make up their minds about this.

In a similar vein, Cutt and Murray reported a series of case studies of 8 nonprofit organizations in two Canadian cities [7,35,8]. They too focused on the relationship between funders and fundees. While both the funders and nonprofit managers subscribed wholeheartedly to the rhetoric of evaluation—that there should be more of it, that it should be objective and emphasize outcomes—in fact they behaved quite differently. Both funders and fundees attempted to influence one another's behaviour in many informal ways outside of the evaluation process. In the end, funders came to definite conclusions about the performance of the organizations they funded but it was heavily influenced by their pre-existing values and the organization's informal reputation in the funder's broader information network. And, in spite of these opinions, their eventual decisions about whether to increase, decrease or terminate funding to these organizations was scarcely influenced at all by their evaluations. Other matters such as economic conditions and political pressures to favor one set of social issues

over another proved much more important. Similar conclusions to these were reached by Forbes [12] and Scotch [32].

Studies of evaluation design and interpretation processes:

Several case studies have looked in detail at the questions of what will be evaluated, how it will be done and how the results will be interpreted. For example, Campbell [4] studied 8 local economic development projects in Northern California with an emphasis on the negotiations over evaluation criteria between funders and project leaders. Lindenberg [21] reported on a detailed case study of the efforts of the head of CARE (the international relief and development organization) to implement a variety of modern management practices drawn from business. Among them was the practice of benchmarking. Ebrahim [9] looked at efforts by funders of two major NGOs in India to control the information coming to them and the resistance they encountered. Ospina *et al.* [28] document even more complex patterns of negotiation between organization leaders, funders and clients in four successful Latino nonprofits in the New York area.

Finally, in a most important book, Paton [30] describes a wide range of evaluation practices in some 27 “social enterprises” primarily in Britain. Among them are 10 organizations required to report Administrative Cost to Expenditure (ACE) ratios to external evaluation bodies, 5 organizations that implemented various forms of Total Quality Management systems, 4 that subjected themselves to external performance audits such as ISO 9000, 3 that attempted to implement outcome measurement systems and one that conducted a social audit to assess its impact on the community it served.

The conclusions from all of these studies are neatly summarized by the following quotation from Paton [30]:

“The limitations, difficulties and pitfalls associated with the various forms of performance measurement are very clear, if hardly a great surprise. Thus, it is probable that, if taken literally, outcome measurement will be impracticable for many social enterprises. More generally, the features that managers hope to find in measurement systems—such as both focus and comprehensiveness, or reliable validity and non-intrusive simplicity—are incompatible and so cannot be realized simultaneously. Moreover, for both internal and external reasons, ‘measurement churn’ seems increasingly to be a fact of life in social enterprises, as it is elsewhere. So the stability on which much of the logic of measurement depends is unlikely to be realized” (p. 160).

Studies of the use of evaluation data:

When evaluators do obtain information on organizational performance by whatever means, there is the question of what they will do with it—the extent to which it will be considered in making decisions about the future. In theory, formal evaluations of what has worked and how efficient a nonprofit organization has been *should* play a prominent role in the making of these kinds of major decisions. The question is: do they?

Research into this question remains skimpy but what there is suggests that it often has only a minimal influence at best, except when the evaluation was carried out as part of a special investigation of a crisis situation. Gebelmann, *et al.* [13], Cutt and Murray [7], Holland [17] and Miller [24] all report studies revealing that many boards of directors in particular are prone to: (a) not proactively seeking more and better evaluation systems; and (b) ignoring or willfully misinterpreting evaluation information that presents “bad news” until a crisis arises. The same tendencies are probable among CEOs though, surprisingly, there is much less research available on them. A similar pattern also prevails when looking at funders as revealed in the work of Hall *et al.* [14], Cutt and Murray [8] and Lohman [22].

This completes our brief review of the past 10 years of research into “what actually happens” in organizational evaluation. As noted earlier, the bottom line is that the results tend only to confirm the conclusion that it is a subjective, “political” process involving a negotiated interpretation of reality by all interested parties.

In spite of the less than sterling performance of evaluators and evaluatees in real life, new evaluation tools are always being developed and being welcomed by nonprofit managers. Therefore, the next section of this paper looks briefly at some of the better-known evaluation tools that have achieved prominence over the past 10 years.

Recent Tools For Improving Organizational Evaluation:

There are too many evaluation tools that have been developed in the past 10 years to be able to report on all of them here (for one recent attempt to do this see Bozzo and Hall, [2]; Cutt and Murray, [8] offer a similar review though of fewer tools). What is done instead is to select a few of the best known and highest potential developments as a sample of what is available.

Program Outcomes: The United Way Approach (www.unitedway.org/outcomes):

This evaluation system focuses exclusively on the identification and measurement of program outcomes for United Way funded agencies in the U.S. While the system starts by evaluating results at the program level, these are supposed to be aggregated at the organizational level so as to report on their overall effectiveness.

Implementation of the outcome measurement system is divided into six stages:

- Building agency commitment and clarifying expectations;
- Building agency capacity to measure outcomes;
- Identifying outcomes, indicators and data collection methods;
- Collecting and analyzing outcome data; improving the outcome measurement system;
- Using and communicating outcome information.

Rather than advocate one particular way to develop outcomes or to collect outcome data, the United Way uses a checklist approach to encourage agencies to think more broadly and critically about its measurement processes. Recent research reported on the United Way website suggests that there is general satisfaction with the six-stage process it recommends and its approach to developing an outcome measurement system. Of all those reviewed, this tool is also the most sensitive to the importance of the implementation process.

The Balanced Scorecard. (www.balancedscorecard.org):

This is a multi-attribute system for conceptualizing and measuring performance designed originally for business organizations and currently being adapted for non-profit organizations [18,27]. In its original form, it assumes that the primary goal of a business is long-run profit maximization. It argues that this will be achieved through a “balanced scorecard of performance attributes” grouped around four “perspectives”:

- The Financial Perspective measuring various financial performance indicators of primary interest to shareholders;
- The Customer Perspective comprising measures of customer satisfaction;
- The Internal Business Perspective which measures internal efficiency and quality; and
- The Innovation and Learning Perspective which attempts to measure the organization’s ability to adapt to changes required by a changing environment.

In the case of nonprofit organizations, their mission statement, rather than the profit statement, becomes the endpoint to be reached through these perspectives. The process starts with defining what that is and identifying outcome indicators that will reveal the extent to which it is being achieved. “Customers” must be replaced by “clients” or “users” of the organization’s services and the “financial perspective” becomes that of the funders or potential funders [27].

CCAF/FCVI Framework for Performance Reporting (www.ccaf-fcvi.com):

This is the most significant effort at tackling the “value-for-money” issue in both the public sector and the non-profit sector. It puts forward 12 “attributes of effectiveness” suggesting that organizations can be audited in terms of how well they manifest these attributes. It is similar in many ways to “The Balanced Scorecard”, however the details of implementation are more thorough.

The system involves developing indicators for the following 12 attributes of effectiveness:

1. *Management Direction*: The extent to which programmatic objectives are clearly stated and understood;
2. *Relevance*: The extent to which the organization or program continues to make sense with respect to the problems or conditions to which it was intended to respond;
3. *Appropriateness*: The extent to which the design of the organization or program and the level of effort are logical in relation to their objectives;
4. *Achievement of intended results*: the extent to which the goals and objectives have been achieved;
5. *Acceptance*: The extent to which the stakeholders for whom the organization or program is designed judge it to be satisfactory;
6. *Secondary Impacts*: The extent to which significant consequences, either intended or unintended and either positive or negative have occurred;
7. *Costs and Productivity*: The relationship between costs, inputs and outputs;
8. *Responsiveness*: The capacity of the program or organization to adapt to changes in such factors as markets, competition, available funding and technology;
9. *Financial Results*: Accounting for revenues and expenditures and for assets and liabilities;
10. *Working Environment*: The extent to which the organization or program provides an appropriate work environment for staff, and staff have the information, capacities and disposition to serve the objectives;
11. *Protection of Assets*: The extent to which the various assets entrusted to the organization or program (physical, technological, financial and human) are safeguarded; and
12. *Monitoring and Reporting*: The extent to which key matters pertaining to performance and organizational or program strength are identified, reported and monitored.

Another approach related to the Balanced Scorecard model is that presented recently by Prof. Rob Paton of the Open University in Britain (Paton, 2003). Called “The Dashboard for Social Enterprises”, it is more specifically designed for nonprofit organizations. The Dashboard is based on asking two sets of questions about the organization’s activities: “Do they work?” and “Are they well run?” These questions are then asked in two contexts, the short-term, operational, context and the longer-term, strategic, context.

Best Practice Benchmarking:

One of the most frequently cited tools of the past 10 years is the application of benchmarking to the evaluation of nonprofit organizations. Benchmarking is a system that compares the organization's practices with those of others doing similar things but who are deemed to be doing it better. It is usually applied to specific programs or functions of the organization so is not, strictly speaking, a tool for evaluating the organizations as a whole. However, it is assumed by those that advocate it that a thorough program of benchmarking will "roll up" to provide a good indicator of how well the organization is doing overall.

Paton [30] is one of the few who have actually carried out research into how well benchmarking works for nonprofits in practice. He found that, though there was considerable enthusiasm for the idea of benchmarking, it was not actually implemented very frequently (in Britain, at least). There is also a problem with what Paton calls "measurement churn"—the tendency to frequently change measurement tools so the same data is rarely gathered over long periods. This makes it difficult to compare performance with others over time. A final difficulty with "best practice" comparisons is that there is no way of knowing if they are the reason for other organizations being more successful. It is possible that practices which "work" for one organization may not work for one's own due to unique situational characteristics of history, culture, personalities, economic conditions and so on.

Charities Rating Services:

As noted in the introduction to this paper, there are several evaluation systems developed to help funders and the public decide how effective various charities are. All of these systems offer generic "standards" that will reveal how well a nonprofit organization is managed. For example, they may report on various aspects of a charity's finances, public availability of audit reports, how fund raising is conducted and the presence of certain policies for its board of directors such as conflict of interest policies.

The problem with these rating services is that, in trying to compare organizations with so many different missions, the standards make no reference to outcomes. They also do not attempt to explicitly reveal the implicit logic models on which they are built. Instead, they assume that organizations that meet their process standards will be more likely to be effective in achieving their goals. Regrettably, little or no published research supports these assumptions though it may be that those who most severely violate the standards they measure are less likely to be effective. Whether those that best meet or surpass these standards are comparatively more effective, however, is unknown [26].

Another problem with all of the evaluation tools except the United Way approach is that very little attention is paid to how the evaluation system is to be implemented. As previously noted, a great deal of research has concluded that, unless all those to be affected by an evaluation system have a strong voice in its design and accept the final product, there is a high probability that the system will fail [7,23,33].

To summarize, it appears that there is still a long way to go before there will be available an evaluation system that can be applied by most nonprofit organizations to reveal valid pictures of how well the organization is performing. Some would argue there is no point in trying, yet decisions are made every day based on untested assumptions and idiosyncratic perceptions of performance. Therefore, the goal of trying to improve the dialogue around this process in a way that takes account of the research into the everyday reality of organizational performance evaluation processes discussed in the first part of this paper is worthwhile pursuing.

*Conclusion:**How Can Nonprofit Organization Leaders Make Practical Use of the Tension Between the Results of the Empirical Research and the Promises of the New Tools?:*

It has been argued here that the inherent subjectivity of the evaluation process makes it almost inevitable that differences will arise at some point between evaluators and evaluatees. Many of the political games that ensue can be destructive in that they distort the information produced and/or result in it being ignored or misused in the way it is applied to future decision-making. They can also create major motivational problems among the evaluatees that can, paradoxically, damage their productivity.

Since one cannot avoid the political dimension in evaluation, attention has to be focused on what can be done to make the differences among the parties constructive rather than destructive. Conflict is not inevitably bad. Indeed, when handled constructively, it can often result in a product that is better than if there were no conflict at all.

In an ideal world, each phase of the evaluation process—its design, implementation, interpretation and application—would be characterized by open discussions among all interested parties during which differences would be aired and resolved by mutual consensus. Unspoken beliefs and assumptions would be made explicit in logic models and standards. This could lead to greater understanding, if not acceptance, of these positions. What has to be done to move toward this ideal of constructive conflict resolution?

Trust building:

The basic secret for creating openness is the creation of an atmosphere of trust among the interested parties. If one of them believes the other is concerned only with their own interests, there is little chance of avoiding destructive game playing. The only approach then becomes one of trying to win the games more often than the opponents.

How is this trust built? Unfortunately for those seeking a “quick fix”, it is usually built over time through many encounters between the parties. During these interactions each must show concern for the other, each must commit fully to communicating reasons for their actions, each allows the other to have a voice in decisions that affect them and each keeps their word when actions are promised or provides explanations if they can't. Paradoxically, though trust takes time to build, it can be destroyed in an instant with only one or two violations of the above rules.

If a prior relationship of trust does not exist before evaluation begins, it must consciously be worked on as the process is developed. This means involving all interested parties in that process, particularly those who are to be evaluated. All must have a voice in deciding the following six questions:

1. What is the purpose of the evaluation?
2. What should be measured?
3. What should be the evaluation methods used?
4. What standards or criteria should be applied to the analysis of the information obtained?
5. How should the data be interpreted?
6. How will the evaluation be used?

Is there a time when this kind of consultation is *not* advisable? Probably the only occasion is when there is strong evidence of malfeasance or willful ineptitude among the evaluatees. This would suggest that there is a high probability that they will consciously suppress or distort information. In such cases, external evaluation by professionals trained in looking for reporting errors, (e.g. forensic accountants) would have to be used.

Logic model building:

Many evaluations founder in destructive politics simply because those designing them fail to articulate the logic models which underlie the system. This failure allows all parties to unconsciously apply their own answers to the questions above which create differences that are not confronted. The time to develop logic models is during the design phase of the evaluation process. For a more detailed discussion of logic models, see Murray (2002).

Some practical problems in certain relationships:

The general guidelines of trust building, participant involvement and logic model creation will help create constructive political climates in evaluation. However, they are more or less difficult to apply depending on who are the evaluators and who the evaluatees. The most straightforward situation is where the evaluator is the management of the organization (or some part of it) and the evaluatees are those running a particular program, function or department within it. Evaluation in this situation is known as “internal evaluation” [33]. The decisions about the evaluation process are under the control of the management who, if they so desire, can ensure that consultation with evaluatees occurs and logic models are thought through before the process starts.

But what if the evaluator is the organization's board of directors? This group is legally responsible for seeing that the organization achieves its mission so the management team is accountable to the board. As the earlier discussion of empirical research shows, however, many boards make little time for, and feel they have no skill in, carrying out evaluations. Many also do not know what to evaluate and feel that probing too deeply into how the organization is doing would suggest that they do not trust the Executive Director. As a result, many boards fail badly in their fiduciary duty to hold management accountable for its actions. Special effort must be devoted to training boards in these duties and providing them with the expertise they need to develop evaluation systems that meet their information needs [8]. However, if the board wants to minimize political game playing with management, the guidelines presented here suggest that the design process should be consultative. In most situations, the best approach would be a board task force on evaluation working together with staff representatives and a professional evaluator (if possible).

An even more difficult relationship exists when the evaluator is an external funder or contractor (e.g. United Way, community foundation, government department). Usually funders or contractors have *many* organizations as their clients. How can they follow the recommended participative model in developing systems for evaluating the performance of these clients? They often express the wish for more and better accountability reporting from them but the participative approach takes time and involves intense face-to-face interactions that are difficult to arrange when there are so many clients. Because of this, as the review of the empirical literature shows, many of these external organizations go to one of two extremes. Some require little or no evaluation other than regular budget and financial reports on how money was spent. Others arbitrarily impose requirements for complex and numerous progress reports as well as final reports. Both are usually unsatisfactory.

In practical terms, the only way to overcome this problem is, whenever a funding arrangement is made, the discussion of it should include a detailed look at the evaluation system that will be used in reporting results. The recipient organization needs to know what kind of information the funder wants and, if, for example, it wants only statistics that obscure important outcomes that can only be assessed qualitatively, the fundee should have an opportunity to influence that decision. It also should have the opportunity to raise beforehand the question of how reported information will be interpreted or used. For example, if measurements reveal low levels of participation or higher than expected costs, will there be an opportunity to conduct research into the reasons for this?

To summarize, organizational evaluation will always contain a subjective element that will lead to differences among those involved. It is only when these differences can be confronted and talked through in a non-threatening, trusting environment that we can avoid destructive game playing and realize the benefits that can be obtained from a well designed evaluation process.

REFERENCES

- [1] American Institute for Philanthropy, 2003. <http://www.charitywatch.org>.
- [2] Bozzo, S.L. and M.H. Hall, 1999. *A Review of Evaluation Resources for Nonprofit Organizations*. Toronto: Canadian Centre for Philanthropy.
- [3] Bernstein, David, 1980. "Comment on Perrin's 'Effective Use and Misuse of Performance Measurement'", *American Journal of Evaluation*, 20(1): 85-93.
- [4] Campbell, David, 2002. "Outcomes Assessment and the Paradox of Nonprofit Accountability." *Nonprofit Management and Leadership*, 12(3): 243-260.
- [5] Canadian Council of Christian Charities, 2003, <http://www.cccc.org>.
- [6] Charities Review Council of Minnesota, 2003. <http://www.crcmn.org>.
- [7] Cutt, James and Murray, 2000. *Vic. Accountability and Performance Evaluation in Non-Profit Organization*. London: Routledge.
- [8] Cutt, James, *et al.*, 1996. "Nonprofits Accommodate the Information Demands of Public and Private Funders." *Nonprofit Management and Leadership*, 7(1): 45-68.
- [9] Ebrahim, Alnoor, 2002. "Information Struggles: The Role of Information in the Reproduction of NCO-Funder Relationships." *Nonprofit and Voluntary Sector Quarterly*, 31(1): 84-114.
- [10] Feller, Irwin, 2002. "Performance Measurement Redux". *American Journal of Evaluation*, 23(4): 435-452.
- [11] Fine, A., C. Thayer and A. Coglán, 1998. *Program Evaluation in the Nonprofit Sector*. Washington: D.C.: The Aspen Institute.
- [12] Forbes, D.P., 1998. "Measuring the Unmeasurable." *Nonprofit and Voluntary Sector Quarterly*, 27(2): 183-202.
- [13] Gebelman, M., S.R. Gilman and D. Pollack, 1997. "The Credibility of Nonprofit Boards." *Administration in Social Work*, 21(2): 21-40.
- [14] Hall, Michael H., Phillips, D. Susan, *et al.*, 2003. *Assessing Performance: Evaluation Practices and Perspectives in Canada's Voluntary Sector*. Ottawa: Voluntary Sector Evaluation Project, Centre for Voluntary Sector Research and Development, Carleton University.
- [15] Herman, Robert D. and Renz, O. David, 2003. "Doing Things Right and Effectiveness in Local Nonprofit Organizations." *Public Administration Review* (in press).
- [16] Herman, Robert D. and Renz, O. David, 1999. "Multiple Constituencies and the Social Construction of Nonprofit Organization Effectiveness." *Nonprofit and Voluntary Sector Quarterly*, 26(2): 185-206.
- [17] Holland, Thomas P., 2002. "Board Accountabilities: Lessons From the Field." *Nonprofit Management and Leadership*, 12(4): 409-428.
- [18] Kaplan, Robert S., 2001. "Strategic Performance Measurement and Management in Nonprofit Organizations." *Nonprofit Management and Leadership*, 11(3): 353-370.
- [19] Letts, C., W.P. Ryan, *et al.*, 1999. *High Performance Nonprofit Organizations: Managing Upstream for Greater Impact*. New York: John Wiley.
- [20] Light, P., 2000. *Making Nonprofits Work*. Washington, D.C.: The Aspen Institute and Brookings Institution Press.
- [21] Lindenberg, Marc, 2001. "Are We at the Cutting Edge or the Blunt Edge: Improving NGO Organizational Performance With Private and Public Sector Strategic Management Frameworks." *Nonprofit Management and Leadership*, 11(3): 247-270.
- [22] Lohman, R.A., 1999. "Has the Time Come to Reevaluate Evaluation?" *Nonprofit Management and Leadership*, 10(1): 93-101.
- [23] Mark, M.M., G.T. Henry and G. Julnes, 2000. *Evaluation: An Integrated Framework for Understanding, Guiding and Improving Policies and Programs*. San Francisco: Jossey-Bass.

- [24] Miller, Judith, 2002. "The Board as a Monitor of Organizational Activity." *Nonprofit Management and Leadership*, 12(4): 429-450.
- [25] Murray, Vic, 2001. "The State of Evaluation Tools and Systems for Nonprofit Organizations." *Accountability: A Challenge for Charities and Fundraisers (New Directions for Philanthropic Fundraising)*, 31: 39-50.
- [26] Murray, Vic, 2002. *Evaluation Games The Political Dimension In Evaluation And Accountability Relationships*. Voluntary Sector Evaluation Project online article: <http://www.vserp.ca/>.
- [27] Niven, Paul R., 2003. *Balanced Scorecard Step-by-Step for Government and Nonprofit Organizations*. Toronto: John Wiley and Sons.
- [28] Ospina, Sonia, Diaz, William and O'Sullivan, F. James, 2002. "Negotiating Accountability: Managerial Lessons From Identity-Based Nonprofit Organizations." *Nonprofit and Voluntary Sector Quarterly*, 31(1): 5-31.
- [29] Panel on Accountability and Governance in the Voluntary Sector (PAGVS), 1999. *Building on Strength: Improving Governance and Accountability in Canada's Voluntary Sector*. Ottawa: Voluntary Sector Roundtable.
- [30] Paton, Rob, 2003. *Managing and Measuring Social Enterprises*. London: Sage Publications.
- [31] Perrin, Burt, 1998. "Effective Use and Misuse of Performance Measurement". *American Journal of Evaluation*, 19(3): 167-179.
- [32] Scotch, R.K., 1998. "Ceremonies of Program Evaluation." *Conference Proceedings*. Association for Research on Nonprofit Organizations and Voluntary Action.
- [33] Sonnichsen, R.C., 2000. *High Impact Internal Evaluation*. Thousand Oaks, CA: Sage.
- [34] Stewart, Walter, 1996. *The Charity Game: Greed, Waste and Fraud in Canada's Charities*. Toronto: Douglas and McIntyre.
- [35] Tassie, B., V. Murray, *et al.*, 1996. "Rationality and Politics: What Really Goes On When Funders Evaluate the Performance of Fundees?" *Nonprofit and Voluntary Sector Quarterly*, 25(3): 347-363.
- [36] The Balanced Scorecard Institute, *The Balanced Scorecard* www.balancedscorecard.org.
- [37] United Way of America, 1996. *Measuring Program Outcomes: A Practical Approach*. Alexandria, Va.: United Way.
- [38] Wise Giving Alliance, 2003. <http://www.give.org>.