Theses on Nonprofit Organization Effectiveness

Robert D. Herman and David O. Renz

Introduction

Contemporary concerns about nonprofit organization (NPO) accountability, outcomes assessment and performance evaluation attest to the importance of taking stock of what we know about NPO effectiveness and the evaluation of NPO effectiveness. Our intention is not to review all the available studies of NPO effectiveness to arrive at set of well-founded conclusions. As recent reviews by Stone and Cutcher-Gershenfeld (forthcoming) and Forbes (1998) demonstrate the research does not permit such a project. Both reviews find that NPO effectiveness studies are characterized by varying theoretical perspectives and varying research objectives, which make accumulation and integration impossible. Forbes (1998) detects a change over of the last two decades from studies that attempt to assess effectiveness of goal attainment toward studies of the social and political processes of effectiveness assessments.

In this paper we draw on the general literature on organizational effectiveness and on the small literature focusing on NPO effectiveness to present theses about NPO effectiveness. Some of the theses we judge substantially supported; others have more limited support. In some cases the support for a thesis is both empirical and theoretical. In other cases the support may be mostly theoretical or conceptual. Our goal is not to descriptively review the literature, rather it is to distill from it assertions and conclusions that help us understand what we know about NPO effectiveness.

Obviously one of the great challenges of studying nonprofit or nongovernmental organizations is their overwhelming variety. The sort of NPOs we explicitly consider in this paper are those that in the United States are defined as publicly-supported charities that is, organizations which provide services or benefits to some part of the larger public (not just members) and which typically receive some financial support from donations. While we are not convinced the effectiveness issues are any different for all-volunteer public-benefit charities than they are for those that operate with paid staff, we have in mind paid staff NPOs. In the interest of brevity when we refer to
NPOs in the remainder of the paper rather than public-benefit charitable organizations.

We distinguish between organizational effectiveness and effectiveness of sub-units or functions. That is, we consider that an organization is not the sum of its parts or functions. It is possible to assess the effectiveness of a program or to assess the effectiveness of the use of service volunteers in an organization, but such assessments do not necessarily indicate much about overall organizational effectiveness. However, the extent to which actual stakeholders of nonprofit organizations make judgments about overall organizational effectiveness on the basis of assessments of the performance of sub-units or of various functions is little studied. For the most part we give our attention in this paper to overall NPO effectiveness, though we do consider issues related to program effectiveness in a thesis concerning outcomes assessment.

A Little Theory (Goes a Long Way)

The academic study of organizational effectiveness has a long and tortuous history. This is not the place for a review of the theoretical literature on organizational effectiveness, but a few key themes and ideas must be noted. The most obvious (and most frequent) approach to both defining and evaluating organization effectiveness is to ask to what extent does an organization reach its goals. Of course, this question assumes that organizations have goals, that the goals can be discovered, that they are at least somewhat stable, that abstract goals can be converted into specific "objective" measures and that data relevant to those measures can be collected, processed and applied in a timely and appropriate manner. Much of academic organizational theory has observed that these are problematic assumptions. As organization theorists challenged the common sense understanding of formal organizations, which Pfieffer (1982) called the purposive-rational and Elmore (1978) the managed systems model of organizations, so too did the limits and assumptions of the goal model get exposed (see Mohr, 1982 for a thorough review of the criticisms of the goal model).

As we have suggested elsewhere (Herman and Renz, 1997a) the literature on organizational effectiveness is the development of alternatives to or modifications of the goal model. One prominent student of effectiveness has distinguished several "models" of effectiveness, including system resource, internal processes, multiple constituencies, competing values, legitimacy, and fault-driven (Cameron, 1986). More recently, Meyer and Gupta (1994) have compared five models: the maximizing,
political, constituency, business and paradoxical.

It might be objected that the proliferation of academic argument and commentary about organizational effectiveness is of no real consequence, as the "real world" uses simple, financial measures of the performance of businesses that of bottom line profit or share price (the implication being that nonprofit organizations can and should focus on goal achievement as defining effectiveness, even if they lack bottom-line financial measures). Clearly some people, most obviously some investors, use such measures in making decisions. However, there is substantial evidence that practicing (business) managers do not and cannot rely solely on a limited or singular measure. Interest in the "balanced scorecard" (see Kaplan and Norton, 1992) and in a multiple measure "business model" (see Eccles and Nohria, 1992) demonstrate that business managers take many goals and indicators into account. Indeed, Meyer and Gupta (1994) observe that over the last few decades the number of measures of business performance has proliferated, in part because measures "run down" and new measures must be created.

Purposive-rational approaches to organization theory imply that not only is it meaningful to conceive of organizations as having goals, even multiple goals, but also that assessment or evaluation of the goals is possible, though perhaps challenging. There is a reality about how well an organization (or its various sub-units) is achieving its goals. Many difficulties may stand in the way of capturing and communicating this reality, but at least in principle it is possible (see Murray and Tassie, 1994 for a thorough consideration of various models of organizational effectiveness evaluation, as applied to NPOs). The resurgence of institutional theory in the mid-1970's has emphasized the non-rational aspects of organizational behavior, paying particular attention to strategies of legitimizing organizations (see Scott, 1995 for an overview). For institutional theory organizational effectiveness is not an "objective" reality. Rather effectiveness is a social construction an achievement by organizational agents and other stakeholders in convincing each other that an organization is pursuing the right objectives in the right way. This way of describing the institutional theory perspective seems to suggest that effectiveness does not exist and, thus, is not important. This is not our understanding. The idea of effectiveness is important and stakeholders care about effectiveness. It is just that rather than existing independently "out there," effectiveness is social created by the actions and interactions of the stakeholders. Effectiveness is real and real in its consequences in the same way that "race" is a (socially constructed) reality with real consequences.

One implication of institutional theory is that the distinction between the reality of effectiveness and the evaluation of effectiveness is perhaps illusory, it is certainly not
meaningful. (This point is well-illustrated by the story of the three baseball umpires discussing their role. The first says he calls balls and strikes "as they are" and the second "I call 'em as I see 'em." The third says, "they ain't nothing 'til I call 'em.") In our discussion of the theses we present we will note what each thesis implies about the independent reality of NPO effectiveness and about evaluating NPO effectiveness.

Theses

1) Nonprofit organization effectiveness is always a matter of comparison. This assertion, we argue, is logically required, whether effectiveness is regarded as an objective reality or a social construction. How could a statement such as "X is a very effective organization" be meaningful unless X is being (implicitly perhaps) compared to other organizations, to some "ideal type" of organizations of its kind, or to itself at an earlier time? While this thesis is obvious, we feel it necessary to include since our experience is that sometimes people seem to be unaware of what the comparison is that leads them to a statement such as the above.

2) Nonprofit organization effectiveness is multi-dimensional and will never be reducible to a single measure. This thesis is much more consonant with a realist view of effectiveness than with a social constructionist view. The first portion is one of the few conclusions that we can safely say is widely (though not universally) accepted in the organizational effectiveness literature. What the dimensions of NPO effectiveness are and what dimensions are typically used are much less clear.

The competing values framework (Quinn and Rohrbaugh, 1981) identifies three dimensions (or sets of values) that constitute organizational effectiveness. One dimension concerns the extent of structure, with high flexibility at one end of the continuum and high control at the other. A second dimension is "focus," with an emphasis on the well-being and development of people to an emphasis on the well-being and development of the organization itself. The final dimension runs from an emphasis on means (processes) to an emphasis on ends (final outcomes). While the competing values model provides substance to the idea that organizations cannot maximize many important criteria at the same time, when actually used to assess organizational performance the multitude of criteria makes interpretation and the implications of the results difficult (see, for example, Rohrbaugh, 1981).
In contrast, the balanced scorecard approach identifies four perspectives for performance assessment: the customer perspective, the internal process perspective, the financial perspective, and the innovation and learning perspective. Organizations develop the measures appropriate to them for each perspective. Paton and Foot (1997) have proposed a scorecard approach for nonprofit organizations, suggesting that NPOs should develop and utilize five general categories of measures: (1) measures of current achievement of each activity or program in the organization; (2) measures of underlying (longer term) performance of each activity/program, taking into account alternative uses of resources; (3) measures of financial and management soundness (addressing the question of how well-run is the organization); (4) measures of renewal or decline (addressing the question of whether organizational capabilities are improving or wasting away); and (5) measures of the impact of organizational development initiatives. These suggestions seem useful though we are aware of no available research that describes the extent to which such performance assessment methods are being used by NPOs.

The multiple constituency model of organization effectiveness holds that various constituencies (or stakeholder groups) define the criteria with which they evaluate a given organization (D'Aunno, 1992, but also see Zammuto, 1984, for a discussion of alternative multiple constituency perspectives on identifying constituencies, their varying preferences and aggregating and using constituency assessments). Thus, different constituencies are likely to use or give different priorities to different criteria, making the possibility of widely applicable dimensions unlikely.

Apparently little empirical work has been done to identify possible dimensions of nonprofit organization effectiveness. Though it was not a purpose of their work, Stone and Cutcher-Gershenfeld's (forthcoming) review of empirical studies of NPO effectiveness suggests a wide variety of general theories or models of effectiveness have been used and a similarly wide-ranging set of criteria have been measured. Their review does not suggest any general dimensions have been proposed or generally used. Forbes (1998) review classed several studies as using a multi-dimensional approach.

In a study of effectiveness in nonprofit arts organizations Kushner and Poole (1996) seem to implicitly argue that not only is effectiveness an objective reality but that it is also unidimensional. They present a model of effectiveness that combines constituent satisfaction, resource acquisition, internal processes and goal attainment. The actual criteria they used to assess the performance of their sample organizations are not clear. Forbes (1998) considers this study as using a multi-dimensional approach.
Though Kushner and Poole use several indicators of effectiveness, the model they present treats constituent satisfaction as leading to resource acquisition effectiveness as leading to internal process effectiveness as leading to goal attainment as leading to overall effectiveness. Thus, we interpret their approach as assuming that these components are cumulative, not independent non-cumulative dimensions.

In our own recent work (Herman and Renz, 1997a) we developed a survey instrument to measure judgements of NPO effectiveness. After pilot testing and revision the final instrument contained items about financial management, fund raising, program delivery, public and media relations, community collaboration, working with volunteers, human resource management, government relations (regulatory compliance, contract administration and advocacy), and board governance. We expected factor analysis to find two or more dimensions, but all the items loaded on only one factor (the Cronbach's alpha, which measures the intercorrelation of the items, is .85). One interpretation of this unidimensionality is that respondents engage in "halo effect," the pattern of judging something consistently on many dimensions because of the human preference for consistency. We certainly consider it possible for an organization to be highly effective (or judged to be highly effective) on some dimensions while being ineffective on others.

The contemporary emphasis on outcomes assessment shows that some stakeholders (perhaps particularly institutional funders) find the program dimension especially important. Clearly more study of what dimensions of NPO effectiveness are actually used and useful needs doing.

That NPO effectiveness might ever be meaningfully expressed in terms of a single measure we believe completely impossible. Though the history of performance measurement in businesses is one of increasing numbers of measures (many of the measures are for sub-unit and functional performance), for certain purposes the share price (and changes in it) is a meaningful measure (Meyer and Gupta, 1994). The great advantage of share price (or of return on investment or other financial "bottom-line" measures) is their comparability across all manner of differences among firms. We can foresee no possibility of developing a similar measure for NPOs.

The fundamental reason why a single measure of NPO effectiveness is an impossibility is that the crucial exchange that NPOs help to enact is one "measured" in moral or value terms. In an important analysis Ostrander and Schervish (1990) argued that philanthropy is not an institutional sector but a particular sort of social (exchange)
relation that must be distinguished from commercial and political relations. Where the media of communication in commercial relations are money and in politics votes, in philanthropy the medium of communication is moral or normative. One implication of this view is that NPO effectiveness should be evaluated not in monetary terms but in moral terms, in terms of the extent to which NPOs respond to the moral values that motivate donors and recipients.

Recent trends suggest that NPOs are moving away from moral language and thinking while adopting commercial language and thinking. For example, a report in the April 9, 1998 issue of the Chronicle of Philanthropy describes the plans of a newly created nonprofit fund called New Profit Inc. This organization, according to the report, will "invest" in charities that are relatively young but successful. New Profit employees, like portfolio managers, will be awarded bonuses based on the successes or failures of the charities they oversee. Donors who think of and describe themselves as investors seeking a return and NPOs that seek to become more entrepreneurial and business-like potentially represent a turning away from a philanthropic (moral) basis of action, a turning away from doing what is right to doing what pays. No doubt advocates of such language, thinking and techniques will respond that they are trying to find better ways of assuring that doing what's right is also doing it well and that doing the right thing well must be rewarded.

3) Boards of directors make a difference in the effectiveness of NPOs, but how is not as clear. Stone and Cutcher-Gershenfeld (forthcoming) note that about half of the empirical studies they reviewed considered the question of the relation between board effectiveness and organizational effectiveness. The prescriptive literature on nonprofit boards has long asserted that not only do boards have crucial legal and moral responsibilities, but that boards can substantially affect the performance of the organization. Both common experience and systematic research have found that many boards do not fully meet the expectations set forth for them (see Middleton, 1987, Herman, 1989, and Fink, 1989 for discussion and analyses of the "board gap"). In response to the difficulties many boards have in fully meeting their prescribed responsibilities, a substantial set of standards about good board practices has developed. The generally accepted view is that boards that follow the right practices or use more of the practices will be more effective than boards that do not and that more effective boards will result in more effective organizations. The available research, though limited, confirms these assertions.

Bradshaw, Murray and Wolpin (1992) found that board involvement in strategic planning, use of good meeting management techniques and low conflict within the
board were related to the chief executives' assessment of board performance. They also found that the board having a common vision and being involved in strategic planning were very modestly correlated with chief executive assessment of organizational effectiveness. Since the measures of board practices, board performance and organizational effectiveness all derive from solely from one respondent (the CEO) for each organization, the extent to which the relationships found are due to common source or to reality is impossible to know.

A recent study by Green and Griesinger (1996) used multiple data sources. Board use of various board practices was used to measure board effectiveness and both board members and CEOs assessed board performance in relation (i.e., in terms of use of) 30 activities, such as "reviews and revises mission," "has term limits," and "formally evaluates board performance." Since board practices and board effectiveness were one and the same thing, the relation between use of board practices and board effectiveness could not be studied. Using a goal achievement approach, they assessed organizational effectiveness in three ways (rankings by practitioners using accreditation reports on study organizations, ranking by a government official whose agency contracted with the study organizations, and one of the researcher's rankings), combining the three sources into a composite.

When examining the relationships between various board practices/effectiveness and organizational effectiveness ratings Green and Griesinger often found divergence between the CEOs and board members. For CEOs organizational effectiveness is related to several board practices, including involvement in short-term and long-term planning, in policy formulation, in fund raising, including making personal contributions, certain aspects of financial management, and several aspects of board development (providing training for new members, formally evaluating board performance and defining specific duties of the board). For board members organizational effectiveness was related to fewer board practices and those that were significant were often different than the practices picked by the CEOs. Board members did not perceive organizational effectiveness as closely related to board involvement in fund raising or in making personal contributions, nor did they see participation in planning as related to organizational effectiveness. In this study which board practices are related to organizational effectiveness differs according to the position of the respondent.

In a study of 64 NPOs Herman, Renz and Heimovics (1997) found that the extent of use of 25 recommended board practices was moderately correlated with CEOs' judgment of board effectiveness but not with judgments of board effectiveness by
board members or funders (foundation, corporate giving and federated fund raising organization officials). More recently we (Herman and Renz, 1997b) analyzed these relationships by selecting out the top ten most effective organizations and the bottom ten least effective organizations (based on averaging the effectiveness ratings from management, board and funder stakeholder groups). We became interested in analyzing the top and bottom organizations when analyses showed that the organizations about which multiple constituencies most consistently judged in the same way were those that were especially effective. We found that the top ten organizations used more of the recommended practices. On only a few specific practices were there notable differences the especially effective more often used board self-evaluation, more often had written expectations about giving and soliciting contributions and the chief executive more often was extensively involved in board nominations. In the earlier study, using all 64 organizations, we found a strong correlation between averaged judgments of board effectiveness and averaged judgments of organizational effectiveness (where different individuals provided the board and the organizational effectiveness judgments). In the top and bottom ten analysis we found an even stronger correlation between board effectiveness and organizational effectiveness.

Recent work by Holland and colleagues in developing a board self-assessment instrument (called the Board Self-Assessment Questionnaire) also provides strong evidence that effective boards are related to effective organizations. Jackson and Holland (1998) studied the relation between the financial performance of thirty-four private colleges and seminaries and various board practices included in the instrument. They found moderate correlations between the overall score on the instrument and its six components. They also report results of a board training intervention. Not only did participating boards show improvement in scores on the instrument, but there was also a modest correlation between change scores on the instrument and gains in financial goal achievement scores (see Holland and Jackson, forthcoming).

Thus, while only a few studies have been done, given that those studies have used different kinds of NPOs, different conceptions and measures of both board and organizational effectiveness the consistent finding that board effectiveness is related to organizational effectiveness must be considered to be strongly supported. Less clear are what specific behaviors and/or characteristics of boards make for greater board effectiveness. Also the causal relation between board and organizational effectiveness remains little researched.
All the previously cited studies of board effectiveness include, in one way or another, various measures of board behavior. Whether more effective boards do more, do different things or do what they do better (or some combination) than less effective boards is not clear. Certainly, there is no specific behavior that seems to identify effective boards across all the studies. Given the great variety of circumstances in which boards and organizations work it seems unrealistic to expect to find one (or a few) simple practices that "guarantee" board effectiveness.

Surprisingly research on board effectiveness has paid little attention to the social characteristics of board members. Some earlier research on nonprofit boards (e.g., Pfeffer, 1973 and Provan, 1980) investigated the relation between social status and funding outcomes and we know that NPO boards differ substantially in their prestige and that relative prestige of boards is widely known in local communities (Galaskiewicz, 1985). More effective organizations (whether conceived in realist or social constructionist terms) are more likely to be able to attract more prestigious people to their boards. Thus, it is possible that the apparent relation between board effectiveness and organizational effectiveness is spurious the true relation is that effective organizations have prestigious boards.

In our study of 64 NPOs we measured the prestige of the organizations' boards and found in our analysis of the differences between the top ten and bottom ten that the top ten did have more prestigious boards. However, the strong correlation between judgments of board effectiveness and organizational effectiveness was virtually unchanged when we statistically controlled for board prestige. Apparently the extent of board prestige has little affect on how various stakeholders judge board effectiveness. It seems plausible to us that board prestige is of little value to NPOs unless it is "activated." That is, a highly prestigious board that does little is unlikely to be effective. Highly prestigious boards must act to make their social clout useful. No doubt it is easier (takes fewer calls) to raise a certain amount of money with a high prestige board. While (as usual) more research is needed on this topic, at present it seems safe to say that most boards can improve their effectiveness by improving what they do and how they do it.

The general assumption, among researchers and practitioners, is that effective boards lead to effective organizations (indeed, our experience is that people want to believe this). There is little evidence, one way or the other, about the causal relation. Effective managers and staff might improve organizational effectiveness and, thus, be more effective at recruiting better board members, resulting in more effective boards as a consequence. The intervention study reported by Holland and Jackson (forthcoming)
4) More effective nonprofit organizations are more likely to use "correct" management practices. Forbes (1998) Stone and Cutcher-Gershenfeld (forthcoming) in their recent reviews observe that, in addition to studies of the relation between board and organizational effectiveness, the small empirical literature on NPO effectiveness has also given some attention to the relation between planning and organizational effectiveness. In a study of 179 churches Odom and Boxx (1988) found a relation between use of planning and growth in such performance measures as Sunday school attendance, offerings, and total membership. Crittenden, Crittenden, and Hunt (1988) in a study of 600 heterogeneous NPOs in a southwestern (U.S.) state selected 28 responding churches for analysis. They measured the extent to which the churches used different elements of the strategic planning process (such as goal setting, financial analysis and forecasting, stakeholder analysis, budget development, implementation management and alternative analysis). They used four measures of performance, growth in contributions, top administrator satisfaction, growth in membership and growth in services, activities and product offerings. Their key analyses were correlations between the six planning elements and the four performance measures. They found statistically significant correlations (in the predicted direction) in 13 of the 24 combinations. The only planning element that was uncorrelated with any performance measure was stakeholder analysis. Implementation management and alternative analysis were significantly correlated with three of the four performance measures. Administrator satisfaction was significantly correlated with four of the six planning elements (though this may be, at least in part, attributable to common source variation).

More recently Siliciano (1997) studied 240 local YMCAs in twelve eastern U.S. states. She used two kinds of performance assessment: a ranking by a regional YMCA official and a financial performance measure, the ratio of total revenues to total operating expenditures. She too measured elements of the planning process, such as mission statement, environmental trend analysis, competitive analysis, long-range goals, short term objectives, action plans, and monitoring of results, where local YMCA CEOs were asked to rate the extent of activity in relation to each element. Siliciano also noted that prior research on planning and performance has been confounded by the relation of size to both of these variables. Thus, she included organizational size (as measured by annual revenues) as a control variable. The results show that most of the formal planning elements are correlated (after controlling for organizational size) with both measures of performance (the mission statement element is uncorrelated with both measures and environmental trend analysis is not...
correlated with financial performance). Having a board strategic planning committee was correlated with three of the planning elements. The results also showed that financially successful YMCAs were more likely to be active planners than the less successful.

The evidence is pretty strong, we believe, for the view that strategic planning is related to effectiveness in NPOs. However, whether planning leads to effectiveness or whether effectiveness leads to planning is much less clear. Planning is one of many forms of "correct management" procedure. There has been very little research on other forms of correct management.

In further analyses of data from our study of effectiveness in 64 NPOs we found that the most effective differed from the least effective in the use of several correct management procedures (Herman and Renz, forthcoming). The correct management procedures were those identified by focus groups of practitioners. The correct procedures they considered indicators of effective NPO management included having a mission statement, a recent needs assessment, a planning document, measuring client satisfaction, formal CEO and employee appraisal process, an independent financial audit, and a statement of organizational effectiveness criteria, among others. The more effective used more of the correct management practices and use of correct procedures is positively correlated with judgments of organizational effectiveness for all three stakeholder groups (funders, board members and senior managers). We also found that the more effective more often pursue strategies of seeking new revenues and strategies to enhance legitimacy.

However, relationships between doing things right and being judged effective were not found in the full sample. This suggests that for many NPOs (those not consistently judged as either highly effective or less effective those in the "murky middle") the connection between management practices and effectiveness (as indicated by constituency judgments) is weak or non-existent. The evidence for concluding that correct management enhances NPO effectiveness is clearly scanty.

5) Nonprofit organization effectiveness is a social construction. This thesis is one which we are sure is likely to find limited agreement. The social constructionist ontology challenges our taken-for-granted understanding of the world and how it works. A social constructionist view holds that some parts of reality do not exist independently of the beliefs and actions of people. Many aspects of the social world are real only because people have believed and acted in ways that are consistent with
creating that reality. The idea and categories of "race" are increasingly recognized as social constructions rather than as physical (biological) realities.

5) Nonprofit organization effectiveness is a social construction. This thesis is one which we are sure is likely to find limited agreement. The social constructionist ontology challenges our taken-for-granted understanding of the world and how it works. A social constructionist view holds that some parts of reality do not exist independently of the beliefs and actions of people. Many aspects of the social world are real only because people have believed and acted in ways that are consistent with creating that reality. The idea and categories of "race" are increasingly recognized as social constructions rather than as physical (biological) realities.

As we observed earlier if one accepts that NPO effectiveness is a social construction, then the distinction between actual organizational effectiveness and evaluation of organizational effectiveness disappears. A good analogy is that of umpiring balls and strikes in a baseball game. There are pitches, but a pitch is neither a ball nor a strike until the umpire calls it. In NPOs there are activities, but there is no effectiveness until someone calls it. Of course, unlike the baseball illustration, in NPOs there is no single umpire (all stakeholders are permitted to call effectiveness, though some stakeholders may be considered more credible than others) and there is no regular procedure for calling effectiveness (all stakeholders are free to send messages, with varying degrees of justification, about effectiveness and to focus on different activities as more relevant to calling effectiveness). Effectiveness is stakeholder judgment, formed and changed in an ongoing process of sense-making and negotiation. Our view is that the social construction of NPO effectiveness is not necessarily stable or complete.

Conceiving of NPO effectiveness as a social construction does not mean that the judgments (and actions taken in relation to those judgments) constituting the social construction are capricious or arbitrary. Conceivably the social processes resulting in judgments of NPO effectiveness could lead various constituencies to develop the same criteria and to evaluate information relevant to those criteria in the same way. If so, research would find that stakeholders reach the same judgments about effectiveness and it would appear that there is a real property, called effectiveness, "out there," existing independently of the people and the process they use to create it. Indeed, if this were always the situation research on NPO effectiveness would be much simpler. However, our recent research shows that stakeholder judgments of NPO effectiveness often differ (Herman and Renz, 1997a). We found consistent judgements of effectiveness much more common among those NPOs that all stakeholder groups agreed were highly effective (Herman and Renz, forthcoming). Apparently NPO stakeholders are more agreed about what high performing NPOs are
than about what more middling or poor performers are.

6) **Program outcome indicators as measures of NPO effectiveness are limited and can be dangerous.** We recognize the value of outcomes assessment, certainly many NPOs would likely benefit from doing (or doing better) outcomes assessments. Our concerns are that outcomes assessment is being promoted as promising more than it alone can deliver and that outcomes assessments might be used in unfair and dangerous ways.

A problem with outcomes assessment is that it encourages people to commit the fallacy of affirming the consequent. The (flawed) logic is "If A, then B. B is observed to occur, thus A must have caused B," where A represents all those program activities that precede B, the outcomes. For example, we might argue that if the sun goes around the Earth, then the sun will rise and set. Establishing that the sun rises and sets (that is, affirming the consequent of premise A), of course, does not require that the sun go around the Earth. Other explanations for the sun's rising and setting are possible.

The idea of outcomes assessment is that program objectives are stated in specific, measurable terms and that these outcomes are consistently tracked. For example, consider a tutoring program for grade-school youngsters. The intended outcome is to improve average scores by a certain number of points or a certain percentage on some standardized test of reading achievement. If such an outcome occurs, of course, there are likely many other possible causes, including maturation, a change in reading instruction in schools, the youngsters' participation in other educational programs, changes in family behavior and so on.

Experts (e.g., Plantz, Greenway and Hendricks, 1997; Martin and Kettner, 1996) note this and other limitations of outcomes assessment, but much remains to be learned about how staff, board members and funders are actually using outcomes assessment. We read about and hear people talk about "proving effectiveness" and "showing results." Such language suggests many people believe outcomes assessment can and will establish causality. There is danger in such unjustified causal inferences in both positive and negative outcome situations. In the positive situation a NPO may take credit for an outcome for which it is only slightly (if at all) responsible. Funders might be convinced and greatly enhance the program, when another use of those resources might have been more effective (in relation to the intended outcome). In the negative situation (say there was no improvement in some measured outcome over time) the inference likely would be that the program is a failure. It is conceivable that "no change" in participants might represent a considerable achievement, if comparable
people who were not receiving program services were declining on the outcome measure. The lack of comparison groups in most outcomes assessments makes for the possibility of "unfair" decisions.

There are other dangers. In an effort to "show results" NPOs may use data they already regularly collect to measure some intended outcome when such data is not the most appropriate measure. Questions of who provides funding for increased (and different) data collection efforts remain to be answered. In some social service programs in the 1960s and 1970s that were evaluated in relation to outcomes, such as job training programs, the practice of "creaming" was frequently observed. "Creaming" occurs when those eligible recipients who are most ready for the service and least likely to need it are selected for the program.

**Conclusion: Possible Futures for Research on NPO Effectiveness**

Beyond reiterating the usual mantra that more research is needed, we think this the place to consider some possible futures of research on NPO effectiveness. Such research has been rather sparse. We believe two principal factors account for the sparseness of such research: one is the unique criterion problem of defining and measuring NPO effectiveness; the other is the general conceptual/theoretical problem of multiple models of effectiveness. Possibly because of the inadequacies of the available research both practitioners and most scholars assume that NPO effectiveness is very similar to business effectiveness. Thus, the general view is that what is learned about business effectiveness can be applied directly to NPOs. However, though more extensive, the academic study of effectiveness in business organizations seems to be about as inconclusive and muddled as for NPOs. Research on strategic planning and business effectiveness, for example, seems not to support any general conclusions. Academic organizational researchers seem mostly uninterested in effectiveness. Among those who are interested in effectiveness studies the analytical challenges of modeling a contingency theory notion of "fit" continue to escalate (see Edwards, 1993 and Edwards and Parry, 1993 for criticisms of the most commonly used computations of fit and for explication of using and interpreting polynomial regression results in relation to fit). It is little wonder that the popular literature on organizational effectiveness is completely unrelated to the academic study of business effectiveness.

Thus, in practice many ideas first instituted and popularized in business (such as strategic planning, visioning, TQM, benchmarking and others) are later adopted by NPOs. Of course, the expectation is that what "works" in business should also work in NPOs. Another view is that keeping up with what are regarded as "best practices" is a
sign of effective management and legitimates a NPO in the eyes of many of its stakeholders. The current trend of seeing the future of NPOs as one of their becoming more business-like promises to further enhance this general pattern.

(1) One likely future, perhaps the most likely future, for NPO effectiveness research is a continuation of the current situation. Few studies will be done and those that are done will be based on differing models of effectiveness, continuing to make accumulation and integration of results difficult. While no "universal" bottom-line indicators will be developed (preventing any meaningful comparisons across different "types" of NPOs), NPO effectiveness will not be regarded as a problem. People will continue to trust their judgments and the social construction of effectiveness will allow participants to all pretend (and believe) that effectiveness they think they agree on is real.

(2) Another future which we think fairly probable is that NPO effectiveness research splits into focusing on multiple "types" of organizations, where sometimes all of the organizations are NPOs and sometimes NPOs are only part of the collection. To some extent, this is true now. Researchers interested in the effectiveness of universities and colleges do not pay much attention to sector location; similarly research on hospital effectiveness appears to give only limited attention to auspices. The clear advantage of studying "types" is the possibility of using standardized measures of financial and/or program outcomes as the effectiveness criteria. The disadvantage (which is true now) is those studying each type focus solely on that type (assuming the type's distinctiveness) leading to even greater fractionating of knowledge and incommensurability of theories and findings.

The question of how to identify "types" would need to be considered. Should "human services" be a single type or should "mental health" be distinguished from "emergency assistance" and so forth? Is some more theoretical basis for distinguishing types available or should "common sense" types be used?

(3) Another possible future is that the general distinctiveness of (public benefit charitable) NPOs continues to be taken seriously and more researchers take up the challenge of trying to develop conceptions and indicators appropriate for NPOs. Though whether and to what extent NPOs differ from business and government organizations continues to be a disputed issue, there seems to us enough convincing evidence to conclude such. A U.S. national probability sample of work establishments (i.e., the specific plant or office rather than the usually larger organization), identified
7% of work establishments as private nonprofit settings (Marsden, Cook and Knoke, 1994). In analyses of data from this sample statistically significant "effects" of an organization being a private, nonprofit organization are frequently reported (see, for example, Marsden, Cook and Kalleberg, 1994; Knoke, 1994; and Marsden, 1994). Such results support the view that NPOs are different from business and government.

Of course, the most important challenge faced by those who would focus on NPO effectiveness is that of the criterion. Is it possible to settle on a small number of fairly easily measured indicators? One general indicator that might be used is size (as measured by total revenues). To some extent size represents the willingness of various constituencies to support an organization. Obviously size is by no means a "perfect" measure. Size is highly correlated with age and big, old NPOs may be more able to "muscle out" upstart organizations that attempt to appeal to the same constituencies. Similarly, growth might be used as another general indicator, as it can be interpreted as indicating current responsiveness to constituency needs and expectations.

We believe that what is crucially needed to advance research on NPO effectiveness is development of a measure of responsiveness as an indicator of effectiveness. As we have argued elsewhere (Herman and Renz, 1997a) a social constructionist multiple constituency conception of NPO effectiveness implies that NPO leaders must see their effectiveness as responsiveness to the needs and expectations of their stakeholders. Tsui (1994) has proposed that stakeholder responsiveness should be the basic criterion of effectiveness, both for individual managers as well as organizations per se. Is it possible to develop a general and easily used questionnaire to assess stakeholder satisfaction with the responsiveness of NPOs? Will the leaders of NPOs see the use of such a measure as useful enough that it might become widely used? If the answers to these two questions were yes, then we would be in a much better position to really make progress in researching NPO effectiveness.

References


