

Multi-Criteria Performance Measurement for Public and Private Sector Enterprises

Patrick T. Hester and Thomas J. Meyers

ABSTRACT

Historically, public and private sector enterprises have been viewed as existing on opposing ends of the performance measurement spectrum, due to seemingly incompatible worldviews. Private sector enterprises are traditionally viewed as profit-driven and focused on a return on investment paradigm, while public sector enterprises are seen as mission-oriented and answerable to a paradigm less focused on investment and more on improving enterprise capability. The authors propose that, in fact, these worldviews are not mutually exclusive, as private and public sector enterprises must both account for investment and mission concerns. In order to leverage real synergy to be gained from distinct but complementary viewpoints, a systemic approach to evaluating organization performance through the novel fusion of operational test and evaluation and multi-criteria decision analysis is developed. Use of this framework is demonstrated within an enterprise that involves consideration of public and private sector concerns. The authors hope that the approach proposed in this chapter will enable public and private sector enterprises to comprehensively address performance.

1. Introduction

The issue of enterprise performance measurement has made many agendas for many years (Neely, 1999), and just as ubiquitous have been related concerns regarding distinctions between public and private enterprise sectors, differences of near- versus long-term perspectives of organizational management or governance, and the need for multi-criteria approaches to performance measurement. This paper seeks to seize what its authors view as an opportunity to advance the cause of organizational performance measurement and performance measurement systems in a way that respects the prominent concerns related to each.

The past forty years have seen significant research devoted to exploring the differences, similarities, and systems for measuring the performance of public and private enterprises. Literature of that same period, however, indicates no real consensus among practitioners and academics as to the meaning of the *public* and *private* labels. Nutt and Backoff (1993), for example, provide the following definitions of the two terms:

"'Public' and 'private' are taken from the Latin: The word 'public' means 'of the people,' as contrasted with 'private,' which means 'set apart'" (p. 210).

Boyne (2002) expands on those definitions in a summary of Rainey, Backoff, and Levin (1976):

"Whereas private firms are owned by entrepreneurs or shareholders, public agencies are owned collectively by members of political communities" (p. 98).

Perry and Rainey (1988) echo the emphasis on ownership with:

"The most common definition of public and private distinguishes organizations according to ownership and/or funding" (p. 184).

Fottler (1981) complements the strict focus on public or private, per se, by suggesting an organization classification continuum extending from *private for-profit*, through *private non-profit* and *private quasi-public* to *public*. With no universally accepted, fundamental distinction between public and private sector enterprises in hand, the authors subscribe to the analytical approach of Rainey et al. (1976) by choosing to distinguish organizations in terms of ownership and funding source factors. Herein, therefore:

Public sector enterprises are those owned by and funded through taxation of the general public, whereas *private sector enterprises* are those owned by a non-government entity and supported by non-government funds, namely sales or private support.

This and other definitions possibly only hint at whatever consequential differences and similarities might be assigned to the public and private sector pairing, and so the authors elaborate on that comparison in Section 2. Section 3 addresses the history and scope of organizational performance measurement as well as how the performance of public or private enterprises might be undertaken equally effectively within a common framework. Section 4 offers the authors' blending of public and private organization performance perspectives together with a common measurement methodology that they demonstrate using a scenario that plainly evokes what have historically been distinguished as public and private concerns. This chapter's final section offers conclusions and thoughts regarding future use and development of the Section 4 methodology.

2. Comparing Private and Public Sector Enterprises

Literature comparing public and private sectors reveals the lack of a universal perspective on their inherent similarities and differences. More precisely, there exists a marked bifurcation of research supporting and rejecting the notion that private and public sector enterprises are fundamentally similar. Bozeman and Bretschneider's (1994) and Scott and Falcone's (1998) in-depth treatments address three schools of thought regarding public and private sector differences and similarities:

1. The "generic" approach: This approach downplays the significance of distinctions between the public and private sectors.
2. The "core" approach: This approach espouses fundamental distinctions between the public and private sectors.
3. The "dimensional" approach: This approach distinguishes organizations based on a number of dimensions that are independent of organizational sector.

Each of these three finds significant support in the literature and should be considered with any attempt to develop a measurement framework or methodology commonly applicable to public and private sector enterprise performance. Following text therefore reviews the generic, core, and dimensional approaches.

2.1. The Generic Approach

Proponents of the generic approach believe that fundamentals of organizational management transcend whatever boundaries that may be perceived to exist between the public and private sectors. Weinberg (1983), for example, contends that many perceived sector differences actually represent distinctions in organizational type rather than sector type, rendering derived notions of sector differences nothing more than the result of "apples to oranges" comparisons. She says of those unfair comparisons:

One reason why the impression of vast differences is so strong is that when generalizing about public enterprises, commentators tend to focus on the public agencies that are most frequently at the center of political controversy: large, complex organizations that command constant attention from top management. Forgotten for the moment are the water bureaus, forest services, agricultural extension services, and fire departments that perform routine, largely noncontroversial functions (Weinberg, 1983, pp. 107-108).

She continues: "When looking for similarities and differences that can be attributed to the effects of being in the public or the private sector, what obviously ought to be compared are similar activities" (1983, p. 108). She plainly maintains that comparisons of significantly dissimilar organizations, regardless of sector affiliation, should be expected to identify significant differences in organizational characteristics, operations, and productivity.

Knott (1993) and Lachman (1985) echo the need to compare similar organizations in order to make objective conclusions regarding their similarities and differences. Knott (1993) indicates that similar organizations have similar management processes, while Lachman (1985) concludes that the activities of similar organizations produce

similar results. Lan & Rainey (1992) indicate there to be only insignificant differences in the goals established by sector-distinct organizations, and Murray (1975) posits the management approaches exercised within each of the public and private domains to be more similar than different, with desired ends possibly differing but with the means typically employed to achieve desired ends actually quite similar.

Several authors have also investigated particular organizational attributes, and have shown there exists little empirical evidence of sector-dependent differences in those. In exploring the claim that public managers are more risk averse than their private counterparts, Bozeman & Kingsley (1998), for example, find no evidence for it, noting, "...the empirical claim that public sector managers are more risk averse than private managers has not been conclusively determined" (p. 109). Baldwin (1987), too, addresses characterizations routinely advanced in the literature to suggest public and private sector enterprise distinctions not grounded in empirical study; characterizations that, for example, have public sector enterprises suffering far more from poorly defined goals and frequent leadership turnovers than their private counterparts, as well as enjoying degrees of job security quite atypical of the private sector. Baldwin understandably proposes wary review or use of such claims.

Despite the literature's strong subjective inclination to dismiss the notion of significant differences between public and private sectors, and despite some empirical evidence promoting the idea of organizational distinctions rooted only in organizational-level attributes, Scott and Falcone (1998) counter that they have found a "...lack of [empirical] support for the generic approach" (p. 141). Perhaps, then, the generic approach fails to fully capture the nuances of the public-private debate. The authors therefore turn their attention to the core approach.

2.2 The Core Approach

The core approach insists that fundamental differences do exist between public and private sector enterprises. Although many of its supporters concede that some similarities may be found in organizational-level attributes such as composition or managerial tasks, this scheme's supporters contend sector differences to be more fundamental than any similarities possibly identified. Many authors have conducted comprehensive reviews of differences in public and private sector enterprises (Nutt & Backoff, 1993; Perry & Rainey, 1988; Rainey, 1991; Rainey, et al., 1976) and so found key differences in areas such as political context, managerial roles, environmental factors, organization-environment transactions, internal structures, and processes. Particular findings require fuller descriptions than what can be offered in the limited space here, but notable among the many studies to have determined sector-based enterprise distinctions are those regarding level of bureaucracy (Boyne, 2002), level of autonomy (Flynn & Tannenbaum, 1993), organizational commitment (Boyne, 2002; Flynn & Tannenbaum, 1993), distributive and procedural justice perceptions and satisfactions with supervisors (Kurland & Egan, 1999), organizational composition and associated ramifications (Perry & Rainey, 1988; Rainey, 1989; Rainey & Bozeman, 2000), personnel and purchasing processes, work attitudes and values (Rainey & Bozeman, 2000), and political context and managerial roles (Rainey, 1989).

Fottler (1981) concludes that differing public and private sector funding sources drive differences in organizations' operating environments, values, incentives, constraints, and management processes. Nutt (2005) reinforces that conclusion by citing decision-making processes as functions of an organization's sector affiliation; precisely, his study shows private sector managers to favor budget decisions determined with disinterested, rigorous analysis, as opposed to public sector managers who commonly arrive at budget decisions through bargaining. Still other researchers have cited sector distinctions as driving markedly dissimilar organizational goals, with private sector goals largely profit-oriented and public sector goals relatively far more multi-dimensional and hence also more difficult to achieve (Hedley, 1998; Kuin, 1968; Rainey & Bozeman, 2000; Tirole, 1994); those relatively complex, public organization goals, moreover, are perceived by managers to foster a greater number of rules and lower perceived efficiency than those linked to private enterprise (Lan & Rainey, 1992). Scott and Falcone (1998) nicely summarize that "core differences clearly exist, and to ignore or discount them runs counter to prior and current empirical evidence" (p. 140). They continue, however, that "At the same time...a singular focus on core differences presents somewhat of a limited picture" (p. 140) and then suggest the dimensional approach as a plausible alternative to or complement of the purely core scheme. Section 2.3 focuses on the dimensional approach.

2.3 The Dimensional Approach

The dimensional approach assumes that organizations can be categorized along several dimensions of "publicness", a term which Bozeman and Bretschneider (1994) define as "a characteristic of an organization which reflects the extent the organization is influenced by political authority" (p. 197), independent of the organization's sector. All organizations may be considered to be more public along certain dimensions (for example, having a large amount of political oversight and accountability) and more private along other dimensions (for example, being market-driven and dependent on private financing for viability). Dahl & Lindblom (1953) and Wamsley & Zald (1973) provide this scheme's theoretical underpinnings, while Bozeman's (1987) assertion that few, if any, organizations are truly public or private points to a fundamental flaw in the core approach as it clearly represents dimensional approach advantages for organizations neither strictly public or private.

Bozeman (1987) further notes the dimensional approach to support identification of, for example, economic elements integral to public enterprises and political elements part-and-parcel of private enterprises. Scott and Falcone (1998) find empirical evidence of the approach's utility, noting their results to "...confirm and extend the growing body of research that has demonstrated the utility of the dimensional approach to the public/private question (e.g., Bozeman & Bretschneider, 1994; Coursey & Bozeman, 1990)" (p. 140).

The authors believe that measurement of public and private enterprises might be only recklessly split across a divide that in many ways does not exist. They also believe the dimensional approach to be greatly aligned with their notion that public and private enterprise performance can and should be commonly measured. Besley and Ghatak (2005) agree, stating that "Not all activities within the public sector are mission-oriented.... Not all private sector activity is profit-oriented" (p. 616). Bhattacharyya (1968) trumpets similar notes with his call for all organizations, regardless of public or private labels, to evaluate their performance as much in terms of profits and return-on-investment as in terms of social and political goals. Gold (1982) also twice echoes a multi-dimensional perspective of performance with "The perception of the importance of the role of organizational mission in success clearly cuts across the public and private sector organizations studied" (p. 569) and with, "No comparative analysis of the public and private sectors can ignore the dimensions of profit" (p. 570). Ramamurti (1987) proposes, among other models, a multi-criteria approach to performance evaluation emplacing profit as well as more mission-centric characteristics as performance criteria (interestingly, his empirical analysis shows a majority of public managers to place an overwhelming emphasis on profit, despite their espousal of a more balanced approach). Lastly, Sheehan (1996) emphasizes the importance of considering mission when evaluating performance, noting "One of the many 'defining models' of effectiveness holds that an organization is effective to the extent that it accomplishes its mission" (p. 110).

Profit- and mission-related concerns, it seems, should together and equally represent elements of measurement systems designed for and applied against the performance of organizations, public or private. That and other considerations will show themselves with the treatment of performance measurement offered with Section 4. With the themes of this section now having been presented, the following section will introduce a final set of interests needed to underlie the methodology that Section 4 will demonstrate: those of performance measurement and performance measurement systems.

3. Performance Measurement

The preceding section noted that the literature of organizational performance measurement in some ways reflects a dual-track perspective of public *or* private enterprises. Many researchers have focused on business-related, or private sector, concerns, while many others have focused on the performance of public sector, primarily government organizations. Although the distinct interests of researchers do not necessarily represent distinctions to be observed with public or private enterprise performance measurement, the relative paucity of publications focused on measurement constructs equally applicable to public *and* private pursuits understandably prompts concern for whether or not the two should be measured differently. This section's review of sector-agnostic measurement constructs drawn from the engineering regimen of test-and-evaluation will complement prior section notions of a "dimensional approach" to sector classifications and, in so doing, will additionally catalyze the authors' ensuing suggestion that performance measurement systems (PMS) may and possibly should ignore whatever significant differences that exist between organizations typically characterized as either public or private. This section will

therefore demonstrate real value to be mutually gained by public and private interests from PMS also characterized far more in terms of their commonalities and robustness than by their distinctions.

3.1. Performance Measurement and Performance Measurement Systems

Public and private sector organizations have for several decades acknowledged the importance of measuring their own performance (Behn, 2003; Bourne, Mills, Wilcox, Neely, & Platts, 2000; Gimbert, Bisbe, & Mendoza, 2010; National Performance Review, 1997; Neely, 1999; Wilcox & Bourne, 2003), and myriad systems have been pursued to meet measurement needs (Behn, 1995; Bourne, et al., 2000; Bourne, Neely, Platts, & Mills, 2002; Griffin & Page, 1996; Neely, Mills, Platts, Gregory, & Richards, 1996; Rouse & Puterill, 2003). With that attention has come a fair consensus among public and private practitioners, as well as academic interests, of what is meant by each of the terms, performance measurement and performance measurement system.

Kerssens-van Drongelen (1999, p. 4) and collaborators (Kerssens-van Drongelen & Bilderbeek, 1999, p. 36; Kerssens-van Drongelen & Cook, 1997) have offered a definition of "company" (p. 347) *performance measurement* that we have here modified ever so slightly to yield a general definition quite representative of many proposed:

Performance measurement is the acquisition and analysis of information about the realization of organizational plans and the factors that may influence plan realization.

Kerssens-van Drongelen (1999, p. 4) and Kerssens-van Drongelen and Cook (1997, p. 347) have likewise defined a *performance measurement system* in a manner able to serve as the kernel of a generalized definition respecting not only their own and others' emphases on performance information collection and processing but also what this work's authors and more (Behn, 1995, 2003; Bourne, et al., 2000; Chiesa, Frattini, Lazzarotti, & Manzini, 2008, 2009; Fitzgerald & Storbeck, 2003; Kaplan & Norton, 1992; Kaplan & Norton, 1996; Kouzmin, Löffler, Klages, & Korac-Kakabads, 1999; Micheli & Manzoni, 2009; National Performance Review, 1997; Sproles, 2000a, 2000b, 2002) recognize as the fundamental need for measurement planning, particularly as it regards the a priori establishment of measurement standards, or performance criteria, appropriate for given purposes and contexts, and equally suited to the domains of public and private enterprise:

A performance measurement system comprises procedures and tools that support some measurement process. It is a mechanism by which performance criteria are established and by which the information to be evaluated against those criteria is identified, collected, and analyzed.

Behn (2003) strongly argues that purpose should drive measurement, and he identifies eight reasons to measure organizational performance within a context of public enterprise management:

- | | | |
|-----------------|------------------------|------------------|
| (1) to evaluate | (2) to control | (3) to budget |
| (4) to motivate | (5) to promote | (6) to celebrate |
| (7) to learn | (8) to improve. | |

He acknowledges interdependencies among many of these and most significantly contends that "the other seven purposes are all subordinate to improvement" (p. 600). Behn's set of purposes and particularly the primacy of (organizational) improvement:

(a) is echoed in notable parallel research (Bourne, et al., 2000; Chiesa, et al., 2008, 2009; Fitzgerald & Storbeck, 2003; Kaplan & Norton, 1992; Kerssens-van Drongelen & Bilderbeek, 1999; Kerssens-van Drongelen & Cook, 1997; National Performance Review, 1997; Neely, 1999; Neely, et al., 1996; Rouse & Puterill, 2003; Simeone, Carnevale, & Millar, 2005);

(b) implies the importance of carefully crafted measurement schemes to public or private organizations intent on performance improvement; and

(c) is reminiscent of the "what's the problem we're trying to solve" adage commonly attributed to engineering circles.

The last two of those traits strongly suggest a generalized utility to be gained from measurement approaches drawn from one of the engineering domain's most prominent components - that of operational test and evaluation - and applied to either public or private organization performance.

3.2. A Test and Evaluation-derived Approach to Performance Measurement

Section 3.1 identified much of the last several decades of study and practice dedicated to the cause of organizational performance measurement. That effort fostered a healthy set of measurement frameworks such as the popular balanced scorecard of Kaplan and Norton (1992) and many of what could be considered more method-level schemes classified by Poh, Ang, and Bai (2001) as either "weighting and ranking" or "benefit-contribution" (p. 64) types. The authors submit that there may be room for a new framework and derived methods complementing established ones and that draws from Stevens' (1979) classic definition of operational test and evaluation (OT&E) as an effort "to determine the performance of a system under the most current operational conditions" (p. 4). Likening public and private enterprises to the systems upon which Stevens' definition is focused, the authors propose a public and private sector-suited framework that respects what Stevens (pp. 4-5) has identified as OT&E's principal objectives:

- To determine whether a system, in combination with its operators (and all supporting parties) can fulfill its current missions and objectives;
- To develop methods and procedures for the optimum employment of new systems or for the use of old systems in new ways to satisfy new missions and objectives or to interface properly with new equipment;
- To establish the limitations, characteristics, and capabilities of a new system to determine how it can best be integrated into an (overarching system's) management structure; and
- To provide information that will assist in the research and development of new systems through documenting needs for improved performance or different performance and in determining the deficiencies from a performance standpoint of the system under test.

The authors maintain that such test, or measurement, objectives clearly relate to organizational improvement, purposes, and operational context. They further posit that approaches to organizational performance measurement could take advantage of three concepts increasingly favored (Campbell, 2004) within the OT&E community: concepts of critical operational issues (COI), measures of effectiveness (MOE), and measures of performance (MOP), all of which may be considered "manifestations of good management practices" (Sproles, 2000a, p. 51) and so every bit as useful for non-engineering applications as for those of test and evaluation or other engineering settings (Sproles, 2000a).

3.2.1. Critical Operational Issues

Stevens' (1979) and particularly Sproles' (2000a, 2000b, 2001, 2002) prompt the following definition of *critical operational issues*:

Incontrovertible, stakeholder-recognized needs directly derived from problems that must be satisfactorily addressed; emergent essentials of capability without which posited problem solutions must be judged as unacceptable on functional grounds.

Critical operational issues represent itches that must be scratched to bring satisfaction; nothing less will do, as a sports team analogy similar to ones used by Blackburn and Valerdi (2007, 2008, 2009) can make plain.

Assume its rabid fans to be, in some context, the principal stakeholders of an American professional baseball team. Also consider that same group - not unlike the rabid fans of many sports teams in many countries - to demand that their club's coming season culminate with a championship or that the team at least perform in a manner the fans could shamelessly defend as "championship caliber." If championship ways did not emerge during the course of the campaign, the resulting collective disappointment would represent the fans' judgment of a failed season or, even more precisely, of a ball club (system) that failed to meet immutable expectations. The need for championship play would, then, represent a stakeholder-recognized COI, one of possibly several fan-desired, emergent team capabilities

termed "show stoppers" by Sproles (2002, p. 256) that, if not achieved, by definition render tested systems of interest - the baseball team in this case - as deficient solutions to the system stakeholders' needs for championship ball or other things. Similarly, stakeholders of some public or private organization might identify a COI regarding a need for their organization to improve a particular operating characteristic vis-à-vis its competitors or its public charter; if whatever might be proposed to achieve that does not yield the desired improvement, then the proposal would have to be judged a failed solution, and the critical issue would remain. Additional and like analogies can explain the role that measures of effectiveness play in determining whether or not critical itches – COI – get scratched.

3.2.2. Measures of Effectiveness

Measures of effectiveness may be defined as (Sproles, 2000a, 2000b, 2001, 2002; Stevens, 1979):

Standards derived from problem stakeholder-determined COI, independent of solutions proposed for problem resolution but representing emergent properties that proposed solutions must exhibit if they are to be judged by stakeholders as having satisfied critical needs.

What should normally be a small number of COI (truly *critical* issues should be expected to be small in number) assigned to any specific need for improvement or problem resolution should each be accompanied by a similarly small number of MOE (Sproles, 2002) that stakeholders then use to make binary, "yes" or "no," determinations of whether or not COI have been satisfactorily resolved. Measures of effectiveness are variables that may be viewed as sliding scales of the "standards" highlighted in the definition above; and because they are drawn only from COI, they have no derivational links to solutions proposed to dispel problems or otherwise afford wanted improvements. The worth of MOE to organizational performance measurement, like COI, may also be explained with a sports metaphor.

Fans hoping to watch their clubs demonstrate championship play on the baseball field might expressly or tacitly judge their teams in terms of "runs scored per game," "team batting average," "number of *Golden Glove* awards earned" for defensive play, the "earned run average" of the team's pitching corps, or other traditional markers – all obviously variable – of baseball excellence. A high average number of runs scored per game or high number of *Golden Glove* awards emerging from a season of play might prompt even the most strident of "championship or bust" fans to adopt a "wait 'til next year" posture; if so, then MOE such as "runs scored per game," "team batting average," "team chemistry," and other qualitative or quantitative standards would have in some way collectively demonstrated the fans to have given their allegiance to a championship caliber ballclub and that, consequently, one of their lives' most critical issues had been resolved!

Measures of effectiveness are as useful to the domains of public and private organizations as they are to the domains of baseball and OT&E. Indeed, the reasoning behind certain of Behn's (2003) eight purposes – such as "to motivate," "to promote," "to celebrate," and "to learn" – in many ways reflects emergent properties able to serve as MOE or to point to other properties that might so serve. To offer a simple but hopefully nonetheless instructive example, a measurement purpose like Behn's "to motivate" readily translates to an effectiveness measure of "degree of employee motivation" that represents a desired emergent property induced (or not) by organizational design and which with other MOE could be used to assess resolution of a COI regarding, for example, a need to be recognized as one of the world's most productive companies. An MOE such as "handling quality" could similarly support a stakeholder judgment of whether or not a new line of automobile would satisfactorily address a critical need for crisp sales, just as baseball fans might use MOE such as "team batting average" and "team chemistry" to determine if their essential, championship capability had been achieved.

3.2.3. Measures of Performance

The one-time New York Yankee catcher and American baseball icon, Yogi Berra, is popularly credited with the observation that "it ain't over 'til it's over," words true enough for baseball and just as true in serving to dispel the confusion that often accompanies the concepts of MOE and MOP (Sproles, 2000a, 2001, 2002). Yogi may be said to have unintentionally employed the "over" of "it ain't over" as an MOE describing whether or not the Yankees still had a chance to win the baseball game in question; the "over" of "'til it's over" recognized that the organized system Yogi viewed as a ballgame always concludes in accordance with rules calling for nine innings of play not yet

completed at the point in the game when he uttered his famous phrase. Effectiveness measures are variables and performance measures are precise evaluations of some system's performance as measured by or against associated MOE. The "measures" of "measures of effectiveness" and the "measures" of "measures of performance" therefore sound identical but mean very different things. Here, again, Sproles' (2000a, 2000b, 2001, 2002) work prompts a useful definition:

Measures of performance are evaluations of systems' intrinsic functions by which can be judged, using MOE, those systems' potential for resolving the critical problems or providing the essential capabilities represented with COI.

Unlike the stakeholder-oriented variables that are MOE, MOP are not variables but evaluations of system-intrinsic performance properties demonstrated (Sproles, 2000a, 2000b, 2001, 2002); neither concept may be of much value without the other, but they are quite different. That is why baseball fans might, for example, might establish MOE such as "runs scored per game," "team batting average," "number of *Golden Glove* awards earned" for defensive play, and the "earned run average" of the team pitching corps that they ultimately use to determine the championship caliber of a team demonstrating respective, season-final MOP of 3.2 runs per game, a 0.283 team batting average, two *Golden Glove* awards, and a team earned run average of 3.89. In making such determinations, the fans, coaches, owners, sportswriters and broadcasters, and other team stakeholders would be employing team-demonstrated MOP in concert with their own previously established MOE and COI to determine in strict "yes" or "no" fashion if their team had displayed the stuff of champions during the season.

Performance measurement systems applied to determine resolution (or not) of public or private organization's COI would closely mirror those used by baseball fans. The COI and derivative MOE would be determined by appropriate stakeholders, and MOP suited to the MOE would be determined, collected, and analyzed for merit. Quantitative MOE would generally call for the acquisition of quantitative MOP, while qualitative MOE such as motivation (a construct in the vein of team chemistry) might call for some combination of qualitative and quantitative MOP or even the translation of qualitative MOP to quantitative terms. Such a translation process could often prove quite useful (Tangen, 2004), particularly if the measurement approach illuminated with this section were applied within the arena of multi-criteria performance measurement.

4. Evaluation Approach

The authors have developed an OT&E- and multi-criteria decision analysis-derived approach to the common evaluation of public and private enterprises. Figure 1 displays that approach.

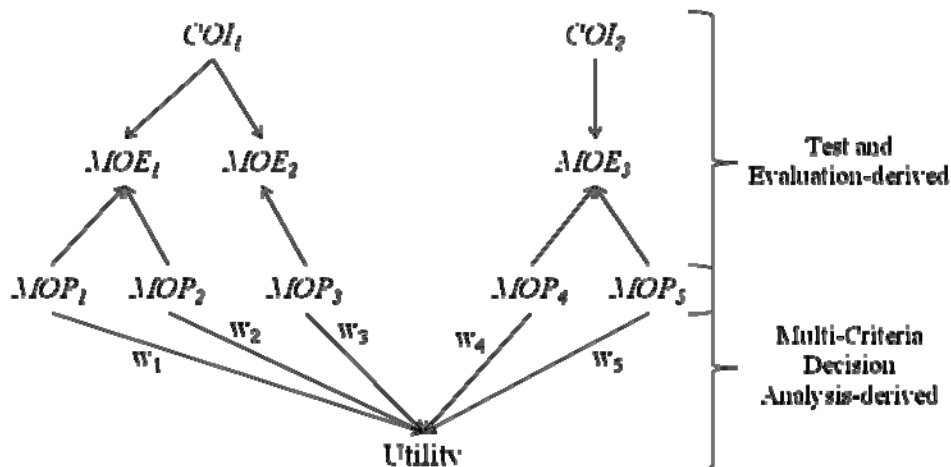


Figure 1: Approach to Assessing Enterprise Utility

For any particular enterprise, COI and derivative MOE must first be identified in accordance with the guidance respectively provided in Sections 3.2.1 and 3.2.2. Enterprise stakeholders should identify critical operational issues respecting what the authors maintain to be sector-independent concerns of mission, investment, or others such as those discussed in Section 2. Once an appropriately small number of COI has been identified, stakeholders should

next derive from each an equally appropriate number of effectiveness measures. The COI identification and MOE derivation processes will together constitute the set of emergent properties desired by stakeholders to alleviate certain organizational problems or to provide certain organizational capabilities, and they will most importantly for PMS purposes represent the gauges against which stakeholders will ultimately evaluate their organization's performance. By then identifying measures of organizational performance to be evaluated – likely not exclusively and just as likely in many-to-one fashion – using established MOE, stakeholders will have fully instantiated the Section 3 OT&E themes and set themselves on a traceable and defensible course for judging whether or not their critical issues will have been resolved. The following identification of needed, enterprise MOP will represent the full realization of Figure 1's upper, or "Test and Evaluation-derived," portion, and it will also represent commencement of the figure's lower portion elements and processes emphasizing a multi-criteria decision analysis effort with which organizational performance can be quantitatively evaluated in accordance with procedures detailed next.

Figure 1 depicts a novel fusion of OT&E concepts with those of multi-criteria decision analysis (MCDA). That fusion allows MOP to be combined into a single evaluation – in the figure generally termed *Utility* and hereafter more specifically termed *Performance Value* – that can serve relevant stakeholders' cumulative desires regarding organizational performance or similar notions of desirability, satisfaction, or such. *Performance Value* may be mathematically expressed as follows:

$$\text{Performance Value} = \sum_{i=1}^n w_i MOP_i \quad (1)$$

where n is the total number of MOP_i and w_i is the weight (or importance) assigned to the i^{th} MOP. The evaluation of a combined measure such as *Performance Value* will require a method for simultaneously determining the combinations of all MOP_i that maximize its value. It may be assumed that the relative importance (i.e., weights) of each MOP_i is not known with certainty and that MOP values can therefore not be easily combined to create a single objective, per equation (1). If, however, the weights associated with each MOP are assumed by appropriate parties to hold particular values (as might be done with a method such as swing weighting), those assumptions and accompanying weight value assignment must precede the determination of MOP values in order not to bias performance value calculations by, for example, associating high weights with high-valued MOP. All MOP must be normalized in order to ensure their combination against a standard scale allowing maximization of individual performance measures; with that maximization stricture driving reformulations, whenever necessary, of MOP as (1 - MOP) or similar constructs that conform with multi-criteria conventions embodied in equation (1).

An organization's *Performance Value* must also be treated as an ordinal rather than a cardinal measure because stakeholders should restrict their performance-related interests to the relative merits of competing enterprise configurations or designs and avoid wrongly placed emphases on seductive but wholly inapplicable absolute values. Enterprise configurations may nonetheless be manipulated to adjust MOP values or to introduce new or discount old MOP. Such manipulations would, of course, drive new evaluations of *Performance Value* and so serve stakeholders' desires for enterprise improvement, however possible. A legitimately maximized *Performance Value* of real worth to stakeholders will result from a mathematically-derived best mix of MOP and associated weight values.

Obtaining optimal mixtures of MOP requires the application of rigorous optimization techniques and the determination of what will normally be initially unknown weights of equation (1); this renders schemes like the weighted sum approach as natural candidates to support *Performance Value* optimization. With the weighted sum approach, a function of two or more objectives (MOP in this case) is combined into a single objective by constructing a weighted sum of the original, multiple objectives and formulating it as an optimization problem of the form:

$$\max \text{Performance Value} = \sum_{i=1}^n w_i \text{MOP}_i \quad (2a)$$

s. t.

$$0 \leq w_i \leq 1 \quad (2b)$$

$$\sum_{i=1}^n w_i = 1 \quad (2c)$$

This formulation can be optimized using standard techniques once weight values have been determined and, as noted earlier, the implications of knowingly or unknowingly assigning weights that illegitimately alter *Performance Value* must be recognized and avoided. The individual MOP, themselves, may be evaluated using any number of defensible techniques such as direct measurement or simulation as well as with the analytic hierarchy process, or AHP, of Saaty (1980), evidence theory (Dempster, 1976; Shafer, 1976), fuzzy techniques (Zadeh, 1965), or Likert scale methods (Likert, 1932), all particularly useful for portraying qualitative MOP in the numeric terms required for MCDA.

The assignment of weight values must respect the manner in which MOP pertinent to any particular organizational need are identified. Optimization formulations must ensure that no incorporated MOP is favored over another, and this stipulation largely stems from the manner in which MOP are identified; the MOP of a specific formulation have been most fundamentally identified in response to previously established COI, or Sproles' (2002, p. 256) "show stoppers." Each MOP could then, itself, possibly stop or at least severely impair the formulation unless paired with a legitimate weight value, and full-factorial design techniques offer a defensible way to promote just such pairings.

Full-factorial designs require a (number of factor levels)^(number of factors) number of configurations. For example, a problem with three MOP (factors) and two (weight) levels of *high* and *low* requires $2^3 = 8$ configurations and thus could be evaluated using weight values linked to the configurations of Table 1. Note Configuration 8, the null solution to the problem, to be a trivial case because the across-the-board assignment of zero weight values will yield no *Performance Value* formulation solution of worth. All full-factorial designs include such a null solution.

Table 1: Full Factorial Weights for Three MOP and Two Weight Levels

Configuration	w_1	w_2	w_3
1	1/3	1/3	1/3
2	1/2	1/2	0
3	1/2	0	1/2
4	1	0	0
5	0	1/2	1/2
6	0	1	0
7	0	0	1
8	0	0	0

Given each Table 1 configuration of weights, an optimal set of MOP can be calculated using the equation (2) problem formulation, and that is typically termed a Pareto optimal set (Pareto, 1971). A set of points is said to be Pareto optimal if, in moving between any two points in the set, an improvement in the value of one objective (MOP) causes the value of at least one other objective (MOP) to worsen. Theoretically, Pareto optimal sets yield infinite numbers of solutions, and so increased discretization of the weights (i.e., more levels of discretization) must be used to ensure a correspondingly more precise optimal solution. The increased discretization will, however, simultaneously increase the optimization problem's computational complexity, as more evaluations of *Performance Value* will be required. Determining the appropriate level of weight discretization is a subjective undertaking.

Example Problem

The concepts presented in this section are best illustrated through an example problem spanning public and private sector enterprises. Morash (2000) discusses transportation planning and public policy and their associated effects on private sector performance through investment policy, financing arrangements, tax policy, and infrastructure improvement. The authors utilize Morash's results to build a notional assessment of a small shipping company.

Suppose we assume that COI – "show stoppers" – for the shipping company are 1) maintaining customer satisfaction and 2) maintaining financial solvency. Utilizing what Morash (2000) identifies as the most company's important performance characteristics (which we term as effectiveness measures), we can identify the associated MOE and MOP depicted in Figure 2.

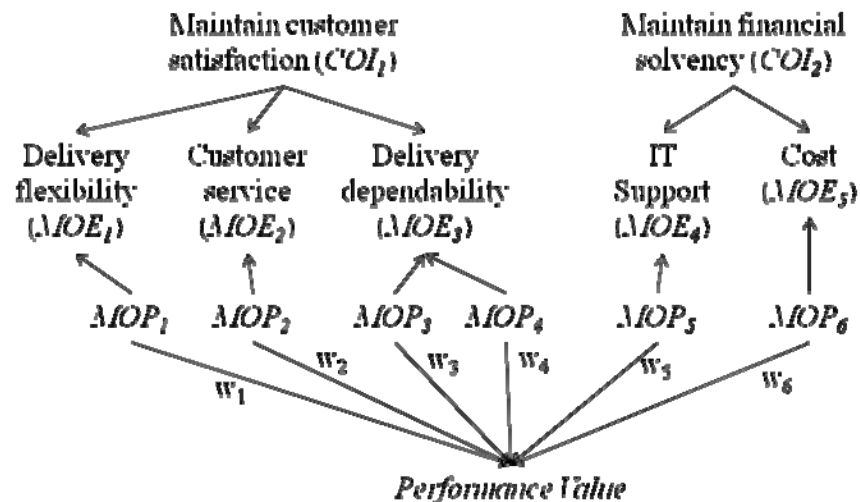


Figure 2: Example Problem COI, MOE, and MOP relationship

Those MOP associated with quantitative MOE, such as MOE_5 , can be evaluated directly. The MOP_6 should thus be evaluated, like cost, in terms of dollars, though its construct must be altered to match other MOP constructs for which maximization will be preferred; neither the MCDA scheme nor organization stakeholders would wish to maximize cost. To likewise respect other requirements of MCDA, quantitative MOP can be developed that provide for the evaluation of qualitative MOE as they also meet MCDA needs. The MOP_3 and MOP_4 , for example, may be respectively conceived as mean percentage of on-time deliveries and variance of that same percentage. The complete set of MOP, combined with a set of corresponding weights, can be utilized to evaluate the shipping company's *Performance Value* by using Eq. (1) and the approach detailed in this section. That evaluation will represent an ordinal measure with which the shipping company owner, its investors, or other stakeholders may judge company performance in terms of the concerns, fundamentally the COI, that they view as most critical.

Consider an additional scenario wherein the company owner chooses to evaluate the potential of a capital investment possibly to be made to improve company performance; any improvement potential could be gauged in a manner like that evident with the previous example. Given the solution-independent nature of COI and MOE, those emergent properties would remain just as desirable as before. However, given their solution-dependent nature, MOP *may* change. That is, the number of MOP may increase, decrease, or remain as is, and their particular evaluations may also change due to a new company configuration induced by the anticipated investment. Stakeholders could apply the framework of this section to MOP identified for the capital investment contemplated, with that application then evaluate relative to one another the *Performance Value* of all company configurations expected to derive from the investment, and lastly determine what if any actions should follow the evaluation. Thus, if the investment represents an improvement in company performance, the organization may wish to pursue it.

Both of these examples highlight anticipated benefits offered with the author's proposed fusing of OT&E and MCDA to yield a multi-criteria performance measurement system equally suited to public and private enterprises. The authors believe their approach allows for justifiable decision-making that well captures enterprise desires for improvement, regardless of public or private sector affiliation.

5. Conclusions

This chapter developed a systemic approach to evaluating organizational performance through the novel fusion of operational test and evaluation and multi-criteria decision analysis. This combined approach provided a framework by which decisions can be made regarding organization performance and potential for improvement. Use of this

approach was demonstrated on a notional example problem taken from the transportation industry. The approach is sector-agnostic, incorporating sector-spanning notions of mission and investment, and it represents a novel manner of performance evaluation which can be utilized regardless of an organization's publicness or other likeminded categorization. Given the foundations upon which this approach is built, the authors welcome others to adopt their Section 4 framework so as to improve organizational performance measurement and associated PMS in a manner scientifically defensible and directly derived from critical operational issues identified by relevant stakeholders as crucial to organizational success.

Acknowledgement

This study was supported by funds from the U.S. Department of Homeland Security's Science and Technology Directorate, Borders and Maritime Division. The authors gratefully acknowledge the support received from this source.

References

- Baldwin, J. N. (1987). Public versus private: Not that different, not that consequential. *Public Personnel Management*, 16(2), 181-193.
- Behn, R. D. (1995). The big questions of public management. *Public Administration Review*, 55, 313-324.
- Behn, R. D. (2003). Why measure performance?: Different purposes require different measures. *Public Administration Review*, 63, 586-606.
- Besley, T., & Ghatak, M. (2005). Competition and incentives with motivated agents. *The American Economic Review*, 95(3), 616-636.
- Bhattacharyya, S. K. (1968). Performance of public undertakings: A Framework for appraisal. *Economic and Political Weekly*, 3(21), M25+M57+M29+M31.
- Blackburn, C., & Valerdi, R. (2007). Measuring systems engineering success: Insights from baseball. Massachusetts Institute of Technology.
- Blackburn, C., & Valerdi, R. (2008). *Measuring systems engineering success: Insights from baseball*. Paper presented at the 18th Annual International Symposium of the International Council on Systems Engineering. Retrieved from <http://lean.mit.edu/docman/2008-conference-papers/view-category.html> (April 18, 2011)
- Blackburn, C., & Valerdi, R. (2009). *Navigating the metrics landscape: An introductory literature guide to metric selection, implementation, & decision making*. Paper presented at the 7th Annual Conference on Systems Engineering Research. Retrieved from <http://cser.lboro.ac.uk/papers/S04-30.pdf> (April 11, 2011)
- Bourne, M., Mills, J., Wilcox, M., Neely, A., & Platts, K. (2000). Designing, implementing, and updating performance measurement systems. *International Journal of Operations and Production Management*, 20, 754-771.
- Bourne, M., Neely, A., Platts, K., & Mills, J. (2002). The success and failure of performance measurement initiatives: Perceptions of participating managers. *International Journal of Operations and Production Management*, 22, 1288-1310.
- Boyne, G. A. (2002). Public and private management: What's the difference? *Journal of Management Studies*, 39(1), 97-122.
- Bozeman, B. (1987). *All organizations are public*. San Francisco: Jossey-Bass.
- Bozeman, B., & Bretschneider, S. (1994). The "Publicness Puzzle" in organization theory: A test of alternative explanations of differences between public and private organizations. *Journal of Public Administration and Research Theory*, 4(2), 197-223.
- Bozeman, B., & Kingsley, G. (1998). Risk culture in public and private organizations. *Public Administration Review*, 58(2), 109-118.
- Campbell, P. L. (2004). *Measures of effectiveness: An annotated bibliography (Sandia National Laboratories Report SAND2004 2902)*. Albuquerque, NM: Sandia National Laboratories.
- Chiesa, V., Frattini, F., Lazzarotti, V., & Manzini, R. (2008). Designing a performance measurement system for the research activities: A reference framework for an empirical study. *Journal of Engineering and Technology Management*, 25, 213-226.
- Chiesa, V., Frattini, F., Lazzarotti, V., & Manzini, R. (2009). Performance measurement in R&D: Exploring the interplay between measurement objectives, dimensions of performance and contextual factors. *R&D Management*, 39, 488-505.

- Coursey, D., & Bozeman, B. (1990). Decision making in public and private organizations: A test of alternative concepts of "publicness". *Public Administration Review*, 50, 525-535.
- Dahl, R., & Lindblom, C. (1953). *Politics, economics, and welfare*. New York: Harper & Row.
- Dempster, A. P. (1976). Upper and lower probabilities induced by a multivalued mapping. *The Annals of Statistics*, 28, 325-339.
- Fitzgerald, L., & Storbeck, J. E. (2003). Pluralistic views of performance. *Management Decision*, 41, 741-750.
- Flynn, D. M., & Tannenbaum, S. I. (1993). Correlates of organizational commitment: Differences in the public and private sector. *Journal of Business and Psychology*, 8(1), 103-116.
- Fottler, M. D. (1981). Is management really generic? *The Academy of Management Review*, 6(1), 1-12.
- Gimbert, X., Bisbe, J., & Mendoza, X. (2010). The role of performance measurement systems in strategy formulation processes. *Long Range Planning*, 43, 477-497.
- Gold, K. A. (1982). Managing for success: A Comparison of the private and public sectors. *Public Administration Review*, 42(6), 568-575.
- Griffin, A., & Page, A. L. (1996). PDMA success measurement project: Recommended measures for product development success and failure. *Journal of Product Innovation Management*, 13, 478-496.
- Hedley, T. P. (1998). Measuring public sector effectiveness using private sector methods. *Public Productivity & Management Review*, 21(3), 251-258.
- Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard-Measures that drive performance. *Harvard Business Review*, 70(1), 71-79.
- Kaplan, R. S., & Norton, D. P. (1996). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, 74(1), 75-85.
- Kerssens-van Drongelen, I. C., & Bilderbeek, J. (1999). R&D performance measurement: More than choosing a set of metrics. *R&D Management*, 29, 35-46.
- Kerssens-van Drongelen, I. C., & Cook, A. (1997). Design principles for the development of measurement systems for research and development processes. *R&D Management*, 27, 345-357.
- Knott, J. H. (1993). Comparing public and private management: Cooperative effort and principal-agent relationships. *Journal of Public Administration Research and Theory*, 3(1), 93-119.
- Kouzmin, A., Löffler, E., Klages, H., & Korac-Kakabads, N. (1999). Benchmarking and performance measurement in public sectors: Towards learning for agency effectiveness. *The International Journal of Public Sector Management*, 12, 121-144.
- Kuin, P. (1968). Management performance in business and public administration. *The Academy of Management Journal*, 11(4), 371-377.
- Kurland, N. B., & Egan, T. D. (1999). Public v. private perceptions of formalization, outcomes, and justice. *Journal of Public Administration Research and Theory*, 9(3), 437-458.
- Lachman, R. (1985). Public and private sector differences: CEOs' perceptions of their role environments. *The Academy of Management Journal*, 28(3), 671-680.
- Lan, Z., & Rainey, H. G. (1992). Goals, rules, and effectiveness in public, private, and hybrid organizations: More evidence on frequent assertions about differences. *Journal of Public Administration Research and Theory*, 2(1), 5-28.
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, 140, 1-55.
- Micheli, P., & Manzoni, J. F. (2009). Strategic performance measurement: Benefits, limitations and paradoxes. *Long Range Planning*, 43, 465-476.
- Morash, E. A. (2000). Linking public and private performance measurement. *Journal of the Transportation Research Board*, 1729, 42-50.
- Murray, M. A. (1975). Comparing public and private management: An exploratory essay. *Public Administration Review*, 35(4), 364-371.
- National Performance Review. (1997). *Serving the American public: Best practices in performance measurement (Benchmarking Study Report of June 1997)*. Washington, D.C.
- Neely, A. (1999). The performance measurement revolution: Why now and where next? *International Journal of Operations and Production Management*, 19, 205-228.
- Neely, A., Mills, J., Platts, K., Gregory, M., & Richards, H. (1996). Performance measurement system design: Should process based approaches be adopted? *International Journal of Production Economics*, 46-47, 423-431.
- Nutt, P. C. (2005). Comparing public and private sector decision-making processes. *Journal of Public Administration and Research Theory*, 16, 289-318.

- Nutt, P. C., & Backoff, R. W. (1993). Organizational publicness and its implications for strategic management. *Journal of Public Administration and Research Theory*, 3(2), 209-231.
- Pareto, V. (1971). *Manual of political economy*. New York: A.M. Kelley.
- Perry, J. L., & Rainey, H. G. (1988). The Public-private distinction in organization theory: A critique and research strategy. *Academy of Management Review*, 13(2), 182-201.
- Poh, K. L., Ang, B. W., & Bai, F. (2001). A comparative analysis of R&D project evaluation methods. *R&D Management*, 31, 63-75.
- Rainey, H. G. (1989). Public management: Recent research on the political context and managerial roles, structures, and behaviors. *Journal of Management*, 15(2), 229-250.
- Rainey, H. G. (1991). *Understanding and managing public organizations*. San Francisco: Jossey-Bass.
- Rainey, H. G., Backoff, R. W., & Levine, C. H. (1976). Comparing public and private organizations. *Public Administration Review*, 36(2), 233-244.
- Rainey, H. G., & Bozeman, B. (2000). Comparing public and private organizations: Empirical research and the power of the a priori. *Journal of Public Administration and Research Theory*, 10(2), 447-469.
- Ramamurti, R. (1987). Performance evaluation of state-owned enterprises in theory and practice. *Management Science*, 33(7), 876-893.
- Rouse, P., & Puterill, M. (2003). An integral framework for performance measurement. *Management Decision*, 41, 791-805.
- Saaty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw Hill.
- Scott, P. G., & Falcone, S. (1998). Comparing public and private organizations: An exploratory analysis of three frameworks. *The American Review of Public Administration*, 28(2), 126-145.
- Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.
- Sheehan, R. M. (1996). Mission accomplishment as philanthropic organization effectiveness: key findings from the excellence in philanthropy project. *Nonprofit and Voluntary Sector Quarterly*, 25(1), 110-123.
- Simeone, R., Carnevale, J., & Millar, A. (2005). A systems approach to performance based management: The national drug control program. *Public Administration Review*, 65, 191-202.
- Sproles, N. (2000a). Coming to grips with measures of effectiveness. *Systems Engineering*, 3, 50-58.
- Sproles, N. (2000b). *Complex systems, soft science, and test and evaluation or 'Real men don't collect soft data.'* Paper presented at the Systems Engineering Test and Evaluation Conference, Brisbane, Australia.
- Sproles, N. (2001). The difficult problem of establishing measures of effectiveness for command and control: A systems engineering perspective. *Systems Engineering*, 4, 145-155.
- Sproles, N. (2002). Formulating measures of effectiveness. *Systems Engineering*, 5, 253-263.
- Stevens, R. T. (1979). *Operational test and evaluation: A systems engineering process*. Malabar, FL: Krieger.
- Tangen, S. (2004). Performance measurement: From philosophy to practice. *International Journal of Productivity and Performance Management*, 53, 726-737.
- Tirole, J. (1994). The internal organization of government. *Oxford Economic Papers*, 46, 1-29.
- Wamsley, G. L., & Zald, M. N. (1973). *The Political economy of public organizations*. Lexington, MA: Heath.
- Weinberg, M. W. (1983). Public management and private management: A diminishing gap? *Journal of Policy Analysis and Management*, 3(1), 107-115.
- Wilcox, M., & Bourne, M. (2003). Predicting performance. *Management Decision*, 41, 806-816.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and Control*, 8, 338-353.