Action Research Validation of an Inventory of Effectiveness Measures

Roland J. Kushner, Ph.D.
Management Advisory Services
P. O. Box 4213
Bethlehem, PA 18018
610-865-9697
roland@rjkushner.com

Action Research Validation of an Inventory of Effectiveness Measures

ABSTRACT:

In a systems model of organization function, resources are consumed and transformed, creating outputs, which dynamically affect the next iteration of resource use. In a closed systems model, no exogenous effects disturb the system; in an open systems model, environmental factors affect all elements of the system. This paper presents an inventory of performance measures based on an open systems model of organizational effectiveness. In this model, performance is examined by looking at the effects of five separate components: satisfying constituents, mobilizing resources, effectively using these resources, setting and attaining goals, and adapting to environmental change. Each component has an independent effect on overall organizational performance and each one is necessary for an integrated concept of effectiveness, but none by itself is sufficient.

The paper addresses the issues of validity and the ability of managers to appropriately use a measurement model in environments where measures are frequently challenged. Its key contribution is a performance measurement inventory developed to operationalize the systems model of effectiveness in a manner that fits into schemas and scripts of managers. The action research is a multi-year process of iteratively applying and improving the measurement model and its tools. The paper proceeds as follows: first, the issue of validity is framed in terms that connect research and practice; the model of organizational effectiveness is presented; this is then mapped onto an inventory of measures; the process of validating the inventory and its tools is described; measurement techniques are presented that can be used in organizational assessment and organizational development situations; and lessons are presented for research and practice in the measurement process.

Action Research Validation of an Inventory of Effectiveness Measures

Introduction

Nonprofit organization managers generally want to measure how well their agency is doing, and also to know how their agency is doing well. The difference between the two areas of inquiry follows from the framework of "doing well," i.e., performance, as a multi-dimensional phenomenon actually present and (perhaps) visible in an organization. But, for individual agencies and their leaders, the multi-dimensionality in and of itself is a problem that makes most measurement difficult to manage. In effect, performance measurement is bounded because the dimensions of the problem, exceed -- sometimes by a wide margin -- the degrees of freedom available to study a particular data set.

In practical terms, this means that general managers and leaders face difficult challenges in getting meaningful and helpful performance measures. They may want to know answers to certain questions. First, they want to know their organization's overall performance, to be able to answer the question *"How well are we doing?"* This is a systemic question, applied to a whole organization. Second, and simultaneously, they try to understand *"How are we doing well?,"* by examining performance as a disaggregated phenomenon, categorized (for example) by elements such as department, program, budget line, or other means for planning, action, or control. This helps to focus attention on specific issues. Performance assessment at the organization level takes place in context of the strategic management process (Stone, Bigelow & Crittenden 1999), so leaders want to know how to implement the findings of their assessments to plan and act ... and assess again. Critically, they want to do both in their own terms and frames of reference, using schemata and scripts that are particular to their operating and social contexts.

This last point is essential. OE is an individual and social construct (Herman & Heimovics 1994), defined as much by the personal and social situation of managers in place as by any external standards. A manager's view of effectiveness stems from individual careers, life

stages, experiences, their environment, their expectations, and also (clearly) from the social structure and culture of an organization. Language and rhetoric also matter as elements of the measurement process in the realities of an organization.

Scholars and external evaluators have eagerly investigated performance assessment. Though this paper looks at nonprofits, organizational effectiveness is widely studied and reported in mainline business journals, in settings such as ARNOVA, and numerous scholarly journals in the disciplines where organizations are studied. It has been seen, for example, as a component of program evaluation (Leland 1999), accountability (Tassie, Murray, Cutt & Bragg 1996), and strategy (Stone, Bigelow & Crittenden 1999). Yet problems remain in the methodological areas of designing appropriate and tractable measurement tools. Ultimately, researchers want to know that their analyses are valid. But it is a resource-intensive job to create validated performance evaluation tools between agencies because of challenges in creating normalized measures. Validity, as recognized and used in multi-observation research, fits uneasily in the $n = 1$, single-site perspective of performance evaluation in a single agency.

The perspectives of practicing managers are another challenge for researchers. Performance assessment of a whole organization by its leaders is rarely done comparatively; more often, it is an act of reflection, an inward look. To validate measurement tools in this subjective and political setting, it's important that they stem from the language, discourse, and rhetoric that drive management action. When observers, researchers, managers, constituents, and funders share common language, it is easier to reach a consensus on the two questions that opened this paper.

Validity in Concept and Application

The concept of validity relates the underlying meaning of words to how they capture what is true about the construct they describe. Qualifiers used to describe how trustworthy a measure are in Campbell & Stanley (1966) where *internal, external,* and *construct* validities

refer respectively to the abilities of a measure to capture causation in an effects model, to apply to broader populations, and to be sensible on its face. Neumann (1991) adds *criterion* (known to be accurate), *face* and *content* (basically rephrasings of construct validity), *concurrent* (operating in parallel with other accepted construct measures), *predictive* (leading to a construct that it intended to forecast), *convergent* (essentially, reliability, through being positively correlated to others) and *discriminant*, negatively correlated to others. Fornell & Cha (1994) similarly emphasize convergent validity, particularly in the link between latent and observed variables. There, it operates in parallel with the characteristic of being robust, persistent over time.

For purposes of this paper, validity of an organization-level performance measurement instrument means that the instrument usefully informs managers, leaders, and stakeholders in an organization and meaningfully contributes to their ability to improve performance. That is, it has (primarily positive) diagnostic capabilities but contributes to (primarily normative) prescriptions for change. Validity of instruments or measurements to managers must have some key capabilities: they can communicate outside (external validity); they accurately tell their story (criterion); they are efficient procedurally and in the sense of minimizing bias, relatively easy to implement; they are robust, persistent over time; and they provide guidance for change.

These are the *optima* of validity. To understand the validity of making sense of effectiveness, we need to explore the fluid perceptions embodied in the concept of performance. To managers, the terms may cloud more than illuminate basic truths of nonprofit performance measurement. The word "effective" may appear impenetrable, an insoluble puzzle, particularly for small agencies where performance measurement is mainly responsive to external funders. For program-driven funding, performance measurement focuses on program level outputs, and only rarely on overall capacity. Resources are frequently lacking to evaluate overall organizational effectiveness according to any model.

Ultimately, effectiveness measures are used by managers to communicate, within organizations and with external stakeholders or opponents. But language exists in organizational contexts, and neither language nor context are unitary. There are few environments with unambiguous attributes known to all observers. They operate in environments that they "enact" (Weick 1998), by focusing on and interpreting particular elements of their observation that they deem to be particularly important, and founding their actions on those elements. The environment is individual to each actor, is in effect created by each actor as they attempt to make sense of it.

In dealing with individual perception, individuals tend to organize their knowledge using schemata, cognitive models that help to give structure and meaning to their expectations about social and organizational phenomena (Poole, Gray & Gioia 1990), such as "measurement" or "effectiveness" or "quality." While schemata are cognitive, they are employed in bilateral and/or multilateral organizational contexts, such as meetings. These become the contexts for a specific kind of schema: scripts. Scripts are dynamic frameworks that individuals enact to process their communications (Gioia & Poole 1984, Poole 1989). When effectiveness measures become part of an organization's discourse, it is a signal that they are organized in schemata and scripts. These organizing structures, often pre-determined and rooted in habit, have powerful influences over performance measurement. Managers tend to look repeatedly at the same measurements, interpret them in the same manner, and assume (or at least hope) that their validity remains constant even as the organization and the environment that give them meaning are in constant flux.

Some of this is natural, a heuristic or habit of measurement, just as scripts are heuristics for intra-organizational communication. And because rationality is limited, it is natural for managers to try to package their performance assessments into more compact forms so that they do not always have to re-invent and redesign performance assessment each time. Compact is not *per se* a problem; surely there are virtues to efficiency in examining organizational effectiveness. What might be more questionable is that in enacting their

particular environment, managers with intra-organizational and interpersonal scripts of limited breadth may expect too much comprehensive information out of measures that are too compact or too large, too narrow or too broad.

What is a Valid Theory of Effectiveness?

Over time, we have seen theories develop that emphasize one or another element of organization performance as being of prime importance. From early administrative science writings to the present, management writers typically use measures of organizational performance as independent, dependent or classification variables. Scholars look in micro and macro organizational study for evidence of performance, quality, success, efficacy, approval, achievement, improvement, satisfaction, efficiency, or some other contextually-determined indication to which to attach a "positive" connotation. Historically, goal attainment was the *ne plus ultra,* the received wisdom of performance in both concept and practice. Alternative theories and themes have emerged as management theory has been influenced by contingency theories, Japanese corporate practice in the 1980s, systems thinking, the quality movement, and stakeholder and network models of management. As thinking about what makes an organization has evolved, so too has the theory of what makes a good organization. Instead of a dominating theory that is accepted as valid, a multitude of approaches contend for primacy. Not surprisingly, some of the contention comes from the different disciplines where theory and empirical study occur, (e.g., operations research, industrial psychology, economics / finance, social interaction, political science, public administration, industrial psychology).

The Open Systems Model of Organizational Effectiveness

Given all of the above, a single theory of effectiveness is unlikely to be held as valid by all managers (to say nothing of researchers). However, organization theory has evolved from deterministic models of organizations towards holistic approaches that perceive organizations as networks and systems. These newer models create a venue for an inclusive and wide- ranging approach to effectiveness. The open systems model of organizational

effectiveness, shown in Figure 1, is presented as such a generalizable conceptual model for structuring measurements of organizational performance (Kushner 1994, Kushner & Poole 1996, Kushner 1999). Reflecting the multi-dimensional nature of performance, it is built on five separate "components:"

● How organizations identify and satisfy constituents, resource owners and others

● How they mobilize resources by identify and acquiring them[1]

● How those resources are used internally

● How they define and achieve their desired outputs

● How they respond and adapt to environmental change

It is important to note that none of them are treated as inherently more or less important. Each affects overall performance both directly and indirectly. None are sufficient; all are necessary. They conceptually frame organizations as open systems, sets of interconnected parts, that compete in changing environments. Because the model portrays constituents as both causing and being caused by overall performance, the model is dynamic, changing over time. No single component dominates, because each is affected by others and affects others, too. In fact, it is more inclusive, making it possible to pay attention and measure all components, making it possible to fit into the scripts and schemata of more managers.

This paper proceeds by applying the model to performance measurement by describing how the components were used in various nonprofit settings.

---

[1]The model and measurement process have developed over a period of years. In early iterations "resource mobilization" was called "resource acquisition." This earlier nomenclature appears in some exhibits.

Figure 1 An Open Systems Model of Organizational Effectiveness
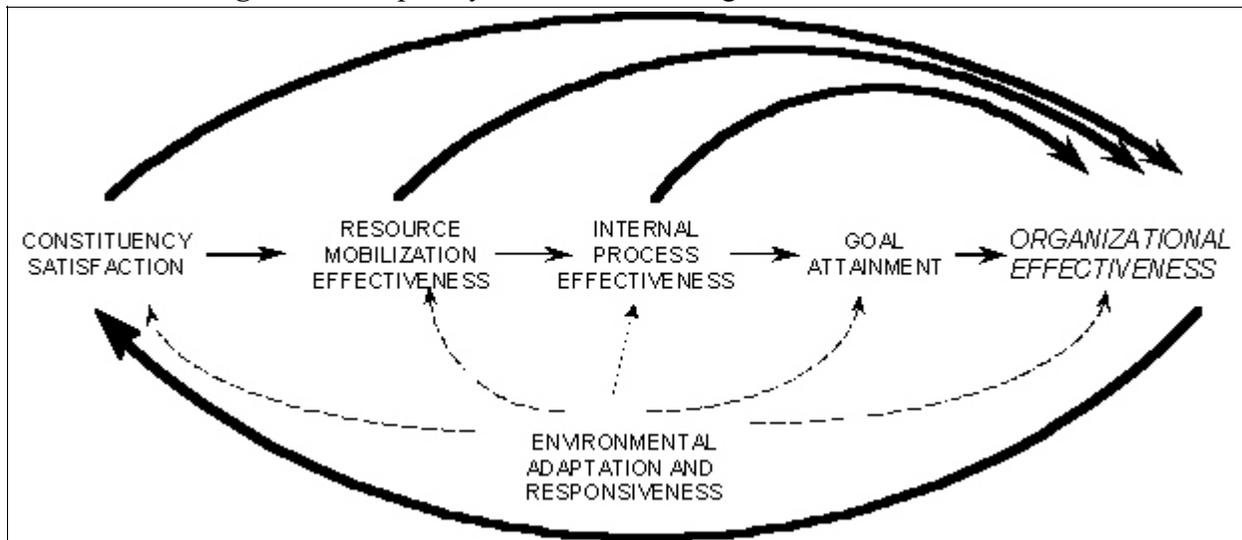
The Action Research Process

From 1994 to the present, I have had opportunities as a consultant and researcher to see how this model has worked, and sometimes not worked. The reports that follow include experiences implementing a methodology in performance evaluation and planning while providing consulting services in a variety of settings in nonprofit organizations. Most have been in one region of the U.S., but others are national or global in scope. Several describe clients of a university-based management support organization in an Eastern U.S. metropolitan area. The action research report is primarily on activities from 1999 to 2001. As is often the case in small sample research, external validity (in the classic sense) of conclusions is not based on random selection of a wide data set, but rather on readers' acceptance of a self-conscious report that relates findings to the underlying research issue.

Implementing the effectiveness model, and making it a useful tool for managers have been persistent themes in my work. Successive engagements produced additions to the record of how the systems model works as well as numerous ways in which it did not. Over time, the systems model led to a measurement model based on response and comment from its users. The process has benefitted by collaborators acting as clients, consultants, trainers, students,

and researchers. Some settings where the model gave rise to a measurement instrument are in Table 1.

Table 1. Pseudonyms and Settings of Action Research Sites

Organization	Setting where the systems model played a role
Health: Multi-state, federally funded STD program run by states.	Strategic planning retreat and write-up
Society: Five divisions of a global engineering professional society	Strategic planning retreat and write-up
Environment: Multi-county conservancy	Framework for "strength and weakness" within strategic planning
Faith: Metropolitan human service organization	Built in to strategic planning methodology as "Planning for Performance" in planning and training
School: Multi-county special education provider	Overall organizational assessment as a diagnostic in restructuring
ARC: U.S.-based humanitarian organization	Developing organizational performance assessment methodology for field use

The last can be explicitly identified as American Red Cross, in connection with its work in the International Red Cross / Red Crescent Movement. *ARC* has adopted the systems model as a framework for its organizational development efforts assisting other Red Cross and Red Crescent National Societies. *ARC*'s use has been extensive, but the systems approach has also been used in other settings, and knowledge from other organizations has helped it develop. The following paragraphs describe that development, in order of complexity rather than chronologically.

In three organizations, the components were used at a summary level to focus attention on key elements of concern during strategic planning. In *Health* and *Environment*, the components were used as frameworks for two-hour discussions of Board, members, and staff

regarding strengths and weaknesses. With *Health*, the Environmental Adaptation Component was not included so four components were used. In both situations, the components framed a fairly open discussion with one or more external facilitators. In *Health*, there were approximately 30 discussants; in *Environment*, there were about 15.

In *School*, top managers were presented the five components and asked to evaluate them on a 7-point rating scale as part of a questionnaire in a day-long organization self-assessment. The questionnaire used is shown in Figure 2, and the results are in Table 2.

Figure 2. Questionnaire used for *School* (excerpt)

PLEASE READ THE SECTIONS ON THIS PAGE AND PREPARE TO MEASURE *School* IN THE FIVE AREAS OF PERFORMANCE. USE A SCALE OF 1 - 7, WHERE 1 = LOWEST EFFECTIVENESS AND 7 = HIGHEST EFFECTIVENESS, OR "DK" = "DON'T KNOW"

Satisfying all constituents

Basically: *School*'s effectiveness in keeping people happy.

Examples of constituents are: clients, families, funders, staff, Board, community, society, professions.

1. Your measurement of how effective *School* is in this area:_____

Mobilizing the appropriate resources

Basically:*School*'s effectiveness in getting what is necessary from the outside environment.

Example of resources are professional skills, facilities, financial resources, leadership, work force.

2. Your measurement of how effective *School* is in this area:_____

Effective internal process:

Basically: Using resources appropriately inside the organization.

Examples are areas like service delivery, communications, HR, efficient operations, accountability, site management

3. Your measurement of how effective *School* is in this area:_____

Achieving goals:

Basically: *School*'s effectiveness in achieve objectives.

Examples are: service goals, financial goals, budget goals, long-term goals, having a good planning process.

4. Your measurement of how effective *School* is in this area:_____

Adapting to the Environment:

Basically: *School*'s effectiveness in recognizing and responding to outside change

Examples: Being successful after changes in constituents, where resources come from, when there are opportunities to be more efficient, the organization's ability to learn.

5. Your measurement of how effective *School* is in this area:_____

Table 2. Mean Responses to *School* Questionnaire

How effective is *School* in specific areas? A scale of 1 to 7, 7 = best, n = 26

a. Constituency Satisfaction	4.00	All t-statistics are > 2.25
b. Resource Acquisition	3.64	
c. Internal Process	2.96	
d. Goal Attainment	3.77	
e. Environmental Adapation	4.60	
Average of all Components	3.83	

It is worth noting that most respondents have graduate degrees and are familiar with evaluation instruments. The low scores they awarded their own agency were a contributing factor to a restructuring process. The relative ranking of the five scores correlated positively with other expressions of their feelings about their work. The fact that "Internal Process" received the lowest score was telling to managers and leaders in the organization, who instinctively sensed difficulties in these areas but needed verification.

A single rating scale measurement such as was used in *Health* masks much important detail. The other three cases were more complex, and needed more involved activities to customize the systems model and make it more precise, and especially to conform it to the terms that the managers understood (i.e., their scripts). This occurred through the development of "System Performance Drivers." This term came about through the process of working with *ARC*. *ARC* accepted the systems model as a conceptual framework and sought to make of it a measurement instrument as part of an organizational development toolkit for use by their overseas delegates. Their specific area of application is that *ARC* provides OD support to numerous Red Cross and Red Crescent National Societies around the world, and has been seeking to improve the knowledge base for OD activities. Performance assessment is, naturally, a key element. *ARC*'s specific need was for ways to move from the component level towards an operational guide to performance that was precise, could be accepted and used by National Societies and *ARC* delegates and could persist over time ... that is, a valid instrument.

Some explicit assumptions were made: The detailed instrument should target areas where managers could plausibly be expected to have knowledge and responsibility, for accountability and to link the assessment to the strategic management cycle. Each of the components ought to have some weight in the overall assessment. The performance assessment ought to prioritize management actions for the future. Assessors in place ought to have the means to measure key areas in terms they believed were meaningful.

The Inventory of System Performance Drivers

The instrument evolved through several steps. The first was to lay the individual areas within each of the five broad components out in more detail. The "System Performance Drivers" constitute an inventory, a series of management decision areas that individually drive performance, both as elements of one component or another, and as contributors to systemic organizational effectiveness. The first iteration of this inventory was developed for *ARC*, and has since continued to evolve; a recent iteration of the SPD list follows in Figure 3. The evolution has been affected by users, but certain themes have persisted:

● Each component has approximately the same number of performance drivers, so that no single component was seen as privileged
● The drivers reflect a managerial choice approach to organizational outcomes; based on a premise that strategic choice and implementation, not environment, are what drives performance. Success is not exogenous but results from what managers do.
● To the extent possible, they are framed at the same overall level of detail
● To the extent possible, the list of drivers is collectively exhaustive (i.e., every significant element of organizational decision-making is contained somewhere in the list)
● To the extent possible, the drivers are mutually exclusive

Organizations where the SPDs could be used were qualitatively different from the three agencies that used the components at a summary level. In *Society*, the setting was a full day planning session for technical associations of a global (primarily US) society of technical / engineering professionals. Two divisions met in a roundtable format in the same room on

one day, and then three the next, for a total of five such groups. Additional training was provided to two facilitators who conducted and recorded the other meetings. Figure 4 shows the SPD list that *Society* used.

Because only a short time was allocated to these sessions (fit into a technical conference), the groups of ten to fifteen discussants found different ways to implement the SPD evaluation. In this case, as in *Environment* and *Health*, the SPDs were introduced as the "strengths and weaknesses" element of situation analysis in strategic planning. They were asked to identify a set of five or six most important issues from the SPD list. Different groups approached this in different ways. One theme was to measure performance in a three-point scale (high, middle, low). Others (more often) used them as cues for discussion, resulting in a consensus list of key performance areas, but overall a less precise measurement.

For *ARC*, the SPDs were set in a facilitators guide for organizational development delegates (Kushner 2000). To make it a tractable tool for managers, additional support was provided in how to implement the inventory and make it useful in short- and long-term organizational development situations.

Figure 3. System Performance Drivers for *ARC*

System Performance Drivers in Internal Process. *The organization is characterized by:*

CS1. *A legal status and statutory base that meets its needs.*
CS2. *A mission statement that is current, applied, and helps to guide action*
CS3. *Responsiveness to beneficiary needs.*
CS4. *Accountability to the community via the governing Board*
CS5. *Involvement of beneficiaries and other stakeholders in program decisions*
CS6. *Dissemination and communication of results and its needs*
CS7. *Building of coalitions, partnerships, and networks*
CS8. *Excellent relationships with principal government departments*
CS9. *Excellent relationships with principal sources of funds*
CS10. *A public image of integrity, cooperation, and capability*

System Performance Drivers in Resource Mobilization. *The organization is characterized by ...*

RA1. *Diversified financial resources, and resource development programs*
RA2. *Human resources, both volunteer and paid*
RA3. *Leadership resources who provide vision and strategic direction*
RA4. *Human resource diversity consistent with local setting*
RA5. *Adequate physical resources and infrastructure in HQ, branches, sites*
RA6. *Intellectual capital: knowledge, judgment, unique skills and/or individuals*
RA7. *Purchasing and procurement skills to make best use of financial resources*
RA8. *Technology and infrastructure resources: IT, logistics, telecom, MIS*
RA9. *Ability to exploit interactions between resources*

System Performance Drivers in Internal Process. *The organization ...*

IP1. *Facilitates decision-making, team development, and conflict resolution*
IP2. *Has effective internal communications*
IP3. *Has logistical capability appropriate to serve the vulnerable.*
IP4. *Works to develop staff, volunteer, management, and leadership resources.*
IP5. *Has and uses processes for organizational learning and development.*
IP6. *Has control and reporting via budgeting, planning, reporting, and audit*
IP7. *The Board provides guidance and review*
IP8. *Management is accountable via strategic planning and evaluation of service*
IP9. *Is efficient, with growing productivity and with appropriate capacity.*
IP10. *Uses its resources in an ethical and just manner.*

System Performance Drivers in Goal Attainment. *The organization ...*

GA1. *Uses long-term, strategic planning to achieve the mission*
GA2. *Achieves its long-term, strategic goals*
GA3. *Has a program development strategy to respond to vulnerability*
GA4. *Achieves its program goals*
GA5. *Has business planning to match services to strategy*
GA6. *Achieves its business goals*
GA7. *Makes operational plans*
GA8. *Achieves its operational goals*
GA9. *Evaluates outcomes against goals at all levels*
GA10. *Plans to improve performance.*

System Performance Drivers in Adaptation and Responsiveness *The organization is responsive ...*

AR1. *To constituency change (vulnerable, political / social environments)*
AR2. *To resource changes, e.g. economic change*
AR3. *To changes in managerial practice*

and adaptive ..:

AR4. *Adapts goals to changing circumstances*
AR5. *Changes actions in response to evaluations*
AR6. *Learns, implementing new knowledge and innovation from these scans*

Figure 4 SPD list used in Society

Constituency Satisfaction: ***How do these statements serve to describe the Association?***

❏   *A current and relevant mission statement.*
❏   *Responsiveness to constituent needs.*
❏   *Accountability to constituents*
❏   *Beneficiaries and other stakeholders are involved in program decisions*
❏   *Is able to disseminate and communicate its results and its needs*
❏   *Shares and/or leads in coalitions, partnerships, and networks*
❏   *Has excellent relationships with principal sources of funds*
❏   *Has a public image of integrity, cooperation, and capability*

Resource Acquisition: ***How do these statements serve to describe the Association?***

❏   *Sufficient human resources (volunteer and staff) to deliver service, attract resources, and promote mission*
❏   *Leadership resources to provide vision and strategic direction*
❏   *Knowledge resources such as special information, skills, tools, and judgement*
❏   *Diversity in human resources consistent with the community*
❏   *Programs to develop financial resources*
❏   *Physical resources and infrastructure to meet service needs*

Internal Process: ***How do these statements serve to describe the Association***

❏   *Appropriate decision-making, team development, and conflict resolution*
❏   *Effective internal communications*
❏   *Has and uses logistical capability appropriate to serve beneficiaries.*
❏   *Ways to develop staff, volunteer, management, and leadership resources.*
❏   *Learns and develops as an organization.*
❏   *Has control and reporting via budgets, plans, reports, and audit*
❏   *The Board provides guidance, help, and review to support the mission*
❏   *The staff and Board work together to accomplish the mission.*
❏   *It is efficient, with growing productivity and with appropriate capacity.*
❏   *Uses its resources in an ethical and just manner*

Goal Attainment: ***How do these statements serve to describe the Association?***

❏   *Uses long-term, strategic planning to achieve the mission*
❏   *Achieves its long-term, strategic goals*
❏   *Develops programs to respond to beneficiary needs*
❏   *Achieves its program goals*
❏   *Makes administrative plans (operations, structure, budget, resource use)*
❏   *Achieves its administrative goals*
❏   *Has operational planning to deliver services according to program goals*
❏   *Achieves its operational goals*
❏   *Evaluates outcomes against goals at strategic, program, administrative, operational levels*

Environmental Adaptation and Responsiveness: ***How do these statements serve to describe the Association?***

❏   *Responds to constituency change (e.g., vulnerability, political / social change)*
❏   *Responds to resource changes, e.g., economic change*
❏   *Learns as an organization*
❏   *Adapts goals to changing circumstances*
❏   *Responds to lessons from evaluations*
❏   *Scans the environment for innovation and uses it to help*

Faith is a multi-divisional organization with service programs in a variety of human service areas. Throughout this structure, the SPD list met considerable resistance. When initially presented, it had not been sufficiently adjusted from the *ARC* setting, and some terms did not fit in the situations where *Faith*'s managers operated. When used in individual unit planning sessions, it aroused antagonism because it was seen as too general, and (in effect) not respectful of *Faith's* uniqueness. In the long run, the leaders *Faith* agreed that a performance measurement instrument customized to their needs was necessary, and ultimately incorporated a revised SPD list into their ongoing planning procedures.

The experience with *Faith* pointed out the importance of customizing the SPDs to an organization's setting. They also sought efficiency, to focus on problem areas rather than to celebrate successes, and to conclude the measurement process with a specific list of problem areas that needed solutions. In the consulting activity, it was recommended that the SPD list be culled to remove drivers that were clearly not informative. That is, when reviewing their performance using the SPD list, organizations ought to start by eliminating particular drivers that describe performance areas where things were unambiguously going well. The next stages of the assessment were to identify a set of five to seven high priority SPDs and, and then a similar set of second priority drivers. This would result in a set of performance issues, targeted by the organization's managers, drawn from a "menu" of possible performance areas. One strong recommendation was that cumulatively, at least two of the SPDs chosen from the set of 10 to 14 ought to come from each of the five component areas. This would preserve the systemic view of assessment, and not emphasize too few of the components.

Some managers wanted to measure more precisely. The SPDs as descriptive statements fit easily in short-run sessions, when assessment is only a step in the path. Leaders in more institutional settings wanted tools to use over time, to seek out and solve problems. They sought more detailed and precise instruments. The measurement framework recommended to *Faith* is shown in Figure 5. This is effectively a menu of measurement options that the organization could choose to implement, depending on their resources and level of sophistication. For each of the 10 or so SPDs that the organization might choose to measure,

they could use these tools to create a customized measurement scale that was specific to their needs. Rather than a single scale for each one, this approach lets users measure any given SPD in terms of reference that are most meaningful within their regular scripts. This gives ownership of the assessment process to managers, making them stakeholders and not just informants. It recognizes that organizational performance on individual drivers deviates from the optimum for different reasons. Some things are just plain wrong. Some things aren't right yet. Some things are better than they used to be.

Figure 5 Tools for Measuring System Performance Drivers

Standard	What is there to measure?	Indicator / Measure	Example
Determinate	Binary – yes or no	A clear rule	A successful audit or accreditation
Rating Scale	From "acceptable to unacceptable," or "excellent" to "poor"	Measure on a 3-, 5-, or 7-point scale	Public awareness of the organization's mission
Evaluation	Compare to an adopted goal	Percentage reached or distance to goal	Decline of specific vulnerability
Developmental	Step-by-step	Milestones reached	Complete stages of a strategic plan
Benchmark	Compare to a recognized leader	Proximity to a standard of excellence	Compare to a competitor organization
Historical	Compare to past experience	Previous performance	Compare to past status of the organization

Table 3, excerpted from a report to *Faith*, shows examples of how to use how different measurement standards from the previous figure.

Table 3. Sample Measurement Standards

Sample System Performance Driver	Standard Used	Performance in this area?
CS3. Responsiveness to beneficiary needs.	Determinate	Good in general, need to help population x
CS5. Involvement of beneficiaries and other stakeholders in program decisions	Rating Scale	Measured on a multi-item rating scale
RA4. Diversity in gender, age, location, and ethnicity consistent with local setting	Developmental	Division is behind schedule in reaching diversity targets
IP8. Management is accountable via strategic planning and evaluation of service	Historical	Not so good up until 2000, but look at us now!
IP9. Is efficient, with growing productivity and with appropriate capacity.	Benchmark	Less efficient than other service areas in *Faith*
GA9. Evaluates outcomes against goals at strategic, program, administrative, operational levels	Evaluation	Only half of the program activities are regularly evaluated. Goal setting is uneven.

Is it Working?

The ultimate internal validity of the SPDs will be demonstrated as organizations that use this approach and conduct subsequent evaluations using the same measurement indicators. So far, this has not occurred in enough organizations to present a report. Interim results are favorable, however. Feedback has already been received that indicates how the inventory is meeting its original objectives, and where additional work in validating is still needed. *ARC* supports organizational development efforts in several regions globally. Three National Societies are undergoing OD activities in which organizational assessment using the SPDs plays a central role, in central Europe, and southern Africa. *ARC* delegates in two and an *ARC*-contracted OD consultant in another have each implemented the assessment methodology, with minor modifications to the SPD inventory.

What is similar between these cases is how the SPD inventory has been both preserved and altered. In all cases, measurement inventories have preserved the component framework and many of the SPDs. However, the original inventory is generally slightly altered, giving more or less emphasis to specific areas of management concern. Thus, field users are finding that it can be adapted to circumstances in place.

Findings of the Action Research

The action research process has generated some problems, and considerable promise. Some themes have recurred in all applications. The detailed SPD measurement inventory requires resource commitment, but the more compact approach was effective in some settings. The results do not appear to vary with industry, if sufficient care is taken to modify the terms to fit particular situations. Because it is focused on identifying problems, the inventory helps to focus management attention on the performance path ahead, rather than being a retrospective. More headlights than rear-view mirror, one might say.

Over the six cases and other situations where the methodology could be used, some key development processes have been recurrent on the parts of managers and leaders trying to implement the systems approach and the SPD inventory. These stages have included:

- Alignment with strategic planning scripts of clients
- Identifying measurement frameworks
- Creating measurement tools
- Matching performance assessment to resources
- Integrating performance assessment into planning

Some positive possibilities appear likely from the inventory and the systems model. The systems framework gives a consistency not only across organizations, but is also a clarifying mindset for some managers. Thus, it contributes to a common language. It is both scalable (e.g., moving from five components to 40-plus SPDs), and upgradeable (can be modified, can be customized).

However, problems remain. Continually improving performance is heavy lifting; in the end, performance measures may regress towards the mean, seeking easier performance targets (Campbell & Stanley 1966). Or perhaps there is a phenomenon similar to Gresham's law, which states that bad money will drive out good. Perhaps bad performance components drive out good ones, in which case, the flexibility and adaptability is a weakness as organizations seek lowest common denominators and externally recognition of their performance standards.. Linguistics are a challenge, given the need to use adjectives to describe performance. Other problem areas are political, because ownership of performance measurements is power. In sum, however, users were able to meaningfully use the inventory to both conduct performance assessment and to use the results to plan change in their organizations.

Subsequent research will attempt to validate the inventory in more conventional (statistical) ways to supplement the experience-based validity described in this paper. The systems model ought to apply equally well in business as it does in nonprofit service areas, but the SPD list for different industries (manufacturing, service, retail, etc.) will perforce be different. There, of course, it must compete with Balanced Scorecard and similar technologies and models. Extending to more nonprofits and businesses will require resources, in order to apply and test the method in multiple environments. In the end, the most meaningful way in which the systems model and performance drivers will be seen as valid must follow from managers, who, in their own terminologies and frames of reference, adopt them as their own standards for their own organizations' effectiveness.

* * * * *

References

Campbell, Donald T. & Julian T. Stanley, (1966), <u>Experimental and Quasi-Experimental Designs for Research</u>, Boston: Houghton-Mifflin

Gioia, Dennis A. & Peter P. Poole, (1984) "Scripts in Organizational Behavior." Academy of Management Review. 9 (3) 449-459.

Herman, Robert D. & Richard D. Heimovics, (1993). Social Construction of Nonprofit Organization Effectiveness: An Interim Research Report Presented to Association for Research into Nonprofit Organizations & Voluntary Action, Toronto, Ont.

Fornell, Chael & Jaesung Cha, (1994) *Partial Least Squares* in Bagozzi, Richard, ed. <u>Advanced Methods of Marketing Research</u>. Cambridge: Blackwell

Kushner, Roland J.(1994) "What is it? vs. What it is?: An Open Systems Approach to Organizational Effectiveness." In Mainero, L. A. & Palmer, D. D. Proceedings of the Eastern Academy of Management. Albany, NY: Eastern Academy of Management.

Kushner, Roland J. (1999) "The Shape of Effectiveness: An Evolving Conceptual Model of Organizational Performance,"Presented to the Association for Research on Nonprofit Organizations and Voluntary Action, Arlington, VA

Kushner, Roland J. (2000) <u>Guide for Assessing Organizational Effectiveness in National Societies. A Resource for Organizational Development Facilitators in the International Red Cross / Red Crescent Movement</u> Washington, DC: American Red Cross International Services Department.

Kushner, Roland J. & Peter P. Poole, (1996) "Structure-Effectiveness Relationships in Nonprofit Arts Organizations" Nonprofit Management & Leadership 7(2).

Neumann, W. Lawrence (1991) <u>Social Research Methods</u> Boston: Allyn and Bacon

Poole, Peter P., (1989) "Influence Modes, Schema Change, and Organizational Transformation." Journal of Applied Behavioral Science. 25 (3), 271-289.

Poole, Peter P., Barbara Gray, & Dennis A.. Gioia, (1990) "Organizational Script Development Through Interactive Accommodation." Group and Organization Studies. 15(2), 212-232.

Stone, Melissa M., Barbara Bigelow, & William Crittenden, (1999) "Research on Strategic Management in Nonprofit Organizations. Syntehsis, Analysis, and Future Directions" Administration & Society 31(3) 378-423

Tassie, Bill, Vic Murray, James Cutt. & Denise Bragg, (1996) "Rationality and Politics: What Really Goes on When Funders Evaluate the Performance of Fundees?" Nonprofit and Voluntary Sector Quarterly 25(3), 347-363

Weick, K. E. (1988). Enacted sensemaking in crisis situations. Journal of Management Studies, 24(4).