

Performance Management and Evidence Based Practice: Similarities and Differences

Edward T. Jennings, Jr., Ph.D.
Martin School of Public Policy and Administration
University of Kentucky
415 Patterson Office Tower
Lexington, KY 40506-0027
859-257-5596
Pub714@uky.edu

Jeremy L. Hall, Ph.D.
Department of Government
University of Alabama at Birmingham
238 Ullman Bldg.,
Birmingham, AL 35294-3050
205-934-0509
jeremyhall@uab.edu

Paper prepared for presentation at “A Performing Public Sector: The Second Transatlantic Dialogue,” sponsored by the European Group on Public Administration and the American Society for Public Administration, Leuven, Belgium, June 1-3, 2006

Performance Management and Evidence Based Practice: Similarities and Differences

Evidence-based practice and compendia of best practices have proliferated in recent years at the same time that governments have implemented performance management systems. In this paper, we explore the theoretical and practical connections between these two mutually-reinforcing policy and management approaches to improved governance. We examine the underlying logic of each approach. We ask whether they share the same purposes, serve the same functions, and develop in the same manner. We compare their strengths and limitations. We conclude by examining how the approaches might be integrated to enhance the management of public programs for public purposes. We do not attempt to answer the question of what difference these approaches make or what affects their impacts on policy and performance (see, e.g., Moynihan, 2006; Coe and Brunet, 2006; Yang and Holzer, 2006).

Understanding the Concepts

Performance measurement and performance management are interrelated concepts. The terms are often used interchangeably, although measurement is a tool that is used in performance management. Related concepts in the public sphere are results based government, managing for results, report cards, and performance oriented government. Whatever the terminology, the general idea is to focus attention on the results of government activity, rather than process or outputs, and to manage public policy and public programs to enhance performance.

In *Managing for Results*, Kamensky and Abramson assert that “Managing for results is shorthand for a conceptual framework that reflects a fundamental change in the management culture of governments across the globe. It is a culture that is fact-based, results-oriented, open, and accountable (2001: 5). They note that the Canadian Auditor General has said:

“Managing for results allows managers to make changes once they know what is working and what is not. It represents a significant difference in the way government programs are managed. Managers can then pay more attention to finding out whether programs are meeting their objectives and less to only carrying out activities or setting up structure and processes (Office of Auditor General, 2000: 20).”

Broom, Caudle, Jennings, and Newcomer (2002) have pointed out that a diverse set of reform pressures in recent years have focused the attention of citizens, officials, and public administrators on performance, results, and accountability. In the United States, the Government Performance and Results Act of 1993, the National Performance Review of the Clinton Administration, and the Program Assessment Rating Tool of the Bush Administration have all focused attention on managing for results. State and local governments across the country

have developed a similar focus. This parallels international developments, sometimes captured under the New Public Management Label (Moynihan, 2006). Broom and her colleagues suggest that measurement alone is not enough. Instead, “the data must be used by program managers and decision makers to assess and improve results (Broom, *et al.*, 2002: 2).”

For almost a century now, there has been interest in our ability to measure the activities of government and to assess government efficiency. An early example is found in the Bureau of Municipal Research and its efforts to study municipal government (Williams, 2003). Herbert Simon’s mid-century dissertation was directed at measuring municipal efficiency. For at least the last thirty years, Harry Hatry and his colleagues at the Urban Institute have provided guidance in measuring local government efficiency and performance.

Despite this long history of attention to measurement, the contemporary movement for performance measurement and performance management emerged and matured over the last twenty or so years in response to citizen dissatisfaction with government performance, growing constraints on public spending, and a developing public management philosophy that shifted focus from hierarchy and process to competition and outcomes. We can see the evidence of that shift and chart its growth by looking at examples drawn from the international movement for a New Public Management, the development of performance measures in American state and local governments, the National Performance Review of the Clinton administration, and the focus of the Government Performance Project.

The outcome orientation can also be observed in the plethora of community indicators that have been promulgated in recent years. These are found in a large number of communities around the country. Some prominent examples are **Jacksonville Community Council, Inc.** (JCCI) and the Boston Indicators Project carried out by the Boston Foundation. JCCI (2003) is an internationally recognized community-based organization that has issued a quality of life report annually since 1985 for Jacksonville, Florida and its surrounding counties. The Boston Foundation’s Boston Indicators Project measures the community’s progress across ten different quality of life dimensions, making information available to community leaders, public officials, and citizens alike (Boston, 2006).

A wide range of organizations have developed report cards intended to chart accomplishments, highlight weaknesses, and foster improvements in governments, agencies, and programs (Gormley and Weimer, 1999). Exemplary of these indicators are the Morgan Quitno (2006) Most Livable State rankings, and the Corporation for Economic Development’s Development Report Card of the States (2006). Such indicators focus the competitive nature of states on improved quality of life and economic performance.

Paralleling the growing attention to performance measures has been increased interest in best practices, or more specifically, evidence-based practice or decision

making. Best practice generally refers to approaches to policy and/or management that have been found to be highly effective. This might be done, for example, by identifying high performing organizations and then asking what makes them effective. It might also be done by convening a panel of experts and asking them to identify best practices in a particular field of endeavor. A recent symposium in the *Journal of Policy Analysis and Management* highlights four such compendia and their uses, but also considers some of the differences among them (Bardach 2003).

Evidence-based practice, which is a form of best practice, refers to practice based on systematic, objective evidence from empirical research. It has scientific findings as its basis. As Maynard (2006: 249) points out, “Over the past 10 years, there has been growing emphasis on evidence-based policy and practice in the United States and elsewhere around the world.” Its use is widespread in medical fields, and there is growing demand for it in other, increasingly diverse, arenas. Two years ago a new journal, *Evidence and Policy*, came into being to provide “comprehensive and critical treatment of the relationship between research evidence and the concerns of policy makers and practitioners.”

Examples of the growing attention to evidence-based practice abound. For example, the No Child Left Behind Act in the U.S. requires that Title I school-wide and targeted assistance programs use effective methods and instructional strategies that are grounded in scientifically based research. The Council for Excellence in Government has a Coalition for Evidence-Based Policy that has worked on criminal justice. With the Justice Department’s Office of Justice Programs it began a collaborative initiative to advance evidence-based crime and substance-abuse policy. The Council has also worked on education and social programs and has a page listing a series of randomized control trials (RCTs) that it thinks should provide a basis for decision making: <http://www.evidencebasedprograms.org/>. A 2006 Kentucky law provides for a Forestry Best Management Practices Board to be established for the purposes of updating guidelines for water quality management, regulations regarding timber harvesting, and supervising the implementation of recommended forestry best management practices (Kentucky Legislative Research Commission, 2006).

The roots of evidence-based practice may be less clear, but the social experiments of the 1970s (e.g., the negative income tax experiments) and the growth of evaluation research provided an evidence base for decision making. It ought to be clear that this was a particular type of evidence grounded in scientific approaches to analysis. It is that scientific approach that we take to be the distinguishing characteristic and hallmark of evidence-based practice. At its heart are the notion of causal relationships and the capacity of the analyst to separate out alternative explanations for an outcome in order to establish the true cause. The ideal evidence for EBP comes from randomized control trials in which subjects are randomly assigned to treatment and non-treatment status. As Maynard (2006) points out, however, the evidence will often be mixed and much of its use depends on synthesizing a diverse set of studies.

Because developments in evidence-based practice have paralleled the growth of performance-based management of the new public management, it is useful to ask how these two seemingly independent concepts might naturally complement each other. Such complement may result from common underlying logic, similar purposes, functions, and through mutually reinforcing simultaneous development. Before exploring these connections, we first consider a number of examples in practice.

Using the Concepts in Practice

Florida was one of the early leaders in performance measurement in the American states. Its Office of Program Policy Analysis and Government Accountability (OPPAGA) led the way in reporting a wide variety of performance measures for the efforts of government agencies. Available on its web site is a 100 page report containing hundreds of performance measures for the full range of public agencies and programs (Performance Measures and Standards Approved by the Legislature for Fiscal Year 2005-2006) (accessed March 25, 2006 at http://www.oppaga.state.fl.us/reports/pdf/2005-06_Measures.pdf). For Vocational Rehabilitation alone, there are 17 measures, ranging from indicators of workload and activity (e.g., number of customers reviewed for eligibility) to outputs (percent of eligible workers receiving reemployment services) to outcomes (percent of VR customers retained in employment after one year) to unit costs (average cost of case life). The state has measures for all dimensions of performance. It makes the information available to the public through written reports and through OPPAGA's web site. These measures are used in the state legislature's performance-based budgeting process which relies on measures and standards established in agency long range plans, thereby encouraging performance improvements. The emphasis throughout is to improve government performance by developing measures that can be used to hold agencies accountable and guide decision making.

Oregon has had one of the most extensively developed performance measurement systems. The Oregon Benchmarks program is often cited as an exemplar of a state establishing a set of important goals and tracking outcomes over time. Rather than linking the goals to particular programs, the state established a set of broad goals and identified a set of indicators for each area. At the present time, there are seven general areas, 26 sub-areas, and 90 measures. The Oregon Progress Board (2003) issues a report detailing progress on the benchmarks. While the benchmarks are not specific to particular agencies, and there are few that would be within the sole control of specific agencies, all agencies have to tie budget requests to efforts to improve the state's standing on the indicators. The agencies also tie their strategic plans and performance measures to the benchmarks, as well as to statutory, legislative, and executive requirements (Oregon Progress Board, 2004). Thus, hundreds of program and agency level performance measures link to an overarching set of state social and economic objectives.

At the federal level, The Government Performance and Results Act of 1993 requires all federal agencies to develop a strategic plan, measure the accomplishment of results, report the results to Congress, and manage programs to enhance results. GPRA has a clear emphasis on measuring program results, rather than activities and outputs,

although Frederickson and Frederickson (forthcoming) have pointed out the devilishly hard time some agencies have demonstrating results. This is particularly true of agencies that rely on third parties, especially when those third parties are autonomous state governments. In pursuit of GPRA requirements, federal agencies have developed hundreds of performance measures. The idea is to set targets that can be pursued by federal agencies as they fulfill their legislative mandates and planned missions. According to the U.S. Government Accountability Office, President Bush's Program Assessment Rating Tool (PART) has increased agency attention to results reported under GPRA (U.S. GAO, 2005a, 2005b).

As with performance measurement, several exemplary programs reflect current use of evidence-based policy and practice. We begin a review of several examples by looking at the Coalition for Evidence-Based Policy (2006). The Coalition, sponsored by the Council for Excellence in Government, is an example of an initiative to promote evidence-based practice in social programs in the United States. Its eighteen member advisory board consists of prominent social scientists that have played key roles in the design and analysis of major social program experiments. Its avowed purpose is to foster the use of social programs whose worth has been demonstrated through randomized control trials (RCTs). The initiative draws attention to a set of studies that meet particularly rigorous standards, including studies of both successful and unsuccessful programs.

As the Coalition puts it:

The central problem that the Coalition for Evidence-Based Policy seeks to address is that U.S. social programs are often implemented with little regard to rigorous evidence, costing billions of dollars yet failing to address critical needs of our society -- in areas such as education, crime and substance abuse, and poverty reduction. A key piece of the solution, we believe, is to provide policymakers and practitioners with clear, actionable information on what works, as demonstrated in scientifically-valid studies, that they can use to improve the lives of the people they serve (2006).

To support that solution, the Coalition's web site summarizes the findings of experimentally tested programs that have strong policy implications. The focus is on programs that have been demonstrated to have positive effects and widely-used programs that have been found to have little or no effect. The evidence has to come from randomized control trials in which subjects are randomly assigned to treatment and control groups. The members of the Coalition require such studies to meet the gold standard in social research, providing evidence that is superior to that obtained from other studies.

To be selected for inclusion on the Coalition's web site, a program and its associated studies have to meet rigorous standards. They have to be well designed and implemented and must have significant policy implications. They take their criteria from

an Office of Management and Budget (OMB) document, [What Constitutes Strong Evidence of a Program's Effectiveness](#), which they point out includes such items as:

- Adequate sample size;
- Few or no systematic differences between the intervention and control groups prior to the intervention;
- Low attrition, and little or no difference in attrition between the intervention and control groups;
- Few or no cross-overs between the intervention and control groups after randomization;
- Placebo controls, where appropriate;
- Intention-to-treat analysis of study outcomes;
- Valid outcome measures, preferably well-established tests and/or objective, “real-world“ measures (e.g., arrest rates for a crime intervention);
- Blinded evaluators, where appropriate;
- Preferably long-term follow-up;
- Appropriate tests for statistical significance (in group-randomized trials, “hierarchical” tests that are based both on the number of groups and the number of individuals in each group) (Coalition, 2006)

Applying these standards, the Coalition has to date identified 24 programs for inclusion on its list, including 19 successful programs and 5 unsuccessful ones. The programs are in diverse areas of social policy: early childhood development, education, substance abuse, mental health, crime, employment and welfare, retirement, and international development.

The Drug Effectiveness Review Project (discussed below) is but one of many efforts to use science and evidence to guide health policy and program practices. The practice of medicine is shaped extensively by evidence-based practice. Timmermans and Mauck (2005: 18) note that evidence-based medicine is “commonly defined as ‘the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients.’” They explain that the term is used many different ways, with meanings ranging from meta-analysis to randomized clinical trials; a common implementation of evidence-based medicine is the use of clinical practice guidelines.

Tanenbaum (2005: 163) states that evidence-based practice refers in a general way to the application of scientific research findings to the treatment of individual patients; she adds “evidence-based medicine is one field of EBP; evidence-based mental health care is another.” The purpose of evidence-based medicine and practice guidelines is “to provide a stronger scientific foundation for clinical work, to achieve consistency, efficiency, effectiveness, quality, and safety in medical care” (Timmermans & Mauck, 2005: 19).

Many states and the federal government are already engaged in activities to produce or utilize evidence-based practice in health policy, moving the use of such evidence well beyond the individual practitioner’s application in clinical practice. At the

federal level, Medicare relies on evidentiary support to make national coverage decisions that dictate which treatments are allowed or disallowed for particular ailments, and under what conditions. Neumann, et al, (2005) evaluated sixty-nine technologies reviewed by Medicare between 1999 and 2003 to determine if there is a relationship between quality of evidence and the decision to approve coverage. They found that “[t]echnologies with good evidence were more likely than those with fair or poor evidence to be covered” (Neumann, et al, 2005: 246-247), and they conclude that “decisions have generally been consistent with the strength of supporting evidence” (2005: 250).

Another initiative that is promoting the use of evidence in policy making is the Drug Effectiveness Review Project, a collaboration of mostly public organizations that gathers the best available evidence on the comparative safety and effectiveness of drugs in the same class (Oregon Evidence-Based Practice Center, 2006). Fifteen states, a foundation, and a Canadian agency are sponsors of the project: Alaska, Arkansas, California, California Health Care Foundation, Canadian Agency for Drugs and Technologies in Health, Idaho, Kansas, Michigan, Minnesota, Missouri, Montana, North Carolina, New York, Oregon, Washington, Wisconsin, Wyoming. The project has already completed 12 reports that were commissioned by three states that were the original partners in the initiative. Another 14 studies are currently underway with expected release by September, 2006.

Each drug class review is based on a careful assessment of evidence from multiple studies. In selecting studies for inclusion in the review, the reviewers are attentive to the scientific quality of the study, with an emphasis on internal validity. This includes study characteristics such as “methods used to enroll patients, assign people to treatment groups, assess outcomes, and analyze data...”

The Drug Effectiveness Review Project carefully states that:

The purpose of the DERP reports is to make available information regarding the comparative effectiveness and safety profiles of different drugs within pharmaceutical classes.

Reports are not usage guidelines, nor should they be read as an endorsement of, or recommendation for, any particular drug, use or approach. Oregon Health & Science University does not recommend or endorse any guideline or recommendation developed by users of these reports (Oregon, 2006).

In other words, the Oregon Health Science University, which provides the home for the project, and the participating sponsors are not endorsing particular uses of the guidelines that states might develop. Despite this, the project collaborators clearly intend for the reviews to guide practice in state operated health programs, like Medicaid and the Child Health Insurance Program. The Commonwealth Fund (2006) points out that the states participating in the collaborative have joined together to reduce costs and maximize benefits in their

pharmaceutical benefits programs. Each partner has its own local process for deciding how to use the information provided by the reviews.

Each participating state has a committee or commission which reviews the reports and decides how to use them. The committees generally are composed of a mix of medical practitioners and pharmacists, although a few states (New York, Oregon, Wisconsin, Wyoming) include consumers on the committee. Several states provide typical examples of their use of the reviews. In Missouri, the Drug Prior Authorization Committee and the Drug Use Review Board use the evidence-based review reports along with other information in selecting medications for a preferred drug list for the Medicaid program. The Arkansas committee uses the reports, along with other information, to develop a preferred drug list. Medicaid will pay for drugs on that list without requiring prior authorization. In Wisconsin, the Medicaid Prior Authorization Review Committee uses the reports in its process for identifying preferred drugs that can be prescribed without prior authorization. It is not yet clear how much impact these reviews have on decisions, but the Commonwealth Fund (2006) has a study underway to assess the impact of this project. That study asks:

1. What strategies are states using to apply evidence-based medicine to manage pharmacy costs and promote appropriate access to prescription drugs, and what has been their experience in using these strategies?
2. How much shift occurs in the utilization of various drugs within a particular therapeutic class when strategies, such as preferred drug lists, are applied to that class of drugs?
3. How do various factors, including the source of evidence, degree of stakeholder input, and choice and strength of enforcement mechanisms, make a difference in the development, operation, effectiveness, or acceptance of the strategy? (Commonwealth Fund, 2006)

At the state level, evidence-based practice has become broadly influential in select policy fields. Hawaii has created the “Empirical Basis to Services Task Force, which searched and evaluated controlled studies in childhood mental health; this task force continues to monitor the literature to establish practice guidelines (Tanenbaum, 2005: 165). The guidelines establish what services will or will not be provided to clients in the Child and Adolescent Mental Health Division of the Hawaii Department of Health by defining the scope of work for contracts to provide services on behalf of the state.

The District of Columbia Department of Mental Health has proposed a policy regarding evidence-based psychotherapy. Under that policy, “all psychotherapy services delivered to community-based adult consumers will conform to a short list of EBPs;” (Tanenbaum, 2005: 168). Oregon has instituted—through the legislature—a requirement that a number of state agencies spend 25% of their program budgets on evidenced-based programs beginning in 2005, 50% in 2007, and 75% in 2009, with budgetary consequences for non-compliance (Tanenbaum, 2005: 170). Wyoming made systematic reviews from the Cochrane Library available electronically to all state citizens free of

charge in 2004, and also established a Prescription Drug Resource Center to provide online access to systematic reviews of drugs within classes (Fox, 2005). In Washington, legislation requires that the best available scientific and medical evidence, derived from systematic research, should guide coverage decisions for every agency of state government that purchases health care (Fox, 2005: 116-117). Though use in health policy areas is on the rise, evidence-based practice is probably not as widely embraced across other policy areas. However, the rationality of the process makes evidence-based practice appealing to policymakers—a fact that is likely to fuel more universal efforts to use evidence-based practices.

Theoretical and Practical Connections

While EBP and PM developed simultaneously, with common goals, they are unique concepts. We evaluate, in turn, various theoretical and practical commonalities that further connect these concepts in public management. Among the dimensions we consider are outcome orientation, reliance on causal theory, role of public engagement, relative pervasiveness in the system of modern governance, the role of policy goals in their utility, and their reliance on the scientific method for direction.

Reasonable consideration suggests that both performance management and evidence-based practice are part of the long effort to rationalize governance and maximize general outcomes, which Stone (2002) has described as the rationality project. Whether they actually attempt to maximize benefits is less critical than the fact that both call on methods of science to produce better outcomes of public spending by focusing on results measurement. Both offer the potential and have the goal of making policy and programs more effective and efficient.

Both PM and EBP are outcome focused. They ask managers, policy makers, and citizens to attend to the results of public policy rather than the processes of public agencies. Both necessarily have process components, and the evolution of PM has broadened its perspective to include process considerations through balanced scorecard approaches. Whereas traditional approaches to public education might have focused on inputs (e.g., dollars, teachers) and processes (e.g., phonetics, new math), contemporary practice focuses on end results reflected through a variety of testing procedures. Ultimately, both seek to improve the operations and results of public programs, so both are results oriented.

Both measure outcomes, but EBP asks whether the outcomes are a product of the policy or program, whereas PM tends to assume that outcomes are a product of the policy or activity. This is a critically important distinction. There is little in the promulgation or practice of performance management to suggest that it is attentive to systematic evidence about the causes of the outcomes of public programs. Indeed, PM demands that agencies manage resources so as to improve the attainment of results, but nothing in GPRA or state initiatives calls on agencies to provide evidence that the policy and the agency's efforts are responsible for producing the desired outcomes. So long as the outcome is measured, the agency can take credit or absorb the blame. In cases such as Oregon's

benchmarking system, the agency may select outcomes that are not directly linked to the programs and policies in question. In other words, there is potential, through a cognitive disconnect, to assume and even convince legislators that a program is effective because of unrelated improvements to outcome measures of interest.

Burt Barnow (2000) has suggested that we need to move PM toward a concern with whether or not it is the program that makes the difference when performance measures improve. In the way of an example, a layperson's assumption is that decreased crime rates are the result of improved police work and law enforcement, while in fact this outcome may result more directly from a growing economy with full employment. In practice, causal linkages may not exist, or may not be well-understood. Managers are likely to rely on qualitative assessment or process monitoring to guide decisions with respect to the management of the program. The goal is to move beyond guesswork, theory, and selective observation to develop a more systematic, evidence grounded basis for policy and action. One is led to the question, is it better to de-link programs from goals and outcomes in the interest of creativity and innovation, or to require linkage that may artificially limit our frame of inquiry?

Evidence-based practice, of course, is quite different. Its entire focus is scientific evidence that a particular program or policy produces the desired result. It explicitly seeks to provide evidence that rules out alternative explanations for the result. It is grounded in scientific studies that produce internally valid results. Its holy grail is the randomized control experiment. When that is not available, it insists on the best quasi-experimental research. It typically relies on the accumulation of research results through a variety of studies with the weight attributed to different studies determined by evidence of validity and reliability.

Both PM and EBP have to account for the reality of the political process and the likelihood of goal multiplicity; each can fail to do this. That is, any given program may have a core identifiable goal, but it may also represent a plurality of varied goals that were linked to the policy during the legislative process. Take, for instance, investment in higher education; one goal, obviously, is to educate students. However, in the minds of individual legislators, the goal might be expressed as research, or innovation, or outranking benchmark states, or the employment that a university would provide, and so on. When goals are diverse, measuring program effectiveness is difficult. PM systems can end up incorporating multiple goals without much guidance on how to balance them or set priorities. Or, they may focus on one set of goals to the exclusion of others. Evidence-based findings with respect to a particular goal may provide little guidance on how to use the evidence in a setting of multiple, sometimes competing, goals.

With both EBP and PM, questions arise with respect to who sets the goals, who selects the targets, and who chooses priorities. There is considerable discussion in the PM literature with respect to whether performance measurement is more effective when it is developed internally or imposed from the outside. Typical texts discuss performance measurement and management as internal processes (e.g., Hatry, 1999), although they often recognize the need to engage multiple stakeholders in strategic planning, goal

setting, and selection of targets. Many analysts, however, suggest that performance measurement has its greatest impact when it comes from outside of the government or organization in the form of a report card (Gormley and Weimer, 1999; Coe and Brunet, 2006). The same would be true of evidence-based practice. Evidence often comes from research conducted independently of the organizations that might use it; it can indeed be driven by researcher determination of the appropriate goals or outcomes. By the same token, much evidence-based practice is driven by decision makers searching for the best information on which to make their decisions. The Drug Effectiveness Review Process, for example, is very much guided by decision makers.

Moreover, as we mentioned earlier, there is the issue of moving targets. As interest groups and advocacy coalitions exert influence on administrators within the policy subsystem, capture may occur, or agency goals may shift. Such was the case with tobacco policy in the United States as health interests invaded historically close-knit agriculture subsystems. Thus, while the goals of the legislature once mattered a great deal, the interpretation of those goals by skilled professionals may be different, and may in fact shift over time.

Once a program is approved by law, it has grown increasingly common for implementation to be carried out by professional administrators who are afforded broad bureaucratic discretion in exercising their duties (Meier, 2000). Professionalization of the bureaucracy has led to an increasing level of technical knowledge and expertise, and administrators who are more likely to understand the mechanisms and causal relationships at work in their field. As efficiency goes, administrators may better understand the policy and its goals than did the legislature that approved it. O’Leary (1994) describes such a case where a group of administrators that came to be known as ‘The Nevada Five’ initiated grassroots policy change in spite of contrary rules and direction provided by the law and their superiors.

In general, both performance management and evidence-based practice assume fixed targets for policy and action. It is the presence of a specific target or outcome that allows scientific research to take place. If the goals of a policy, program, or practice were to change from day to day, month to month, year to year, they would not be of much use as guides for agency management; nor would they be susceptible to scientific research. Both enterprises assume that goals can be specified with sufficient clarity to be measured and sufficient to provide a basis for action.

There is an additional problem with evidence-based research: its usefulness is limited by the resources available to decisionmakers. Randomized control experiments (or their nearest equivalent) are complex, difficult to carry out (particularly in field settings), and therefore very costly. In many policy areas, it is simply not feasible to expect experimental evaluation to be possible, in part due to the nature of the policy, and in part due to the impossibility of creating a “sterile” environment in the field. Furthermore, the time associated with collecting evidence may be beyond the decision timeframe. Faced with the challenge to “do something,” public managers are likely to satisfice rather than wait for compelling experimental evidence about their program.

Neither performance management nor evidence-based practice is particularly susceptible to public engagement, although some types of PM are community-based initiatives and there have been some efforts to foster citizen participation in PM. For example, the PEW charitable foundation sponsored a set of initiatives in the 1990s to identify ways to integrate PM with citizen engagement (Epstein, 2002). Yang and Holzer maintain that the success of PM will ultimately turn on citizen trust which can be enhanced by measuring what citizens care about and building participatory processes (2006). In addition, the community indicators movement, while not focused on management of public programs, provides a variety of models of citizen engagement in the development of community report cards composed of social indicators. EBP, with its close attention to science, presumes expert guidance of decision making. Hence, Maynard's plaint in discussing policy making, why don't they listen to the evidence? Surprisingly, her answer turns on the quality of the evidence rather than addressing features of policy making that might diminish attention to scientific evidence (Maynard, 2006). Likewise, much performance management is ill-tuned to the public, carried out within the agency as part of strategic planning and management. The Government Performance and Results Act does not call for public involvement in the development of agency plans and measures. Florida's performance-based budgeting is carried out within the normal budget process, one dominated by agencies and legislative committees, not public engagement. The Oregon Progress Board, which develops the Oregon Benchmarks, includes civic leaders and public officials among its members and has sought to engage citizens in a recent initiative, but the process by which state agencies develop performance measures and use them to guide performance is largely internal to the agencies.

This history, while not indicative of widespread citizen participation and engagement, does not preclude such engagement. The use of strategic planning, at its core, presumes that there is flexibility in determining goals within a framework of mandates handed down from political and administrative powers-that-be. Whether citizens are involved in this process or not depends first on the nature of the program and whether relevant stakeholder groups outside the agency exist. Where there is room for participants of the planning process to determine the mission, goals, and strategies of an organization, planning is more likely to be engaging. Communities and their direction-seeking leaders are more likely to benefit from the use of strategic planning efforts to set broad goals than are agencies with already narrow focus. Strategic planning's focus on creating public value through increased responsiveness, performance and accountability is a key part of the PM goal. The emphasis on selecting preferred strategies and on evaluation and reassessment of strategies closely resembles the goals of EBP.

The critical difference between PM and EBP is that evidence-based practice relies on scientific method to analyze results; performance management does not. Performance managers measure outcomes; they assess their attainment; they adapt policy, resources, and organization to enhance performance. In general, they do not rely on scientific studies to make the adaptations. Evidence-based practitioners, on the other hand, are entirely dependent on scientific research. If it has not been studied with attention to

sampling, measurement, internal and external validity and a host of other scientific concerns, there is no basis for action. A subtext of the Coalition for Evidence-Based Policy initiative is that we are better off not spending large sums on social programs until we have solid scientific evidence of efficacy.

This leads to a final major difference between PM and EBP. Performance measurement is ubiquitous. Measures of performance can be developed for virtually all policies and programs. For most policies and programs, those measures can be quantitative. States like Florida and Virginia have thousands of measures for the performance of their agencies. When Congress adopted GPRA, it required all federal agencies to develop strategic plans and measurable goals. In the end, it turned out, of course, that some goals are inherently difficult to measure with any significant degree of validity and reliability and it can be difficult to hold agencies accountable for many outcomes (Frederickson and Frederickson, forthcoming). Evidence-based practice is inherently more limited. It takes a heavy investment to develop valid, reliable evidence with respect to the outcomes of a practice or program. The range of agency goals for which scientific evidence can be produced is quite limited. And, as Maynard (2006) points out, the quality of the evidence and its applicability vary considerably.

Integration or Separation?

We have examined the commonalities and differences between EBP and PM, highlighting several strengths and shortcomings of each approach. We conclude that their joint use provides potential for improved policy and management, though not in strict terms. How might further integration of the concepts lead to improved outcomes for governments and agencies that use them? One method that stands out is to incorporate the desire for better causal understanding into PM efforts. Examining policy and program outcomes in the presence of theoretical expectations should increase organizational learning, and lead to greater efficiency and outcome improvements than examining outcomes as the assumed result of the programs and policies in place. As we previously alluded, such efforts will be restrained by the extent to which outcome measurement is feasible. Likewise, EBP may be more efficacious when applied with less rigid expectations for scientific validity where risk of harm is low, where experimental design is difficult, and with recognition that policies and programs (or interventions) may simultaneously serve multiple goals.

References

- Bardach, Eugene. 2003. "Creating Compendia of 'Best Practice.'" *Journal of Policy Analysis & Management*. 22(4): 661-665.
- Barnow, Burt S. 2000. "Exploring the Relationship Between Performance Management and Program Impact: A Case Study of the Job Training Partnership Act." *Journal of Policy Analysis and Management*, 19:1 (Winter), 118-41.

- Boston Foundation, 2000. Bostony Indicators Project, accessed April 28, 2006 at <http://www.tbf.org/indicatorsProject/>.
- Broom, Cheryle, Sharon Caudle, Edward T. Jennings, Jr., and Kathryn Newcomer, 2002. "Meeting the Challenges of Performance-Oriented Government," Chapter 1 in Kathryn Newcomer, Edward T. Jennings, Jr., Cheryle Broom, and Allen Lomax, 2002. *Meeting the Challenges of Performance-Oriented Government*. Washington DC: American Society for Public Administration.
- Coalition for Evidence-Based Policy, 2006. "What Works and What Doesn't Work in Social Policy?: Findings From Well-Designed Randomized Controlled Trials" accessed April 21, 2006 at <http://www.evidencebasedprograms.org/>.
- Coe, Charles K. and James R. Brunet, 2006. "Organizational Report Cards: Significant Impact or Much Ado About Nothing," *Public Administration Review*, 66, 1 (January/February): 90-100.
- Commonwealth Fund, "Using Evidence-Based Medicine to Design State Drug Programs," accessed at http://www.cmwf.org/tools/tools_show.htm?doc_id=232254 on April 22, 2006.
- Corporation for Economic Development, 2006. Development Report Card of the States. <http://drc.cfed.org/>.
- Epstein, Paul, Lyle Wray, Martha Marshall, and Stuart Grifel, 2002. "Twenty-First Century Governance: Better Results by Linking Citizens, Government, and Performance Measurement," in Kathryn Newcomer, Edward T. Jennings, Jr., Cheryle Broom, and Allen Lomax, 2002. *Meeting the Challenges of Performance-Oriented Government*. Washington DC: American Society for Public Administration.
- Florida Office of Program Policy Analysis and Government Accountability (OPPAGA), 2006. "Performance Measures and Standards Approved by the Legislature for Fiscal Year 2005-2006" accessed March 25, 2006 at http://www.oppaga.state.fl.us/reports/pdf/2005-06_Measures.pdf.
- Fox, Daniel M. Evidence of Evidence-Based Health Policy: The Politics of Systematic Reviews in Coverage Decisions. *Health Affairs*, Jan/Feb 2005 (24:1) pp. 114-122.
- Frederickson & Frederickson, forthcoming. *Measuring Performance in the Hollow State*. Washington, DC: Georgetown University Press.
- Gormley, Willliam T., Jr. and David L. Weimer. 1999. *Organizational Report Cards*. Cambridge: Harvard University Press.

- Hatry, Harry P. 1999. *Performance Measurement: Getting Results*. Washington, D.C.: The Urban Institute Press.
- JCCI. 2003. *Quality of Life Progress Report*. Jacksonville Community Council, Inc. Accessed May 5, 2004 at <http://www.jcci.org/newerhome.htm>
- Kamensky & Abramson, 2001. *Managing for Results*. Lanham, MD: Rowman & Littlefield Publishers, Inc.
- Kentucky Legislative Research Commission. 2006. Senate Bill 147: An Act relating to Forestry” accessed April 25, 2006 at <http://www.lrc.ky.gov/record/06RS/SB147.htm>.
- Maynard, Rebecca A. 2006. “Presidential Address: Evidence-Based Decision Making: what Will It Take for Decision Makers to Decide,” *Journal of Policy Analysis and Management*, 25, 1 (Spring): 249-266.
- Meier, Kenneth J., 2000. *Politics and the Bureaucracy: Policymaking in the Fourth Branch of Government*, 4th Edition. Fort Worth: Harcourt College Publishers.
- Morgan Quitno Press, 2006. Most Livable State Awards. <http://www.morganquitno.com/awards.htm>.
- Moynihan, Donald P., 2006. “Managing for Results in State Government: Evaluating a Decade of Reform,” *Public Administration Review*, 66, 1 (January/February): 77-89.
- Neumann, Peter J., Nomita Divi, Molly T. Beinfeld, Bat-Sheva Levine, Patricia Seliger Keenan, Elkan F. Halpern & G. Scott Gazelle. 2005. “Medicare’s National Coverage Decisions, 1999-2003: Quality of Evidence and Review Times,” *Health Affairs*, 24, 1 (Jan/Feb 2005): 243-254.
- O’Leary, Rosemary, 1994. “The Bureaucratic Politics Paradox.” *Journal of Public Administration Research and Theory*, 4:43-67.
- [Office of Canadian Auditor General, 2000.](#)
- Oregon Evidence-based Practice Center, accessed at <http://www.ohsu.edu/epc/> on April 22, 2006.
- Oregon Progress Board, 2003. *Is Oregon Making Progress? The 2003 Benchmark Performance Report*. Accessed May 13, 2004 at http://egov.oregon.gov/DAS/OPB/studies_reports.shtml.
- Oregon Progress Board, 2004. Annual Performance Measure Reports. Accessed August 3, 2004 at http://egov.oregon.gov/DAS/OPB/apr_Jan04.shtml.

- Oregon Progress Board, 2005. Report by the Advisory Committee on Citizen-Friendly Reporting to the Oregon Progress Board on the Government Performance Reporting Demonstration Grant, December, 2005. Accessed April 21, 2006 at <http://egov.oregon.gov/DAS/OPB/docs/kpm/NCCI/NCCIfinal.doc>.
- Stone, Deborah, 2002. *Policy Paradox: The Art of Political Decision Making*. Revised Edition. New York: W.W. Norton and Company.
- Tanenbaum, Sandra J. Evidence-Based Practice as Mental Health Policy: Three Controversies and a Caveat. *Health Affairs*, Jan/Feb 2005 (24:1) pp. 163-173.
- Timmermans, Stefan & Aaron Mauck. The Promises and Pitfalls of Evidence-Based Medicine. *Health Affairs*, Jan/Feb 2005 (24:1) pp. 18-28.
- U.S. GAO, 2005a. "OMB's Part Reviews Increased Attention to Improving Evidence of Program Results."
- _____, 2005b. "PART Focuses Attention on Program Performance, but More Can Be Done to Engage Congress."
- Williams, D. W. (2003) Measuring government in the early twentieth century. *Public Administration Review*, 63, 6 (November/December): 643-659.
- Yang, Kaifeng and Marc Holzer, 2006. "The Performance-Trust Link: Implications for Performance Measurement," *Public Administration Review*, 66, 1 (January/February): 114-125.