De-PART: A Multiplicity of Perspectives on PART's Mixed Legacies

Matthew Dull
Assistant Professor
Center for Public Administration & Policy, Virginia Tech
mdull@vt.edu – 202-821-3807

Abstract

Falling in the succession of presidential management and budget reform initiatives that preceded it, the Bush administration's Program Assessment Rating Tool (PART) did not last long following the president's departure. Criticizing PART evaluations as "insular" and "ideological," incoming Obama administration OMB officials promised to "fundamentally reconfigure" OMB's approach to program assessment. At the same time, though the details of its "government-wide performance measurement framework" are not yet fully formed, the continuities in rhetoric and ideas between administrations are striking. Indeed, a long scholarly tradition suggests initiatives like PART leave behind a variety of organizational and institutional legacies. This paper explores a multiplicity of perspectives on six of PART's legacies: 1) framing decisions; 2) increasing transparency; 3) signaling commitment; 4) constructing dialogue; 5) accumulating knowledge; and 6) governing by abstraction. Each yields a mixed picture. Reflecting on this multiplicity of perspectives, I contemplate why the persistent rhythm of comprehensive administrative reform continues, in some basic respects, to confound the field's most imaginative and careful scholarship. Drawing on an extended elaboration of Arthur Stinchcombe's *When Formality Works* (2001), I conclude that scholarly efforts to explain and understand the role of performance information would be well-served by attending to what Stinchcombe terms governing abstractions.

Draft 1.0
Prepared for the 10th Public Management Research Association (PMRA)
Research Conference, Columbus, OH, October 3, 2009. As always, feedback is welcome.

Introduction

Summer 2009 saw a welcome departure for OMB's program examiners, most of whom spent part of the previous six summers puzzling over the yes's and no's of the Bush administration's Program Assessment Rating Tool (PART) – sometimes to the great consternation of agency personnel. PART may have become as unpopular among OMB's examiners as it was among administrators of the programs PART's 30-odd weighted binaries tagged "effective," "ineffective," or some finely calibrated shade of gray. Program examiners serve as the point of contact for a wide range of program-related issues. They are busy people; and many came to resent the days and weeks spent completing PART assessments.[1] Few regretted the Obama administration's decision to abandon it. None found the hours difficult to fill.

Within weeks of the transition, the Obama administration announced it planned to "fundamentally reconfigure" the PART process. Echoing PART's most acute critics (Radin 2006), OMB promised to "open up the insular performance measurement process to the public, the Congress and outside experts," to "eliminate ideological performance goals and replace them with goals Americans care about…" (OMB 2009). This was a bold rebuke, to be sure, and a reflection of the real disdain that had been building for years around PART and the Bush administration's broader President's Management Agenda (PMA). At the same time, there are few subjects on which President Obama sounds more like George W. Bush than when he promises to review the budget "line by line," rooting out wasteful and ineffective programs. Indeed, there is a real continuity in the two administrations' commitment to integrating performance information government decision-making. Describing plans for the new administration's, "government-wide performance measurement framework," OMB Deputy Director Jeff Zients promises, like PART, it "will be focused on outcomes, allow comparisons across programs and agencies, and show trends over time…" (Zients 2009).

A long scholarly tradition holds that initiatives like PART may come to an end, but they never fully fade away, leaving behind a variety of organizational and institutional legacies (Arnold 1998; Schick 1966). This paper reflects on PART's legacies, drawing on the increasingly robust and varied scholarly perspectives that have marked the "middle aging" of the performance movement.[2] Like the end of government-wide mandated planning-programming-budgeting (PPB) the schools and offices of planning and analysis it left behind embodied and enacted PPB's varied legacies (Schick 1973). Agencies, OMB, Congress, and the community of consultants and academics who watched it develop depart along a path defined by the PART experience. But, again consistent with prior experience, the lessons derived from PART yield no easy answers, no neat principles or proverbs to guide future action.

This paper offers a multiplicity of perspectives on six of PART's varied legacies: 1) framing decisions; 2) increasing transparency; 3) signaling commitment; 4) constructing dialogue; 5) accumulating knowledge; and 6) governing by abstraction. The list is neither comprehensive, nor very tidy in its organization and treatment. Each perspective yields a mixed picture. Together though I think they highlight the ongoing sometimes discordant rhythm of comprehensive administrative reform in

---

[1] Program examiners are OMB's "frontline" staff, positions designed by the Clinton-era OMB 2000 reorganization in order to integrate management and budget functions. It is worth noting that this is not OMB's first effort to merge the management and budget functions (Rudalevige 2003). Indeed Radin (129-130) sees the OMB 2000 reorganization as less an integration than a depletion of OMB's management function. Regardless, the program examiners are the inheritors of OMB's role as the president's "eyes and ears" in policy and relations with the agencies (Neustadt 2000, 51-2; Berman 1979, 67; Tomkin 1998, 81).

[2] With apologies to Hood and Peters (2004).

institutional politics. They also help to illustrate why, after more than a century of careful study, administrative reform continues to confound scholars and theorists. To lend the treatment at least a semblance of order, the next section begins by mapping a handful of common perspectives framing the way observers view PART and the prospects of performance management more broadly.

Mixed Results and Multiple Perspectives

In their influential review of 20th century reform James March and Johan Olsen (1983, 1989) organize reform-thinking into two "rhetorics." The first embraces a language of administration, emphasizing efficiency, competence, and managerial control. March and Olsen label this the rhetoric of orthodox administration. The second follows a realist logic, which they label realpolitik, emphasizing interests, power arrangements, and political control. There are other ways to frame this broad distinction – the first is prescriptive, the second political – but as March and Olsen compellingly illustrate both are easily discernable in thinking and talking about reform across eras. But in developing this dichotomy, March and Olsen suggest another set of dynamics, which rests at the core of their project – and, indeed, helps to define a whole era of sociology and institutional theory. They write:

> The orthodoxy of administration is the voice of the prologue to comprehensive administrative reform; the orthodoxy of realpolitik is the voice of the epilogue; the myths of the first shade into the myths of the second over the course of a major effort at reorganization; and both sets of myths are needed for a normatively proper interpretation of the reorganization saga (March and Olsen 1989, 92)

In the logic of orthodox administration, the language and the substance of reorganization are presumed to be closely linked by an instrumental logic: Reform is common because government needs fixing; reforms fail from flaws in design or execution. By contrast, the realist logic suggests reform is common because actors seek power. As a consequence, language and substance are often divorced; and reforms must be judged not by empty language, but by actions and outcomes.

Underlying March and Olsen's dueling rhetorics is the notion of "loose coupling," central to much sociological new-institutional theory (Meyer and Rowan 1977), which emphasizes the decoupling of formal structures from the substance of organizations. Like the game theorists' "cheap talk" the concept of decoupling is fundamental to the research tradition and, as a consequence, it is often misused.[3] March and Olsen's project is more ambitious and more interesting than simply an effort to demonstrate reform's empty promises; instead they use comprehensive reform to explore problematic attention and the elaboration of meaning through institutional politics. Nonetheless, the degree to which reform is linked to the "stuff" of organizations orients much contemporary thinking. This is particularly true of results-model reform and the performance measures that are its byproducts, which are criticized at once as meaningless and as a threat professional self-governance. Very often both contain a kernel of truth, but just as often they are conflated by critics of performance management in ways that obscure not only the pathologies of performance measurement, but its usefulness as well.

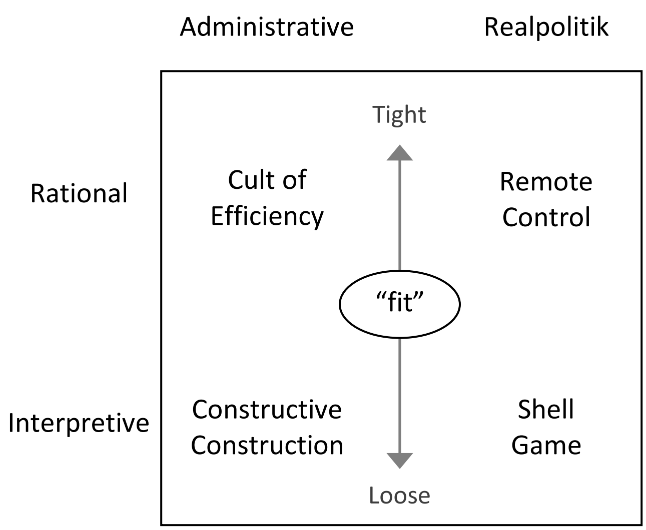
Figure 1 maps a handful of common perspectives on PART's legacies, organized around the two dimensions described above. The administrative and realist logics are represented on the horizontal axis. The diagram's vertical axis represents the distinction between rational accounts, portraying

[3] Scott (2008, 171-4) offers a critical assessment.

performance information as tightly linked to the substance of organizations, and interpretive accounts, emphasizing reform values, rhetoric, and ideas that are distinct and often disconnected from the objective conditions performance indicators are presumed to measure. Though the multiplicity of perspectives articulated by observers of results-model reform do not always fit neatly into the resulting two-by-two matrix, this diagram offers at least a point of reference for the discussion that follows.

By way of illustration, Figure 1 identifies four types of perspectives on performance measurement, each associated with particular quadrant. In the diagram's northwest corner, the *cult of efficiency* draws together March and Olsen's administrative orthodoxy with the treatment of measurement as a rational organizational process, tightly linked to the substance of the organization. With its roots in a progressive era conceptions of scientific administration, this analytic tradition is by a significant margin the longest lasting and still the most influential in thinking about the measurement of government results.[4]

Figure 1. Results-Model Reform:
Mixed Results and Multiple Perspectives



Scholarly accounts of results-model reforms tend to direct readers south and east from the cult of efficiency (sometimes both at once). Traveling southwest are scholars emphasizing what I label the *constructive construction* of organizational performance and quantitative measurement. This perspective draws on interpretive and constructivist theories of information and organizations, while retaining the basic normative commitments of the broader field. Feldman and March (1981) observe that organizations collect more information than they use, arguing that the collection and processing of data serves vital symbolic and normative purposes, signaling a commitment to rational decision-making. Performance measures offer little to guide decision-makers faced with tasks and conditions that are complex, novel, and unbounded (Schwenk 1989). Actors must contend with ambiguity, problems

[4] Among the truly under-appreciated progressive-era students of results-model reforms, Mabel L. Walker concludes her 1929 analysis of the "results" of municipal government, stating, "The government of a modern city is a tremendously complicated and difficult job. Politics is no longer sufficient to the task; we must make of municipal government a science" (Walker 1929: 134).

characterized by multiple conflicting conceptualizations, which cannot be resolved by new or better information. In this view, performance management is a social and normative, not a technical exercise (March and Olsen 1976).[5]

Building on this research tradition, Moynihan (2005a, 2005b) has developed a robust and varied body of scholarship examining the symbolic, rhetorical, and interpretive dimensions of performance management. While performance measures do not faithfully represent objective organizational conditions and their adoption is often more symbol than substance, Moynihan maintains a focus on the capacity of managers to learn from and make use of performance information. He writes, “Performance management reforms give motivated managers a license, even if it is largely a rhetorical one, to pursue organizational change and actual performance improvement” (Moynihan 2008, 192). Moynihan sees in performance information the capacity to construct and facilitate dialogue, promote organizational learning, and even to shape the terms on which policy decisions are made.[6]

Toward the northwest corner of Figure 1, critics of results-model reform draw on very different intellectual traditions, rooted in historical scholarship examining the development of quantification as an instrument of social and political *remote control*. Quantitative measures reduce and mobilize information, extending the capacity to observe and even, in some cases, the power to manipulate conditions across distances and boundaries far beyond what might previously have been imagined. Rooted in complex, long-term social and technological developments, this varied historical scholarship places the capacity to process and transfer information at the center of fundamental changes in the distribution of power, leading to dramatic centralization in states and the emergence of bureaucratic organizational forms (Beniger 1986; Desrosieres 1998; Porter 1995). Students of performance measurement operate at a somewhat smaller scale, but the claims are similar: The capacity of numbers to reduce and mobilize information alters the balance of control within and between organizations. Along these lines Beryl Radin (2006, 27-9, 54-86) argues that quantitative performance measures undermine humanistic values, erode trust, and erode professional autonomy and self governance.

Finally, the southeast quadrant of Figure 1 is occupied by the *shell game*,[7] where performance measures are more than vacant, but serve as useful distractions in the service of powerful interests. Adherents to

---

[5] March (1994: 179) writes, “Decisions are seen as vehicles for constructing meaningful interpretations of fundamentally confusing worlds, not as outcomes produced by a comprehensible environment.” March and Sutton (1997) go still further, arguing that the very enterprise of attempting to measure and discern the causes of organizational performance places analysts between a veritable rock and a hard place. Researchers educated in the social sciences recognize the complications inferring causation from correlations between measurable conditions and organizational performance, but they are employed to draw just these inferences.

[6] Moynihan’s (2005) study of simple- and complex-learning in two state departments of corrections suggests an intriguing additional feature of results-model reform politics. In one of the cases, he finds state corrections officials using performance information not only to rethink in some basic ways how they operated, but to influence the political environment. Moynihan argues that the example, “…demonstrates how [managing for results] not only enabled double loop learning, but also shaped the external environment as part of a pragmatic rationale for significant policy change” (Moynihan 2005: 213).

[7] The shell game is a classic “short-con” dating back to the Middle Ages. A “mark” is challenged identify which of three shells hides a small ball or “pea.” Using sleight of hand, the conman ensures that the mark cannot win. Participants in the con like “the shill” act as if they are part of the crowd of observers, egging the victim on. The analogy is useful here because of the image it evokes – a subject’s attention is manipulated in order to distract from what is “really” going on. For a brief review of the shell game’s fascinating history, see: (http://en.wikipedia.org/wiki/Shell_game, accessed: 9/23/09).

critical and post-modern approaches sometimes approach this perspective. Indeed, a strong reading of Deborah Stone's *Policy Paradox* (2002) suggests elements of the shell game.[8] However, the "good government" inclinations of must public administration scholarship and practice resist such a bluntly cynical account. That said anyone who has suffered through the incessant performance management "cheerleading" proliferating over the last two decades may have given this notion a passing consideration. Confronted with credulous descriptions of measurement, causation, and behavioral responses to structured incentives and by PowerPoint presentations that flirt with the line between aspiration and fantasy, one can be excused from passing from skeptic to cynic. More importantly, however, the shell game represents a sort of ideal form. Whereas the cult of efficiency takes performance measurement as a concrete representation of organizational substance, the shell game treats this activity as utterly unhinged.

The next two sections begin with the two legacies most commonly articulated by PART's developers and advocates: that its assessments serve as a framework informing decisions and its clarity and availability online enhance government transparency. I find a mixed assessment of each, highlighting variations on the perspectives outlined above.

1) Framing Decisions

The central premise of all performance- and results-model reforms is that government will be improved by measuring its results. Though ideas and rhetoric vary from mechanistic to vague, results-model reforms aim to frame policy and administrative decisions around performance information. PART's designers sought to "make use" of the performance measures developed by agencies under the Government Performance and Results Act (GPRA), which many hoped might be honed as an instrument to identify and eliminate ineffective government programs. Consistent with an account of PART as institutional bargaining (Dull 2006), Lewis and Gilmour (2006a, 2006b) find both that traditionally Democratic programs received systematically lower scores and that PART scores tended to influence budget allocation for Democratic programs and not for those traditionally supported by Republicans.

Yet, by most accounts, PART left little and largely inconclusive evidence of influence either on the Hill or as an instrument for improving agency effectiveness (Blanchard 2008; Gilmour 2008, Moynihan 2008). Even during a period of unified Republican control and disciplined Republican majorities in both chambers of Congress, PART's assessments found little support in congressional budgeting. Though PART scores correlated with presidential budget proposals (Gilmour and Lewis 2006a), it found little support in a notoriously resistant congressional appropriations process (Norcross and Adamson 2007). As a matter of institutional incentives, PART always faced a significant challenge from interest groups protecting federal programs from cuts.[9] Meyers (2009, 219) notes that resistance is accentuated by congressional organization, noting that despite significant progress in developing performance

[8] Stone (2002, 29) asserts, "…crucial information is very often *deliberately* kept secret…Secrecy and revelation are tools of political strategy, and we would grossly misunderstand the character of information in politics if we thought of it as neutral facts, readily disclosed."

[9] One longstanding account for the failure eliminate programs suggests government organizations may be "immortal" (Kaufman 1966), though recent empirical work challenges this generalization. An intriguing analysis of federal programs between 1971 and 2003 drawn from the Catalog of Federal Domestic Assistance (CFDA) by Berry, Burden, and Howell (2009) shows that programs are particularly vulnerable to "death" or "mutation" under partisan turnover and economic decline.

information at the agency and program levels, increasing responsiveness and accountability on the Hill requires a realignment of the committee system, comprehensive reform he concedes faces dim prospects.

2) Increasing Transparency

Among PART's important legacies is its contribution to the broad shift in government information, embracing internet technology in innovative ways with the potential to transform the politics of government transparency. In the words of OMB's Dustin Brown, "The fact that we made ...over 1,000 assessments available online with sometimes painstaking detail attached ... to anyone with an Internet connection is a tremendous accomplishment" (Newell 2008). Thus, PART took up contemporary information and research technologies, synthesizing and transferring standardized, salient information. By embedding information in a results-logic, and simplifying the terms on which program accomplishments are gauged, it enhanced transparency by presenting the results using a compellingly designed website www.expectmore.gov.[10] This logic, if not the URL, has been enthusiastically embraced by the Obama administration as well. In so doing, they extend among the very basic traditions in reform politics, the hope that new and better public information will enhance democratic deliberation. A century ago, the New York Bureau of Municipal Research's Henry Allen hoped that budget reform would both enhance government efficiency and help to create a "socialized intelligence," an educated citizenry that would ensure democratic accountability as the state grew in size and complexity (Kahn 1997).[11]

If the public seemed less than enthused by PART's quantitative scoring, academic researchers responded with keen interest. PART scores provided ready variables, where reliable indicators are often sparse and time-consuming to collect and code. As a result, PART scores are the subjects of a growing list of scholarly papers (Gilmour and Lewis 2006a, 2006b; Jung and Rainey 2009; Olsen and Levy 2004; Stalebrink 2008). Most compelling to date are analyses by David Lewis (2007, 2008) finding that PART programs headed by political appointees receive systematically lower scores for program management than those headed by careerists. Moreover, consistent with broader research on executive tenure and turnover, Lewis finds a curvilinear relationship between leadership tenure and program assessments, rising steeply at first then leveling off as administrators reach longer periods of service.[12] Even critics of PART who question its process and implicit ideological aims may concede the value of cross-agency evaluations, which make compelling research like Lewis's possible. Indeed, it is notable that among the

---

[10] Thus program assessments are reduced to a series of yes's and no's. The professional search for better measures results in measures that are more sophisticated – that is to say they get more complicated. More often than not, sophistication comes at a direct cost to measurement's other stated objective – information that is accessible to the public. As Clarence Ridley and Herbert Simon observed in the 1943 edition of *Measuring Municipal Activities*, "the citizen does not yet have an entirely satisfactory yardstick, nor are prospects too bright that he will obtain it in the visible future" (Ridley and Simon 1943; Lee 2003: 77). Porter illustrates this dilemma in the context of the French experience with social statistics (Porter 1994: ch. 4). For a contemporary review of pessimistic and optimistic accounts of public responsiveness to quantitative indicators, see: Gormley and Weimer 1999.

[11] Graham's *Democracy by Disclosure* (2002, 3) develops a parallel account regarding the influence of technology on regulation, writing, "The Internet has enhanced the power of disclosure by shattering a seemingly immutable law of communication: in-depth information about risks could be shared among experts; only superficial information could be shared with broad audiences. Trade-offs were inevitable between the richness of information and its reach." Graham contends that new information technologies have created the capacity to manipulate large amounts of sophisticated data, altering the dynamics of asymmetric information in politics and markets.

[12] Dull and Roberts (2009) briefly review this empirical literature.

stated goals for the Obama administration's performance evaluation framework is to, "allow comparisons across programs and agencies, and show trends over time…" (Zients 2009).

Christopher Hood (2006) traces the concept of transparency in contemporary public sector governance to political theorist Jeremy Bentham. For Bentham, the free flow of information was a vital feature of well-ordered political institutions. He saw publicity as a critical source of accountability and security against corruption of official power, writing, "The efficacy of this great instrument extends to everything – legislation, administration, judicature. Without publicity, no good is permanent; under the auspices of publicity, no evil can continue."[13] But, Bentham's information politics is perhaps most famously illustrated by his "Panopticon" design. Bentham imagined the behavior of a disordered population – prisoners or the poor – could be corrected by an institutionalized "all seeing eye." Their actions rendered completely transparent, subjects of Bentham's "Panopticon Prison" lack the discretion to engage in unsanctioned behavior and thus would learn to govern their own actions. In the tradition of enlightenment "statists" (Porter 1995), Bentham saw transparency as vital source of useful knowledge to inform the development of state policies. Of course, though his broader account of the "power of publicity" presages much contemporary thinking about policy and governance, to contemporary readers Bentham's Panopticon evokes the use of "thin" quantitative information to increase the reach of state power, sometimes with consequences that are profoundly destructive consequences (Scott 1998).

Thus the account of performance measurement as remote control parallels a broader account of quantification as an instrument of control, which has supplanted professional knowledge. "[T]he transition from expert judgment to explicit decision criteria did not grow out of the attempts of powerful insiders to make better decisions," writes historian Theodore Porter (1995, x), "but rather emerged as a strategy of impersonality in response to their exposure to pressures from outside." Beryl Radin portrays Bentham's "numbering of the modern world" as a sort of cultural affliction. "There is perhaps no element within the performance measurement process that is more important than the reliance on numbers and quantitative presentation of accomplishments," Radin (2006, 27) argues. This quantification fixation reduces the capacity of analysts and organizations to recognize complexity. Increasing reliance on numbers distorts the social organization of policy and administrative expertise and threatens professional autonomy through the encroachment of market and hierarchical modes of organization (Radin 2006, 57-62).[14]

Deborah Stone's *Policy Paradox* (2002) places the proliferation of numbers in a broader cultural commitment to rational decision. Like puzzles or poems with hidden complexity and potent symbolism they require careful interpretation. She too emphasizes the use of measurement as an instrument of control. "The power to measure is the power to control," Stone (2002, 186) contends, asserting, "Measurers have a lot of discretion in their choice of what and how to measure." But for Stone the power of numbers rests in their complexity rather than their coercive capacity. The only remedy rests in deciphering embedded assumptions and by mastering the art of decision-making in a sometimes intentionally misleading policy environment.

There is perhaps at least an element of the policy paradox in PART. Again, analysis of PART scores suggests systematic skew and a tendency for PART assessments to weigh more heavily on traditionally

---

[13] Quoted in Lieberman 2007, 12. Bentham, Jeremy, Political Tactics, Michael James, Cyprian Blamires and Catherine Pease-Watkin (Collected Works of Jeremy Bentham, Oxford 1999: 37.

[14] Wilson (1989, 228) similarly observes the tendency in military history of commanders to employ new technologies—from the telegraph to the computer—to reduce the discretion of frontline officers.

Democratic programs (Lewis and Gilmour 2006a, 2006b). Moreover, on closer inspection, some observers found PART was neither as complete, nor as clear as advertised. Despite OMB's claim that PART covered all federal programs, the scope of PART's scrutiny was something less than comprehensive in a few areas. Tax expenditures are among PARTs glaring omissions. For example, PART did not examine the low income housing tax credit was not examined by PART even though it represents the single largest source of low-income housing construction funds (Posner 2009, 238), nor did it examine the equally far-reaching Earned Income Tax Credit. Others found the apparently uncomplicated objectivity of the PART assessments less than enlightening. A passage from one June 2006 House Appropriations Committee report offers a particularly colorful depiction. In a section entitled "Quality of Budget Documents," the committee weighs in on the helpfulness of PART assessments for decision-making.

> For years, the Committee has directed departments and agencies to improve the budget justification document quality and presentation by including relevant and specific budget information. While the Committee has seen some improvement in a few submissions, most justifications continue to be filled with references to the Program Assessment Rating Tool (PART), drowning in pleonasm, and yet still devoid of useful information (U.S. House 2006: 4).

Who says Hill staffers don't read? Pleonasm, or using unnecessary words to express an idea, certainly does not imply the clarity and decisiveness that were the stated objectives of PART's developers.

3) Signaling Commitment

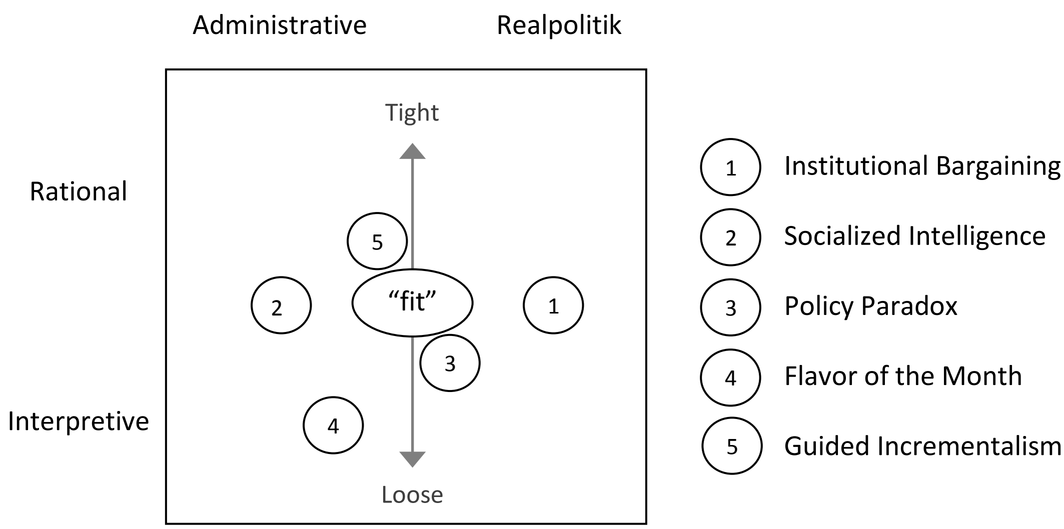
Anecdotally at least one area many observers concede PART and the Bush administration's performance budgeting initiative had an impact is on the performance measures agencies collect and report. At a minimum most agency administrators were eager to avoid having a program tagged "ineffective," or receiving a "red" on the PMA report cards. If PART had not gained their attention before, it is easy to imagine how such a mark might focus the mind of reputation-conscious agency leadership. Administrative reform initiatives like PART are firmly rooted in the politics of attention. Both the credibility and use of performance measures are shaped by clear signals of leadership commitment (Dull 2009). Though by no means universally accepted, PART signaled the administration's commitment to the use of performance measurement to guide budgeting and management with discernable consequences for agency use of performance measures.

That said, one wonders about the depth of PART's influence in the agencies. Are changes in an agency's external reporting indicative of some underlying change in agency management, or did PART simply become an exercise in compliance, as so many Bush administrations early on spoke of GPRA? Are they both simply layers of administrative sediment in Paul Light's *Tides of Reform* (1997)? The *shell game* perspective on performance measurement takes this account a step further. Not only did PART become an exercise in compliance, empty adoption disconnected from the substance of agency administration, but the political principals who initiated and sought to gain reputational advantage from it expected nothing more. Reform, in this account, is nothing but "cheap talk" (Dixit and Nalebuff 1991) or "myth and ceremony" (Meyer and Rowan 1977).

A less cynical account – but one with similar consequences – is that Bush administration's really "meant it," that they were committed to PART, but that their commitments lacked credibility.

Managers and employees form discernable judgments about leaders' credibility and rely on those judgments to guide their own behavior. Prior experience often leaves agency personnel feeling reform initiatives waste time and effort, with little consequence for organizational performance. At the center of contemporary reform thus are questions about the credibility of the commitments policy makers and administrators make to changing how government agencies operate. Stated succinctly, credibility is a necessary condition for performance management. If managers fear that leadership is not credible—if they suspect leaders lack competence, authority, or trust—the capacity to produce credible information will not take hold. As reforms come and go, without a hidden agenda or bad intent, agency personnel understandably begin to think of these initiatives as the *flavor of the month*, utterly disconnected from the substance of their jobs.

Figure 2. Multiple Perspectives:
A Few Mixed Forms



Multiple Lenses and Mixed Forms

Figure 2 locates a series of mixed forms identified through the discussion. The Bush administration's effort to frame budgetary decisions using PART was, I have argued, both a response to and ultimately a casualty of *institutional bargaining* (1). Like other venues of policy influence in the American system, budgeting is a shared power. Presidents must bargain for budgetary leverage across the centers of power on the Hill (Dickinson 1996; Neustadt 1990). The logic of institutional bargaining is realist, focused on PART's capacity (or failure) to enhance presidential policy influence, but the link between PART and policy influence is itself subject to interpretation and bargaining. While a handful of programs targeted by the Bush administration using "ineffective" PART ratings were actually reduced or eliminated (Norcross and Adamson 2007), the consensus seems to have been that most of PART's audience on the Hill found its assessments neither informative, nor particularly persuasive (Radin 126-7).

Now more than a century old, Henry Allen's *socialized intelligence* (2) likewise straddles the blurry boundary between interpretation and positive or rationalistic conceptions of the link between representation and reality. Allen's ideal, which runs like a thread through subsequent budgeting and management reform advocacy, holds that new techniques for organizing and displaying information will not only improve government, but educate the public and enhance the public's political efficacy, supporting the aims democratic deliberation and accountability (Kahn 1997). This ideal is evident in the adoption of both GPRA and PART,[15] and it is on precisely this point that the *policy paradox* (3) perspective diverges. This perspective sees efforts to shape public perceptions of government performance through a realist lens, not simply around the stated intention to enhance democracy, but as a method of framing – perhaps even obscuring – information and interests. The *flavor of the month* presumes no hidden agenda, but suspects performance measurement like the many reform initiatives that preceded it will fail to sink in and will sooner or later be replaced by the next fad or flavor.

4) Constructing Dialogue

Donald Moynihan's (2008, 207) recent examination of PART takes a clear-eyed view of this tendency for reform to come unhinged:

> Performance management doctrine often appears to float away from reality, inflated by its value as a political symbol, the beguiling simplicity of theoretical models that exclude most of the realities of management in government, and best practice stories related in breathless conviction. [16]

Embracing the language and interpretive focus of social constructivism, Moynihan acknowledges that a common constructivist response to the ambitions and rhetoric of results-model might be to reject the enterprise of performance measurement out of hand. Looking across state- and federal-level reform initiatives, he observes their adoption is often partial, misunderstood, and motivated by symbol more than substance. But, particularly at the agency level, he observes the potential for performance information to construct "interactive dialogue" in ways that lend clarity to complex and ambiguous goals and conditions, to shape organizational culture and commitment, and even to reframe policy problems in creative, potentially transformative ways.

This *constructive constructivist* perspective acknowledges the disconnect that often characterizes performance measurement, but sees examples – particularly at the level of agency management – that performance information enhances organizational learning. "The good news," he writes, "is that performance reporting mandates can help learning, if they interact with the right organizational conditions" (Moynihan 2008, 185). Intriguingly, PART represents an exception to the general pattern of

[15] As President Clinton stated when he signed GPRA into law, "The legislation itself mainly involves the inner workings of government, things that most people don't think about and maybe don't ever want to think about. It requires the formulation of strategic plans, of setting yearly goals and targets for every program; of measuring and reporting how well programs actually perform compared to the targets set for them, and more accountability for achieving results. But we should view this structure in much simpler terms, terms that every American should be able to identify with. The law simply requires that we chart a course for every endeavor that we take the people's money for, see how well we are progressing, tell the public how we are doing, stop the things that don't work, and never stop improving the things that we think are worth investing in" (Clinton 1993: 1540-1).

[16] This quote is included in Khademian (2009, 993). She summarizes the essence of Moynihan's account noting, "Put differently, the effort to exert political control in practice can be a resource for creative performance."

performance information embraced at the agency level. Instead, it was undertaken by OMB as a consequence of the perception that agencies had failed to take advantage of the performance information generated under GPRA (Moynihan 2008, 198). Indeed, the capacity of PART to promote constructive dialogue between the agencies and OMB's program examiners was a casualty, in part, of the threat that PART scores might be used to justify budget cuts (Moynihan 2008, 183).

5) Accumulating Knowledge

Central to PART and all results-model reforms is the iterative, routine basis for collecting and analyzing performance information. They create both a clear period-by-period trend intended to enhance decision-making and accountability; as well as new or altered routines for data collection, interpretation, and reporting[17] By clarifying environmental feedback and structuring decision-processes, routines like those that developed around the PART process hold the potential to accentuate and enhance ongoing organizational adaptation and learning (Becker, Lazaric, Nelson, and Winter 2005).[18] Viewed closely, routines are more than patterned behaviors; they are ongoing, "effortful accomplishments" (Feldman 2000: 613; Pentland and Rueter 1994),[19] producing endogenous change. Feldman and Pentland (2003) represent the dynamics between three aspects illustrated in Figure 3.

→ **Ostensive** – The internal representation of the routine. It may be formalized like standard operating procedure or analytical technique. Or, may be informal, such as a "taken-for-granted" norm.

→ **Artifactual** – The forms, copies, schedules, electronic records that record the imprints and serve as the "indicators" of the performative and ostensive aspects.

→ **Performative** – The ongoing "effortful" and "improvisational" enactment of the routine, the performative aspect represents the physical practices associated with carrying out the routine

The ostensive aspects are abstract representations of the routines in the minds of actors, reflect individual "world views."[20] The artifactual aspects of routines are perhaps the most easily observed

[17] Harry Hatry defines performance measurement as, "…measurement on a regular basis of the results (outcomes) and efficiency of services or programs." Hatry emphasizes two components of this definition. First, measures are collected and reported regularly, thus providing a basis for evaluating progress. Second, outcomes are defined in terms of direct and indirect benefits to program "customers" defined in terms of agency or policy mission (Hatry 1999: 4; Italics in original). Regular reporting is an important distinguishing feature in the field of budget policy. For a discussion relating to the formalization of budget routines, see: Meyers 1994: 43-4.

[18] This account departs from traditional treatment of routines as sources of stability (Becker 2005). Described in research and in every day conversation as regular, patterned behavior, for March and Simon (1958) they are "performance programs," for Schumpeter (1955) they are "accustomed circular flows," a contrast to innovation (Winter 2006: 136). For Nelson and Winter (1982), routines are at once "genes" that capture distinctive and persistent firm attributes and "truces," stable patterns that diffuse conflict among interests. Routines enable coordinated action by enhancing the reliability with which actors predict others' behavior (Cohen et al. 2006; Cohendet and Llerena 2003: 274).

[19] Feldman (2000: 613) observes, "One can think of routines as flows of connected ideas, actions, and outcomes. Ideas produce actions, actions produce outcomes, and outcomes produce new ideas. It is the relationship between these elements that generates change."

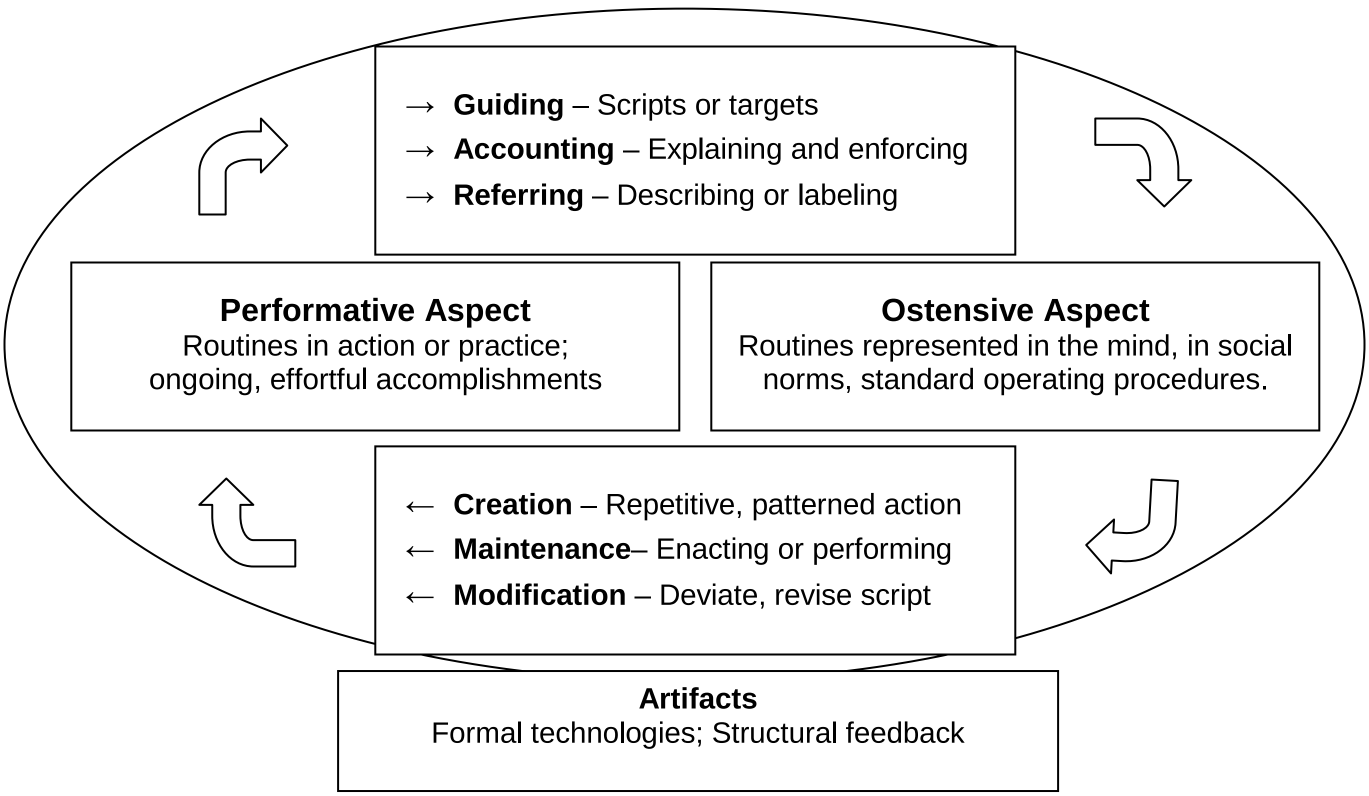
[20] Feldman and Pentland (2003: 101) contend that the ostensive aspect of routines should not be treated as a single entity, noting, "Each participant's understanding of a routine depends on his or her role and point of view

through the use of archival and other recorded data (March, Schulz, and Zhou 2000). But, it is the performative aspect, made up of practices or patterns of action,[21] that form the central element of the model. John Seely Brown and Paul Duguid capture the common theme running through much contemporary research on the social organization of knowledge:

> Part of the skill of work, all work, then is routinization, adapting the particulars of the world so that they fit within the general schemas of the organization. The gap to be bridged lies between reality and process, and it is bridged by the improvisation inherent in practice – so deeply inherent that the practitioners themselves are barely aware of it (Brown and Duguid 2000: 108-9).

Figure 3. Organizational Routines

Dual Aspects of Routines (Feldman and Pentland 2003; Pentland and Feldman 2005)



A long-standing tradition in the study of human social organizations is rooted in the embodiment of skill, knowledge of practice, or "tacit" knowledge that grows from performing the same tasks or routines over

---

…The ostensive aspect of the routine gains in apparent objectivity and concreteness as the views of different participants come into alignment. But it is still only a partial picture because it does not include the performances."

[21] Feldman and Pentland (2003: 102) emphasize the improvisational nature of the performative aspect, noting, "Practices are carried out against a background of rules and expectations, but the particular courses of action we choose are always, to some extent, novel."

and over again, and cannot be articulated (Mantzavinos 2001; Polanyi 1958).[22] The development and valuation of knowledge thus are embedded in a political and normative context. Exchange is embedded in networks of social ties (Granovetter 1985; Uzzi 1997), and the employment contract is rooted not just in extrinsic incentives, but in internal or intrinsic motivation (Austen and Larkey 2002; Osterloh and Frey 2000).[23] Information to oversee and influence organizations depends crucially on the normative commitments of individuals and organizations (Brehm and Gates 1999; Miller 1992; Golden 2000).

Contemporary results-model reforms embody the recognition that normative commitments are vital where reform takes hold. These initiatives call on agencies to align data synthesis and transfer with existing intra-agency culture and commitments.[24] Contemporary reformers are no less ambitious in their schemes to reshape policy administration than were their Progressive-era predecessors. Yet, the models guiding reform are informed by more than a century of frustrated efforts at organizational engineering, and a now-massive management literature emphasizing the inherently social nature of knowledge and practice. Michael Barzelay and Colin Campbell (2003) borrow the concept of *guided incrementalism* to describe this approach to strategic planning.[25] Administrators concede cognitive and organizational the barriers to comprehensive strategic decisions, but they seek to retain strategic influence, searching for opportunities to out-maneuver the forces of incrementalism and organizational inertia. They do so, Campbell and Barzelay observe, by arranging capacities and dispositions within the organization so that they might respond to opportunities as they emerge – a strategy of cultivation rather than design.

Though much remains to be seen, these ideas are evident in the stated commitments of Obama administration officials seeking to build on PART's legacy. In the words of OMB Deputy Director Jeff Zients (2009):

> "…it is time to pay far more attention to the use of Federal performance information as a powerful performance-improving tool – useful for communicating priorities, progress, and raising issues; for illuminating what works that should be continued and what does not work that needs attention; for motivating the best from our workforce and our service delivery partners; and allocating scarce resources wisely."

---

[22] This perspective grounds a broad research program in the fields of organizations and strategy emphasizes what are variously termed competence-based (Nelson and Winter 1982), resource-based (Conner and Prahalad 1996), knowledge-based (Liebskind 1996), and learning theories of organization (Levitt and March 1988).

[23] Another consequence is political conflict. Routines do not simply reduce task-based uncertainties, but also mediate among competing organizational interests (Caiden and Wildavsky 2004; Nelson and Winter 2002). As March (1994: 255) puts it, "Decision-related knowledge is rarely innocent…systems of information and knowledge are instruments of power that favor those who can control them at the expense of those who cannot."

[24] In 2004, OMB Deputy Director Clay Johnson testified about the administration's PART initiative, "the biggest challenge is to change habits," calling PART, "a new habit, a new way of thinking about budgets," and "a new thought process" (Johnson 2004).

[25] This concept draws on Amitai Etzioni's concept "mixed scanning," which Etzioni develops in response to the limitations of both rational and incrementalist models of decision-making. Etzioni writes, "What is needed for active decision-making is a strategy that is less exacting than the rationalistic one but not as constricting in its perspective as the incremental approach, not as utopian as rationalism but not as conservative as incrementalism, not so unrealistic a model that it cannot be followed but not one that legitimates myopic, self-oriented, non-innovative decision-making" (Etzioni 1968: 282-3).

In the spring budget guidance to the agencies, OMB asked each agency to identify a handful of "high priority performance goals," and OMB is working with the agencies to finalize these objectives. Zients (2009) adds, "As agencies achieve these near-term goals, new ones will be added."

6) Governing by Abstraction

A blunt but useful analogy for the task of performance measurement is that the process is like asking program administrators: "What is your organization's dependent variable?" Administrators are essentially tasked with hypothesizing an effect traceable to the programs and activities they administer, and operationalizing this causal model using one or a handful of quantitative indicators. The analogy is useful because of how inadequate most performance information looks against the standards of basic quantitative reasoning. Indicators of questionable validity, under-specified causal models, little or no information about error or uncertainty – through this lens most performance information looks like an inferential minefield. Much performance measurement is bad management because it is bad social science. The analogy is also useful because frames performance measures as operationalized variables – or, as Arthur Stinchcombe (2001) might label them, as formalized abstractions.

Stinchcombe's *When Formality Works* (2001) offers a final perspective on PART's legacies – and a useful means of organizing and reflecting on the confounding multiplicity of perspectives reviewed to this point. As the title suggests, the book responds to what Stinchcombe argues is a tendency in organization theory to associate the formal aspects of organizations with pathology, structures disconnected from the substance of organizations. Returning to Figures 1 and 2, most contemporary thinking is fixed firmly in the southern quadrants. In Stinchcombe's view, formalization should be conceptualized as a variable, "as active, as a direction of change," subject to a dynamic context. Formality and informality are not opposed, he argues, but represent rich variations in flexibility and concreteness across different forms of social organization. Flexibility depends critically on complementary rigid forms. "Formality is not the opposite of the informal substance but a refined version (or versions) of it" (Stinchcombe 2001, 3).

Formalities – or "abstractions that govern" – are defined by the relation to varieties of informality, which Stinchcombe organizes into three categories. *Classical informality* consists of the parts of social life that are left out of governing formalities, sometimes specifically to subvert rigid or empty formalities.[26] Stinchcombe is more interested in two other varieties, constituting the "loose joints" between different varieties of formalization. *Informally embedded formality* represents informal activities or arrangements that are formalized elsewhere. The discretion of a "certified" professional, for example, gains authority by reference to "a complex system of abstractions governing his or her activity," represented by the professional certification (Stinchcombe 2001, 34). *Formality being constructed* responds to evidence that a formality at odds with the reality it aims to govern either by creating a new formal system or patching up the old one.

[26] Stinchcombe (2001, 8) contends that much of the sociological focus on classical informality aims demonstrating that "it is not irrelevant after all." He writes, "Sometimes [classical informality] can be systematically subversive of the formal order, as in informal secrecy about bribery, or a mobilizing conversation for a wildcat strike that looks, from the outside, like friendly conversation. Sometimes it can be very the very first stages of improving the abstraction system, as when a conversation at a professional meeting suggests a new and more accurate way of making the first cut in the set of applicants for a position. In any case, the general point is that the formal system never pretended it was going to govern the activity in question, so it is informality by default."

Formalization, or the degree to which abstractions can be taken as "fact," so that "most people, most of the time, do not have to go behind it," varies relative to three sets of mechanisms. *Cognitive adequacy*, or qualities through which an abstraction yields a "picture" capable of guiding action, which in turn is characterized by four features: 1) Accuracy or the error in the abstraction; 2) Cognitive economy, meaning the abstraction includes only information that is "factually sufficient" takes a form that is "easy to think about;" 3) Sufficiency, meaning the abstraction captures all relevant information; and 4) Scope, meaning the abstraction covers the full range of relevant conditions (Stinchcombe 2001, 23-7).

The second, *communicability,* is the capacity of an abstraction to enable the discourse vital to governing social activity. It is defined by three characteristics: 1) Transmissibility means that those using the abstraction understand, believe it is valid, and it is being applied correctly; 2) Transparency is the degree to which everyone engaged in the abstraction process has the ability to examine all aspects of the abstraction; and 3) Durability means the quality of the communication involving the abstraction does not decay over time or with multiple transmissions (Stinchcombe 2001, 30-2). Finally, *trajectory of improvement* captures the dynamic capacity of abstractions to adapt and apply across different contexts,[27] defined by three sets of factors: 1) Correctability means that an abstraction is on a trajectory of correction, improving through feedback and adapting to changing conditions; 2) Robustness is defined by its capacity to "abstract correctly" across different situations and circumstances; and 3) Algorithms and reasoning devices are formal mechanisms relating abstractions to one another.

Stinchcombe's framework, as this partial summary suggests, is a characteristically dense and, again characteristic of his work generally, his case material eclectic. Appeals courts, the administration of immigration law, scientific discourse – Stinchcombe even includes a chapter describing a university course he took on how to read construction blueprints. Drawing the material together is a simple, organizing theme, "…socially instituted abstraction systems, when properly designed, are a very useful way of governing some kinds of social action" (Stinchcombe 2001, 181). Indeed, at the level of organizations, the interplay between formal and informal arrangements is not a matter of one dominating or subverting another, but of variation on a continuum from flexibility to rigidity, both vital to effectiveness. "Organizational flexibility…depends on the structure of rigidities, just as running depends on a stable relation between flexible muscles and joints" (Stinchcombe 2008, 132).

Budgets are, of course, classic examples of the interplay between formal and informal aspects of organizations (Caiden and Wildavsky 2002). Returning to Figure 3, budgeting routines operate through the interplay of formal artifacts (balance sheets, formal guidance, computer technologies), shared ideas (expectations based on professional training and prior experience), and the ongoing "effortful" performance of budgeting tasks and duties. Budgets organize resources into classes flexibility or liquidity, "debt service being less liquid, operating expenses more liquid, and cash balances most liquid," where an organization's overall strategy, "depends on matching the degree of internal liquidity of resources to the volatility of conditions" (Stinchcombe 20010, 127-8).

The idea that government can be improved by measuring its results has deep roots and varied legacies (Ridley and Simon 1938; Walker 1929; Williams 2003, 2004). A century ago, the link between resources and results was very much on the mind of the NYBMR reformers who developed the modern executive

[27] "Sociologists have a tendency to think that most formality is…fraud and ritual," Stinchcombe (2001, 36) explains, "One of the best ways to check whether it is likely to be fraud and ritual is to ask whether there is an effective set of devices to improve the abstractions in the system."

budget.[28] The line-item budgets that emerged from the first wave of budget reform represented a major advance in the control of government expenditures, clearly delineating by unit and category where government dollars are spent, the 'who and what' of government spending. However, planning the strategic allocation of resources and review of what those inputs have accomplished once the money is spent outstripped the capacity of existing technologies and institutions. [29] Reformers designed superior systems, but government organizations lacked the capacity or political will to make them work.

The same dynamics have fueled subsequent generations of reform, budget or analytical models outstripping state capacity and resistance among institutionalized interests – a pattern that will doubtlessly persist. Yet, in addition to broad cultural and political developments (Kettl 2002), two fundamental developments have formed the basis for integration of performance information into the process of government resource allocation and management. First, changing technology has transformed human capacities to collect, analyze, and transfer quantitative data.[30] Developing technologies have, in turn, helped to frame and fuel a second broad set of developments in the ways we think about and analyze organizational resources.

For decades, estimating firm value was treated largely as a "solved" problem (Mantzavinos 2001), subject to the rules and practices of cost accounting. Accounting offered a set of established, authoritative abstractions for governing management and market activity (Kaplan 1986). Yet, retrospective cost-accounting did not provide useful tools for estimating how much firm value was enhanced by so-called "intangibles" like knowledge and reputation. The growth of internet and other knowledge-based enterprises particularly during the decade of the 1990's drew increasing salience to the problem of accurate valuation among professional organizations like the Financial Standards Accounting Board (FASB 2004; Sherman 2001) and market actors. The balance scorecard and other combinations of financial and non-financial performance indicators developed in part as a means of external evaluation and control (Meyer 2002; Lev 2001).

Contemporary results-model reforms like PART extend the logic, developing a set of formal arrangements designed to value and govern the human dimensions of the enterprise: knowledge,

---

[28] Williams (2003: 648) quotes NYBMR's Henry Bruere on the use of metrics to inform the rational allocation of budgetary resources. "Reluctance to appropriate has been partly attributable to the uncertainty that funds requested would be used with fullest efficiency. But it has also been due to the fact that neither the requesting departments, the community, nor the board of estimate and apportionment, had definitely pictured the wide gap existing between health and charitable services now rendered and the health and charitable needs of the city."

[29] Allen Schick (1966) observes that it was only after experimenting with a variety of techniques that NYBMR's reform advocates reluctantly accepted line-item control budgets. Both NYBMR and the Taft Commission recognized these shortcomings, but found their reform ambitions frustrated. The line-item technique provided a workable solution to a pressing problem of control in the face of an increasingly robust state apparatus. "In an age when personnel and purchasing controls were unreliable, the first consideration was how to prevent administrative improprieties…" Schick explains, "Once installed, object controls rapidly gained stature as an indispensable deterrent to administrative misbehavior" (Schick 1966: 74).

[30] In a commentary accompanying a late edition of *Administrative Behavior*, Simon (1997: 249) observes, "With the development of information-processing technology, we have a growing capacity to consider interactions and tradeoffs among alternatives and consequences; to cumulate our understanding of fragments of the whole problem by embedding these fragments in comprehensive models." Regarding PPB, Schick (1966: 86) noted, "Without the availability of the decisional-informational capability provided by cost-benefit and systems analysis, it is doubtful that PPB would be part of the budgetary apparatus today. The new technologies make it possible to cope with the enormous informational and analytic burdens imposed by PPB. As aids to calculation, they furnish a methodology for the analysis of alternatives, thereby expanding the range of decision-making in budgeting."

attention, strategy, commitment, and reputation. These are traditionally the domain of scholars focused on the informal aspects of organizations. Yet, each is characterized by the same tensions and trade-offs Stinchcombe observes between liquidity and persistence. All organizations, but particularly for public organizations subject to the push and shove of institutional politics, the interplay between stability and flexibility is vital to sustaining robust organizational strategies.

This paper has no-doubt illustrated that there many factors and perspectives contributing to the logic of administrative reform. Much (though by no means all) of this attention tends to either assume performance measures are tightly linked to the substance of organizations or to fixate on how reform is disconnected from what "really happens" through informal organizational practices and arrangements. I see in Stinchcombe's framework a means by which to think carefully about performance measures as abstract formalities, embedding in a complex and changing interplay between formal and informal organizational arrangements.

Conclusion

This paper considered a variety of perspectives on the mixed legacies of the Bush administration's Program Assessment Rating Tool (PART), which departed bureaucratic life this year for the greener pastures occupied by its predecessors Planning Programming Budgeting (PPB), Management By Objectives (MBO), Zero-Base Budgeting (ZBB) – reform initiatives that expired, but never fully faded away. The conceptual muddle may be my own, but as a careful observer and a consumer of scholarly attempts to explain and understand the syncopated rhythms of administrative reform perhaps the single overarching theme I have taken away from thinking and writing about reform is its multiplicity of perspectives. This selective review draws together a range of these perspectives, briefly considering six of PART's legacies, shaping the context of future reforms.

PART's developers and advocates sought with mixed results to frame management and budgeting decisions, increase transparency, and signal the president's commitment to rational and efficient governance. Practitioners and observers saw in PART the capacity to construct dialogue, promote learning, and accumulate knowledge – again, with mixed results. Briefly reviewing each of these accounts, I engaged particularly the degree to which these perspectives emphasized the loose or tight link between the byproducts of reforms – performance information – and the substance of organizations. In the paper's final section I elaborate on the framework outlined in Arthur Stinchcombe's *When Formality Works* (2001) and make the case that his account of the complex interdependence of formal and informal organizational arrangements offers some perspective on the development of contemporary results-model reform. This perspective requires further development, but my hope is that these ideas will help to frame new thinking about how performance measures do and do not "work."

Sources

Allen, William H. 1907. *Efficient Democracy.* New York; Dodd, Mead & Company.

Arnold, Peri E. 1998. *Making the Managerial Presidency: Comprehensive Reorganization Planning, 1905- 1996*. Lawrence: University of Kansas.

Barzelay, Michael and Colin Campbell. 2003. *Preparing for the Future: Strategic Planning in the U.S. Airforce*. Washington, DC: Brookings Institution Press.

Becker, Markus C. 2005. "The concept of routines – Some clarifications." *Cambridge Journal of Economics.* 29(2): 249–262.

Beniger, James R. 1986. The Control Revolution: Technological and Economic Origins of the Information Society. Cambridge, MA: Harvard University Press.

Berman, Larry. 1979. The Office of Management and Budget and the presidency, 1921–1979. Princeton, NJ: Princeton University Press.

Berry, Christopher R., Barry C. Burden, and William G. Howell. 2009. "After Enactment: The Lives and Deaths of Federal Programs." University of Chicago: Working Paper.

Blanchard, Lawrence A. 2008. "PART and performance budgeting effectiveness." In F. S. Redburn, R. J. Shea, & T. F. Buss (Eds.), *Performance management and budgeting: How governments can learn from experience* Armonk, NY: M.E. Sharpe. 67-91.

Brown, John Seely and Paul Duguid. 2002. *The Social Life of Information.* Harvard Business School Press.

Chisholm, Donald. 1995. "Problem Solving and Institutional Design." *Journal of Public Administration Research and Theory.* 5(4): 451.

Desrosieres, Alain. 1998. *The Politics of Large Numbers: A History of Statistical Reasoning.* Boston: Harvard University Press.

Dickinson, Matthew J. 1996. *Bitter Harvest: FDR, Presidential Power, and the Growth of the Presidential Branch.* New York: Cambridge University.

Dixit, A. and B. Nalebuff. 1991. *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life.* New York: W.W. Norton and Company

Dull, Matthew and Patrick S. Roberts, 2009. "Continuity, Competence, and the Succession of Senate-Confirmed Agency Appointees, 1989-2009," Presidential Studies Quarterly. 39:3 (September): 432-453.

Dull, Matthew. 2006. "Why PART? The Institutional Politics of Presidential Budget Reform." *Journal of Public Administration Research & Theory*. 16(2):187-215.

Dull, Matthew. 2009. "Results-Model Reform Leadership: Questions of Credible Commitment." *Journal of Public Administration Research & Theory*. 19(2):255-284.

Feldman, Martha S. and Brian T. Pentland. 2003. "Reconceptualizing Organizational Routines as a Source of Flexibility and Change." *Administrative Science Quarterly.* 28.

Feldman, Martha S. and James G. March. 1981. "Information in Organizations as Signal and Symbol." *Administrative Science Quarterly* 26(2): 171-186.

Feldman, Martha. 1989. *Order Without Design: Information Production and Policy Making*. Palo Alto: Stanford University Press.

Feldman, Martha. 2000. "Organizational Routines as a Source of Continuous Change." *Organization Science.* 11(6): 611-629.

Gilmour, John B. 2008. Implementing OMB's program assessment rating tool: Meeting the challenges of performance-based budgeting. In F. S. Redburn, R. J. Shea, & T. F. Buss (Eds.), Performance management and budgeting: How governments can learn from experience (pp. 21-48). Armonk, NY: M.E. Sharpe.

Gilmour, John B., and David E Lewis. 2006b "Does Performance Budgeting Work? An Examination of the Office of Management and Budget's PART Scores," *Public Administration Review.* Sep/Oct 66(5): 742-753.

Gilmour, John B., and David E. Lewis. 2006a. "Political Appointees and the Competence of Federal Program Management." *American Politics Research* 34(1): 3-21.

Gormley, William T. and David L. Weimer. 1999. *Organizational Report Cards.* Cambridge: Harvard University Press.

Graham, Mary. 2002. *Democracy by Disclosure: The Rise of Technopopulism.* Washington, DC: Brookings.

Hatry, Harry P. 1999. *Performance Measurement: Getting Results.* Washington, DC: Urban Institute Press.

Hood, C. and G. Peters (2004). "The Middle Aging of New Public Management: Into the Age of Paradox?" *J Public Adm Res Theory* 14(3): 267-282.

Jung, Chan Su and Hal G. Rainey. 2009. "Program Types, Goal Ambiguity, and Performance in U.S. Federal Programs." Paper Prepared for the 10th Public Management Research Association Research Conference, Columbus, OH, October 2009.

Kahn, Jonathan D. 1997. *Budgeting Democracy: State-Building and Citizenship in America, 1890-1928.* Ithaca, NY: Cornell University Press.

Kaufman, Herbert. 1976. *Are Government Organizations Immortal?* Washington, DC: Brookings Institution Press.

Kettl, Donald F. 2002*. The Transformation of Governance: Public Administration for Twenty-First Century America.* Baltimore: Johns Hopkins University Press.

Khademian, Anne. 2009. "Ambiguity, Dialogue, and the Underbelly of Performance Management." *Journal of Public Administration Research and Theory.* 992-995.

Lee, Mordecai. 2003. "Is There Anything New Under the Sun? Herbert Simon's Contributions in the 1930's to Performance Measurement and Public Reporting of Performance Results." *Public Voices*. 6:2-3.

Lev, Baruch. 2001. Intangibles: Management, Measurement, and Reporting. Washington, DC: Brookings Institution.

Levitt, Barbara and James G. March. 1988. Organizational Learning. *Annual Review of Sociology*. 14: 319-40.

Levitt, Barbara and James G. March. 1990. "Chester I. Barnard and the Intelligence of Learning." In *Organization theory: from Chester Barnard to the present and beyond*, edited by O. E. Williamson. New York: OxfordUniversity Press.

Lewis, David A. 2007. "Testing Pendleton's Premise: Do Political Appointees Make Worse Bureaucrats?" *Journal of Politics.* 69 (November): 1073-88.

Lewis, David E. 2008. T*he Politics of Presidential Appointments: Political Control and Bureaucratic Performance.* Princeton University Press.

Light, Paul. C. 1997. *The Tides of Reform: Making Government Work, 1945-1995.* New Haven: Yale University Press.

Mantzavinos, C. 2001. Individuals, Institutions and Markets. Cambridge: Cambridge University.

March, James G. 1987. "Ambiguity and Accounting: The Elusive Link Between Information and Decision Making." *Accounting, Organizations and Society.* 12(2): 153-168.

March, James G. 1994. A Primer on Decision-Making. New York: Free Press.

March, James G. and Johan P. Olsen. 1976. *Ambiguity and Choice in Organizations.* Bergen, Norway: Universitetsforlaget.

March, James G. and Johan P. Olsen. 1983. "Organizing Political Life: What Administrative Reorganization Tells Us about Government." American Political Science Review. 77(2): 281-296.

March, James G. and Johan P. Olsen. 1989. *Rediscovering Institutions.* New York: Free Press.

March, James. G. and R. I. Sutton. 1997. "Organizational Performance as a Dependent Variable." *Organization Science.* 8(6): 698-706.

Meyer, John W. and Brian Rowan. 1977. "Institutionalized Organizations: Formal Structure as Myth and Ceremony." American Journal of Sociology. 83: 440-463.

Meyer, Marshall W. 2002a. Rethinking Performance Measurement: Beyond the Balanced Scorecard. New York: Cambridge University Press.

Meyers, Roy T. 1994. *Strategic Budgeting.* An Arbor: University of Michigan Press.

Meyers, Roy T. 2009. "The 'Ball of Confusion' in Federal Budgeting: A Shadow Agenda for Deliberative Reform of the Federal Budget Process," *Public Administration Review*. March/April: 69(2): 211-223.

Moynihan, Donald P. 2005a. Goal-Based Learning and the Future of Performance Management. *Public Administration Review.* 65 (2):203-216.

Moynihan, Donald P. 2005b. "Managing for Results in an Impossible Job: Solution or Symbol?" *International Journal of Public Administration*, 28:3,213 — 231

Moynihan, Donald P. 2008. *The Dynamics of Performance Management: Constructing Information and Reform.* Georgetown University Press.

Nelson, Richard R. and Winter, Sidney G. 1982. *An Evolutionary Theory of Economic Change.* Cambridge, MA: Belknap Press.

Neustadt, Richard E. 1990. *Presidential Power and the Modern Presidents.* New York: Free Press.

Newell, Elizabeth. 2008. "Legacy of program assessment tool unclear." *Government Executive.* September 16.

Norcross, Eileen and Joseph Adamson. 2007. "An Analysis of the Office of Management and Budget's Program Assessment Rating Tool (PART) for Fiscal Year 2008." Mercatus Center, George Mason University: Working Paper.

Office of Management and Budget. 2009. *A New Era of Responsibility: The FY 2010 Budget.* (http://www.whitehouse.gov/omb/assets/fy2010_new_era/Jumpstarting_The_Economy.pdf)

Olsen, Robert., and David Levy. 2004. "Program performance and the President's budget: Do OMB's PART scores really matter?"Paper presented at the Annual Meeting of the Association for Public Policy Analysis and Management, October.

Pentland, Brian T., and H.H. Reuter. 1994. Organizational Routines as Grammars of Action. *Administrative Science Quarterly.* 39:484-510.

Pentland, Brian T., and Martha S. Feldman. 2005. *Organizational routines as a unit of analysis. Industrial and Corporate Change.*

Porter, Theodore M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life.* Princeton, NJ: Princeton University Press.

Posner, Paul L. 2009. "Budget Process Reform: Waiting for Godot," *Public Administration Review.* March/April. 69(2):233-243.

Radin. Beryl A. 2006. *Challenging the Performance Movement: Accountability, Complexity, and Democratic Values.* Georgetown University Press.

Ridley, Clarence E., and Herbert A. Simon. 1938. *Measuring municipal activities; a survey of suggested criteria and reporting forms for appraising administration.* Chicago,: The International city managers' association.

Rudalevige, Andrew. 2003. The ''M'' in OMB: The Office of Management and Budget and presidential management of the executive branch, 1939–2003. Paper presented at the annual meeting of the American Political Science Association, Philadelphia.

Schick, Allen. 1966. "The Road to PPB: The Stages of Budget Reform." *Public Administration Review.* 26. December.

Schick, Allen. 1973. "A death in the bureaucracy: The demise of federal PPB." *Public Administration Review.* 33:146–56.

Schwenk, C. 1989. "Linking Cognitive, Organizational and Political Factors in Explaining Strategic Change." *Journal of Management Studies.* 26(2): 177-187.

Scott, James C. 1998. *Seeing Like a State: How Certain Schemes to Improve the Human Condition Have Failed.* Yale University Press

Scott, W. Richard. 2008. *Institutions and Organizations,* 3rd ed. Thousand Oaks, CA: Sage.

Sherman, Jay. 2001. "A Question of Intangibles." Treasury and Risk Management. September: 49-50.

Simon, Herbert A. 1997. *Administrative Behavior: A Study of Decision-Making Processes in Administrative Organization.* 4th ed. New York: Free Press.

Stalebrink, Odd J. 2008. "National Performance Mandates and Intergovernmental Collaboration: An Examination of the Program Assessment Rating Tool (PART)" *The American Review of Public Administration.* 20(10).

Stinchcombe, Arthur L. 2001. *When Formality Works: Authority and Abstraction in Law and Organizations.* University of Chicago.

Tomkin, Shelley Lynne. 1998. *Inside OMB: Politics and process in the president's budget office.* Armonk, NY: M. E. Sharpe.

Walker, Mabel. L. 1929. "Rating cities according to the services which their citizens are getting." *The American City*, 41, 130-134.

Wildavsky, Aaron B., and Naomi Caiden. 2004. *The new politics of the budgetary process. 5th ed, Longmand classics in political science.* New York: Pearson/Longman.

Williams, Daniel W. 2003. Measuring Government in the Early Twentieth Century. Public Administration Review 63 (6):643-659.

Williams, Daniel W. 2004. Evolution of Performance Measurement to 1930. Administration & Society 36 (2):131-165.

Winter, Sidney G. 2006. "Toward a Neo-Schumpeterian Theory of the Firm." Industrial and Corporate Change. 15(1): 125-141.

Zients, Jeff. 2009. Statement before the Senate Committee on Homeland Security and Governmental Affairs Subcommittee on Federal Financial Management, Government Information, Federal Services, and International Security. September 24.