JPART 20:i143–i159

THE STATE OF AGENTS: A SPECIAL ISSUE

Accountable Agents: Federal Performance Measurement and Third-Party Government

Melvin J. Dubnick
University of New Hampshire

H. George Frederickson
University of Kansas

ABSTRACT

Ordinarily research articles on public sector third-party governance, or the state of agents, turn to the subject of accountability in their conclusions. Rather than leaving it to the conclusions, this article takes up the subject of public sector third-party governance as a problem of accountability. To consider accountability, we studied contemporary performance measurement practices in the federal government, particularly as they are applied in five agencies of the US Department of Health and Human Services. Our findings are presented using a six-part "promises of accountability" heuristic which captured the many and varied uses of accountability in contemporary policy and administration discourse. We found that the characteristics of third parties and the nature of their principal-agent relations are a key determinant of accountability; bureaucratic and hierarchical controls enhance agency and program accountability; building rules and program policies into grants and contracts enhance accountability; agencies that practice lateral trust-based forms of "relational contracting" foster cultures of accountability; auditing and reporting requirements enhance program accountability; federal performance measurement regimes tend to be "top down" and "one size fits all;" federal performance measurement regimes are primarily executive branch forms of accountability; performance measurement often represents attempts to superimpose managerial logic and managerial process on inherently political processes embedded in the separation of powers; there is little evidence that performance-based accountability results in enhancing democratic outcomes or greater justice or equity.

INTRODUCTION

Whatever it is called—third-party government, the hollow state, the shadow bureaucracy, the blended public workforce, articulated chains of third parties, public programs and their partners, steering rather than rowing, the state of agents—it is THE public administration challenge of our time. American governments at all levels have moved steadily in the direction of taking managerial and service delivery functions out of the hands of civil servants and putting them in the hands of nonprofit and for-profit third-party contractors and grantees. Compared to the growth of third-party contractors and grantees, the cohort of federal civil servants has shrunk by nearly 25% (Light 1999).

Address correspondence to the author at gfred@ku.edu.

doi:10.1093/jopart/mup039
Advance Access publication on XX XX, XXX

Public management via contracting out and grant making has grown so rapidly that the contracting practices of modern governance have gotten far ahead of our capacity to build either explanatory theory or theory that helps government officials organize and manage more effectively (Posner 2002; Sclar 2000). To paraphrase Todd LaPorte: public administration by contracting out and grant making may work in practice but not in theory.

The "state of agents" project has wisely chosen to approach the subject empirically and theoretically. At the center of most considerations of the efficacy of third-party government or the state of agents is the matter of accountability. The first purpose of this article is to unpack the application of concepts of accountability to the state of agents using a "promises of accountability" heuristic (Dubnick 2002b). Second, the promises of accountability logic is applied to earlier research on the measurement of performance in five federal agencies, particularly hollowed-out federal agencies that rely almost entirely on third-party agents to carry out their work (Frederickson and Frederickson 2007). This research focused on federal performance measurement particularly through the implementation of the Government Performance and Results Act of 1993 (GPRA) and the Program Assessment Rating Tool (PART), both administered by the Office of Management and Budget (OMB). Finally, some hollow state accountability generalizations are presented.

THE PROMISES OF ACCOUNTABILITY

Accountability is both a word and a bundle of concepts. As a word, accountability is an icon notable for its synonymic qualities—its capacity to stand in for many other terms, making the term a powerful rhetorical tool (Dubnick 2002b). "Used in the appropriate context, accountability holds the promise of bringing someone to justice, of generating desired performance through control and oversight, of promoting democracy through institutional reforms, and of facilitating ethical behavior" (Dubnick 2002b, 2). As a word, accountability is notoriously ambiguous (Dubnick 2005), chameleon like (Sinclair 1995), and subject to promiscuous use. Nevertheless, a definition of accountability provides a launching point for the promises of accountability schema presented below. Accountability as traditionally defined (1) is a social relationship between at least two parties (2) in which the demand or obligation for account-giving is accepted and expected by both parties. Account-giving is, therefore, "after the fact" of an accountable matter. Accountability in modern terms also (3) includes organizational and/or political mechanisms designed to "bring" or "cause" individuals or agencies to account "before the fact" by causing them to act accountably.

Concepts of accountability are worthy of defense and are critically important to modern understandings of governance. It is the burden of our argument that contemporary concepts of accountability are of particular importance to the governance of third parties and to the state of agents project (Bovens 2005).

Approached conceptually[1], accountability emerges as meaningful in at least six such contexts—contexts described here as "promises." These contexts with their attendant promises are empirical, which is to say they are intended to describe the actual usage of accountability in specific ways in particular settings (Grant and Keohane 2005). These six promises of accountability are based on and flow from earlier work describing "multiple

1 A concept is taken here to mean a prototype, an exemplar, a mental construct, or an agreed-upon meaning among a community of actors. Concepts are most meaningful within broad intellectual contexts—they exist with frames, formats, schemes, webs of belief, or paradigms (Kuhn 1970; Dubnick 2002a).

Table 1
Promises of Accountability as Means and Ends

Focus on: (Time)	Accountability Valued as: Means (Mechanisms)	Ends (Virtues)
Inputs	**1A. The Promise of Control** Assumes that hierarchy, standardized procedures, and orders will result in greater accountability.	**1B. The Promise of Integrity** Assumes that individuals and even groups wish to be accountable or can be part of an accountable culture.
Processes	**2A. The Promise of Ethical Behavior/ Good Choices** Assumes that corruption and inappropriate behavior can be prevented, ameliorated, or corrected through procedural accountability mechanisms.	**2B. The Promise of Democracy** Assumes the creation of vertical and horizontal procedures of accountability will result in democratic outcomes.
Outcomes	**3A. The Promise of Performance** Assumes that individuals or groups held to account for their behavior through performance measurements will perform better.	**3B. The Promise of Justice/Equity** Assumes the opportunity to seek justice in light of some claimed act or possible act will result in justice or fairness.

Source: This table is taken from Dubnick (work in progress).

accountabilities" (legal accountability, political accountability, bureaucratic accountability, and professional accountability) and applications of the multiple accountabilities schema to the Challenger Accident (Romzek and Dubnick 1987) and the Ron Brown air crash (Romzek and Ingraham 2000).

Each of the six promises of accountability represents a distinct context in which the logic of accountability is played out. Furthermore, each of the six promises of accountability is logically associated with the contemporary state of agents, associations that we describe in the second section of this article.

Accountability concepts promise control, ethical behavior, performance, integrity, democracy or legitimacy, and justice or equity. Some of these promises are the *mechanisms* or *means* of accountability—concepts of control, ethical behavior, and performance. Some are the *ends* or *virtues* of accountability, the promises of integrity, democracy or legitimacy, justice or equity. In table 1, the promises are arranged by time along a vertical axis from inputs (the promises of control and integrity); to processes (the promises of ethical behavior and democracy or legitimacy); and to outcomes (the promises of performance and justice or equity). Introducing the elements of time to concepts of accountability is a considerable advance over earlier almost exclusive emphases on backward-looking account-giving, such as the oversight hearings and reports on the Challenger and Columbia accidents. It will be argued here that there is a close resemblance between contextually derived promises of accountability and the equally important concept of governance, particularly as the concept of governance comprehends what the state of agents is understood to describe (Frederickson 2005). As rhetoric, the promises of accountability are almost always exaggerated. Nevertheless, embedded in that rhetoric are useful concepts that help explain accountability mechanisms and ends, particularly in the context of contemporary governance. Indeed,

Dubnick argues that "accountability . . . emerges as a primary characteristic of governance in contexts where there is a sense of agreement about the legitimacy of expectations among community members. Conceptually, accountability can thus be regarded as a form of governance that depends on the dynamic social interactions and mechanisms created within a moral community" (Dubnick 2002b, 7).

In the traditional practice of public administration, it was understood that hierarchy, standardized procedures, policy directives, and the other elements of bureaucratic control result in both individual and group accountability to agency and program goals. Despite the empirical challenges to this logic on the part of those describing organizational goal displacement, agency capture, and other "bureau pathologies," it is almost certainly the case that the dominant form of accountability in governance was and continues to be the "promise of control." We know, of course, the many instances when the promise of accountability through control fails. But we also understand that the order, predictability, and reliability of government serve to indicate that in a general sense the promises of accountability through bureaucratic control are kept.

Accountability as a word and a concept promises ethical behavior of public officials, or at least a reduction in corrupt behavior, ordinarily through a variety of procedural mechanisms such as income reporting requirements for elected and appointed public officials. Professional and jurisdictional codes of ethics and oaths of office are thought to encourage ethical behavior and thus accountability. It could be argued that the almost universal adoption of merit-based civil service systems by American governments at all levels during the Progressive Era was an accountability movement. Accountability in the Progressive Era was understood to be based on a broad norm that merit systems would reduce corruption on the part of public officials and thereby make government more honest and therefore more accountable to the public (Frederickson, Johnson, and Wood 2006). In the last 30 years things have changed. Merit-based selection and promotion of civil servants has, in many jurisdictions, and particularly the federal government, been watered down and new layers of politically appointed officials have been inserted above the merit-based civil service. This has been done based on a promise of accountability, defining accountability to be responsiveness to political principals and particularly elected executives. The erosion of the civil service and the widespread practice of contracting out have almost certainly weakened the promises of accountability through bureaucratic control and accountability through civil service merit procedures, codes of ethics, oaths of office, annual audits, and other ethics-enhancing mechanisms (Aberbach and Rockman 2000).

The promise of accountability through performance measurement is the fashion of the day. In the national government, the GPRA and the George W. Bush administration's PART are designed to control agencies by setting goals and annually measuring progress toward those goals. Such performance measurement procedures are thought to be before-the-fact mechanisms by which agencies are caused to improve performance in the direction of program goals. In recent years, GPRA and PART have been "strengthened" by the OMB through the use of performance targets and performance budgeting, linking agency performance to requests for funding (Mihm 2001). The logic behind the implementation of the federal No Child Left Behind law is another example of the promise of accountability by performance measurement, performance based on high-stakes student testing. Many states and localities have developed similar accountability mechanisms based on performance measurement, notably the currently popular CitiStat movement (Behn 2007).

One of the modern promises of accountability is greater individual and institutional integrity. Accountability, in other words, helps make individuals or groups "better" because they are more accountable. Using similar reasoning, the promises of accountability include furthering the development of cultures of accountability in groups or institutions. The modern literature on cultures of accountability builds on earlier literature on the nature of individual behavior in the workplace (McGregor 1960) and are an accountability adaptation of an earlier question: Are workers naturally inclined to be accountable or must workers be forced (or controlled) to be accountable? Achieving accountability through the virtue of integrity appears to turn on trust (Bertelli and Smith, 2005). Several streams of literature including principal-agent theory, prisoners' dilemma games (Axelrod 1984), relational contracting (Bertelli and Smith, 2005), and self-regulating production and commerce (Prakash and Potoski 2006) also conclude that trust is the key determinant of cooperation, responsiveness, or consumer protection.

Accountability promises democratic government, particularly as a set of procedural guarantees. In theory, elected officials are held accountable each Election Day. The problem is that electoral accountability processes are distinctly rigged in favor of incumbents, considerably weakening the promise of democratic accountability. Other electoral forms, particularly initiative, referendum, and recall, are likewise understood to be means by which the people can directly circumvent representative government and either directly legislate or remove from office. Finally, the US Constitution as well as the 50 state constitutions and particularly the Bill of Rights include firmly established norms to which public officials are held accountable.

The accountability promises of justice tend to be played out in judicial settings. The logic of due process is designed to guarantee just treatment at the bar. The accountability promises of fairness or equity are another matter. Fairness issues tend to be policy specific, as in fair housing, equal schooling, and so forth. In this policy specificity, fairness issues also tend to manifest themselves at the stage of policy implementation and are laden with accountability questions. The splendid recent field research by Maynard-Moody and Musheno (2003) details the vexing accountability choices faced by cops, teachers, and counselors as they make choices between clients in the context of scarce time and resources.

Before turning to an application of some research on federal performance measurement and the promises of accountability, the following point will be helpful. Some of the claims or assertions presented in the next section fit the promises of accountability rather well. These claims or assertions are summarized in table 2. Other claims or assertions made on the basis of this research do not fit the promises of accountability frame but are nevertheless important and relevant to accountability.

THE ACCOUNTABILITY OF FIVE HOLLOWED-OUT FEDERAL AGENCIES

We turn now to some research on federal performance measurement in five federal agencies, all of them primarily hollowed out (Milward 1996, Milward, Provan and Else 1993), and the question of accountability. This work is based primarily on the findings presented in Frederickson and Frederickson's *Measuring the Performance of the Hollow State* (2006), to date the most complete study of the applications of the logic of performance measurement to federal programs. The Fredericksons consider specifically five agencies in the Department of Health and Human Services, selected primarily because they are differently hollowed out. The agencies studied were

Table 2
The Promises of Accountability and some Tentative Claims Regarding the Influence of Performance Measurement

Focus	Means (Mechanisms)	Ends (Virtues)
Inputs	**1A. The Promise of Control** Bureaucratic and hierarchical controls enhance agency and program accountability. Whether a federal program is operated by third-party agents is a key determinant of accountability. Building rules and program policies into grants and contracts with third-party agents enhances accountability. The characteristics of third parties and the nature of their principal-agent relations is a key determinant of accountability	**1B. The Promise of Integrity** Agencies that practice lateral trust-based forms of "relational contracting" foster cultures of accountability.
Processes	**2A. The Promise of Ethical Behavior/Good Choices** Building rules and program policies into grants and contracts with third-party agents enhances accountability. Auditing and reporting requirements and procedures enhance program accountability.	**2B. The Promise of Democracy** There is little evidence that there are democratic outcomes as a result of performance-based accountability protocols. Federal performance measurement regimes are primarily executive branch forms of accountability. Performance measurement represents attempts to superimpose managerial logic and managerial processes on inherently political processes embedded in the separation of powers.
Outcomes	**3A. The Promise of Performance** However flawed, GPRA- and PART-based performance measurement applications are believed by most principals to enhance program accountability, at least marginally. Federal performance measurement regimes are "top down" and "one size fits all." Even if there is goal agreement, it is often difficult to find measures adequate to the determination of social outcomes.	**3B. The Promise of Justice/Equity** Little evidence that performance-based accountability protocols improve the prospects for the justice or equity of program outcomes.

(1) the Food and Drug Administration (FDA) to represent regulatory agencies;

(2) the Centers for Medicare and Medicaid (CMS) because it uses the two key state of agents tools—contracts primarily with insurers in the case of Medicare (MCR) and grants to states and cofunding with states in the case of Medicaid (MCD). CMS was also chosen because of scale to represent hollowed-out agencies with complex networks of third parties and very large budgets;

(3) the Health Resources and Services Administration (HRSA) to represent multipurpose agencies with many small federally funded programs with a wide range of third-party agents implementing their programs and policies;

(4) the Indian Health Services (IHS) because of the unique nature of their agents—sovereign Indian tribes;

(5) the National Institutes of Health (NIH) to represent research and development grants-based third-party agents.

This breadth of policy tools and types of third-party agents suggest unique and particular promises of accountability associated with each agency and government program (O'Toole and Meier 2004). To explicate these particular patterns we consider, from the Frederickson and Frederickson study, several differences found between the agencies associated with variations in: the extent to which they are hollowed out; their accountability to federal program purposes; the nature and quality of their network articulations; the level of their goal and policy agreement; and the level of their centralization of policy implementation (Salamon 1989).

The Fredericksons used a qualitative case study methodology designed to develop empirically informed but pre-hypothetical claims.[2]

Because Frederickson and Frederickson studied only hollowed-out agencies, it is almost certainly the case that the primary determinant of an agency's promises of accountability has been left out—that is the degree to which an agency is hollowed out. If they had,

2 A qualitative, comparative case study research methodology is primarily used in the Frederickson and Frederickson study. Qualitative research techniques were chosen because federal performance measurement is relatively new and because the body of extant research on public sector performance measurement is still relatively limited (Jennings and Haist 2002, 2004).

The particular qualitative approach used in this study relies on grounded theory (Dey 1999; Strauss 1987; Strauss and Corbin 1990). Grounded theory refers to theory that is inductively developed during the course of a study. The term "grounded" suggests that the theory is grounded in the data that are collected and analyzed through systematic processes. Grounded theory represents a departure from deductive or standard social science empirical research in that it does not involve predetermined theories and hypotheses conceptually developed from extant literature and tested against empirical data (Maxwell 1996). Instead of the use of hypothesis-driven sampling schemes based on existing literature and theory to determine appropriate questions, in grounded theory ongoing discussions with interviewees influence the framing and reframing of research questions, research categories, and the nature of data generated during the cycles of data analysis (Miles and Huberman 1994).

The primary purpose of grounded theory is to empirically build theory that provides an understanding of the area studied. The research findings of grounded theory "constitute a theoretical formulation of the reality under investigation, rather than consisting of a set of numbers, or a group of loosely related themes" (Strauss and Corbin 1990, 24).In grounded theory, the aim is to find, develop, discover, or generate empirically based claims, hypotheses, concepts, and theories that account for or explain complex phenomena.

The primary data source for the Frederickson and Frederickson research was field interviews with those charged with the implementation of GPRA, PART, and other performance measurement protocols in the selected agencies. An additional data source was the use of agency performance plans. The final sources of data for this study were telephone interviews and e-mail exchanges.

for example, studied the Social Security Administration, one of the few remaining major examples of direct government, there would be a strong basis for comparison against hollowed-out agencies. The only examples of direct government in the Frederickson and Frederickson study are some clinics and services operated directly by the IHS and some "intramural" research done directly by the NIH.

There is some evidence that the accountability differences between direct and hollowed-out government administration are salient. Using PART scores, the direct government agencies average measure of accountability was 68 on a scale of 0–100, with an average of management effectiveness of 16.8 on a scale of 1–20. Agencies operating primarily through third parties had an average measure of accountability of 56 in the case of block and formula grants and 54 in the case of competitive grants, on a scale of 0–100. These same agencies had an average management effectiveness measure of 15.3 in the case of block or formula grants and 14.5 in the case of competitive grants, on a scale of 0–20 (Frederickson and Frederickson 2007, 182). Agencies using third parties to do research and development, however, had an average measure of accountability of 81 (the highest) on a 0–100 scale and an average management effectiveness of 17.2 on a scale of 0–20 (again, the highest). (It is important to remember that the PART rating model has been the subject of considerable criticism as to its methodology and implementation.[3]) So, using PART scores, direct government agencies and programs were found to be more accountable and better managed than agencies and programs doing their work through third parties, with the exception of research and development agencies and programs using third parties. Nevertheless, it appears that these "data" support the claim that the quality of management and the nature of accountability in hollowed-out third-party-operated federal programs are very different from the management and accountability of directly operated federal programs. The accumulation of GPRA- and PART-based performance measurement results shines a strong light on these differences. One might conclude, based on PART comparisons, that third-party-operated federal programs are less accountable and less well managed than direct government programs; or one might conclude that networks of third parties will be inherently, or by their nature, less accountable to federal government objectives and less well managed, at least less well managed in terms of traditional definitions of hierarchical management. And, of course, one might argue that the PART rating methodology is not sufficiently valid to support any empirical claims. Still, it is interesting that the advocates of contracting out and third-party government are also the advocates of the PART rating program, and the PART rating program shows that hollowed-out or third-party agencies and programs are less well managed, and presumably, therefore, less accountable, than direct service agencies and programs.

The Frederickson and Frederickson findings generally confirm the point made more than 80 years ago by Mary Parker Follett, who wrote that "we see that the place in which orders are given, may make all the difference in the world as to the response we get. Hand them down from President to the work manager and the effect is weakened. One might say that the strength of a favorable response to an order is in reverse ratio to the distances the order travels" (Follett 1926, 157). Follett's hypothesis had to do with giving orders in a standard organizational hierarchy. If one substitutes goals for orders, and if one assumes that Follett's description of a "favorable response" is essentially the same thing as

3 See Frederickson and Frederickson (2007, 183–4); Dull (2006), Gilmour and Lewis (2006); Radin (2006).

accountability to program goals, then the Fredericksons' findings support her hypothesis. On the basis of those findings, and adapting Follett's hypothesis to a modern third-party government, we suggest that: "The strength of agency accountability to program goals is in reverse ratio to the distances actual implementation travels and the layers of third parties through which implementation travels."

Modern federal government performance measurement regimes are deeply embedded in vertical networks of third parties. As Follett claims, the length of these articulations matter and the characteristics of third parties matter. For example, agency performance in the context of federal-state relations and state implementation of federal policy is very different from agency performance in the context of contract-based third-party service providers. This suggests the claim that: "The accountability to agency program goals is in reverse ratio to the distances actual implementation travels and to the characteristics of third-party grantees and contractors."

The two CMS programs, MCR and MCD, are notably different. Because MCD funding is shared with the states and because the states are responsible for the implementation of MCD, in the day-to-day work of MCD there is a considerable weakening of accountability to federal program goals. In addition, states ordinarily contract with third-party providers or financial intermediaries to implement MCD policy at the state and particularly the county level, further attenuating accountability to federal goals. MCD is, therefore, an expression of operating American fiscal federalism. Unlike the fiscal federalism features of MCD, MCR is implanted through a series of very large CMS contracts with insurance companies and other financial intermediaries. The details of federal MCR goals, rules, and policies are built into those contracts and act as "capacitors" (voltage-increasing devices) that mitigate federal goal attenuation.

Like CMS and particularly MCD, the grants between the IHS and Indian tribes are a form of fiscal federalism, because grant recipients have a form of sovereignty, resulting in extensive federal goal attenuation. Clinics and other services operated directly by the IHS experience far less federal goal attenuation.

Of the five agencies studied, the most difficult to judge regarding accountability to federal program goals is the FDA. Although FDA is a regulatory agency, it has few food and medicine inspectors and is operated following the logic and protocols of compliance through self-reporting. Even with fairly strict reporting protocols, food producers and pharmaceutical firms will shade compliance in the direction of their goals and away from federal goals (Heimann 1998).

Easily the most interesting finding with respect to accountability is the case of the NIH, judged to be the most accountable of the five agencies to federal program goals although it is almost entirely hollowed out. As the Frederickson and Frederickson study indicates, the NIH practices a sophisticated version of "relational contracting" involving long-standing relations between NIH and its grantees built on interdependence and trust. This is an excellent example of the "accountability promise of integrity" shown in table 1. Indeed, NIH is an outstanding example of why federal research and development agencies get the highest marks for accountability to federal program purposes and for management in the PART rankings. Although not as sophisticated as NIH in its relations with third parties, the HRSA was found by the Fredericksons to be highly accountable to federal program purposes by using many of the techniques of relational contracting.

The promises of accountability are clearly influenced by the state of third-party agents. In the extent to which accountability promises control, it matters whether the third parties are states, contractors, research grantees, or commercial firms engaged in guided self-regulation. Accountability via hierarchy is the strongest but is closely followed by accountability based on relational contracting or grant management. The promises of procedurally based accountability are entirely based on the procedures and the skills of those implementing the procedures. Following the logic of performance measurement, accountability to agency goals will improve as a result of the implementation of GPRA and PART performance measurement. According to the interview respondents in the Fredericksons' study, the five federal agencies being studied have improved their performance measurement capabilities. Therefore, following the logic of performance measurement, these agencies are more accountable for program outputs and even outcomes. However, following the logic of counterfactuals, there is no evidence that performance measurement actually "causes" improved performance. Nevertheless, agency officials tend to believe it does.

Frederickson and Frederickson describe the federal state of agents as "articulated vertical networks of third parties," articulation taken to mean jointed, as in an articulated bus or the articulation of cars on a train. The quality of articulation has to do with the extent to which separate organizations or institutions in a network are coupled, fit together, linked, or combined and the nature and quality of those connections. Articulation can take many forms, including both block and categorical grants from the federal government to the states and territories or to other subgovernments; contracts let by the federal government or by subgovernments to nonprofit or for-profit third parties; the processes of regulating third parties; and the provision of funding to third parties with clients who have claims for services from these third parties. The point is that the wide variation of arrangements between the federal government and third parties and the remarkable variation in the purposes of these arrangements cannot be understood as just grants or contracts with third parties or subgovernments. The subject of federal grants and contracts is far more complex. Although contracting out and privatization have received the bulk of public attention, contracts are only one form of the devolution of federal program implementation by third parties. Formalized contracts between the federal government and third parties are all, however, various forms of articulations, almost all of them vertical.

In third-party grant and contract regimes, the form of articulation between agencies and their third parties ranges from remote control to managed networks. Remote control, according to Salamon (1989), is a loosely coupled grant or contract articulation in which third parties exercise wide discretion and latitude in both the management and the substance of policy implementation. Contracts for health research between NIH and hospitals and universities are a typical example. Managed networks, at the other extreme, are tightly coupled articulations between agencies and grant or contract third parties. Tight network management can be hierarchical or exercised through agency oversight programs, auditing procedures, carefully drafted contracts, tight contract management, negotiated shared understandings between agencies and third parties, and so forth. The direct services of the IHS are an example of a managed network that is direct government and therefore hierarchical. Through grants and contracts, MCR manages very large-scale health financing schemes that are relatively tightly coupled by means of detailed contracts and regular auditing oversight. MCD, on the other hand, is a relatively loosely coupled block granting regime with long strings of third-party health providers and financial intermediaries. Between

MCR on one hand, and MCD on the other, the HRSA is a comparatively simple granting agency which purposely trades some relaxation of grant details in return for grantee discretion and a willingness to be accountable to agency purposes. HRSA openly fosters a trusting relationship between the agency and its third parties, an example of the accountability "promise of integrity." In an idealized state, the FDA-guided self-regulation model would be an example of trust-based loosely coupled form of network articulation. However, because of competing goals (treated below), the FDA alternates between loosely coupled arrangements with food and drug manufacturers and providers on one hand and attempts to "manage" self-regulation on the other hand (Heimann 1998). Finally, the NIH are a prime example of a trust-based granting regime which operates on remote control with minimal oversight, a good example of "relational contracting," or the accountability promise of integrity.

The nature and quality of agency-agent network articulations that are on remote control exhibit the promises of accountability through performance and integrity. By comparison, principal-agent network articulations that are "managed" illustrate the accountability promises of hierarchical control or bureaucratic contract and grant management processes.

Most federal programs have multiple purposes. The level of goal agreement or congruence in an agency's statutory or regulatory foundations and the level of policy agreement among agency principals influence how they respond to the imperatives of accountability. Tensions between the states and the federal government in the MCD program, and certainly the competing interests in FDA's regulatory sphere, appear often in the media and are frequently discussed in the research literature (Heimann 1998). Based on face-to-face discussions with agency officials, goal disagreement between their principals is the primary accountability challenge FDA and MCD face. But annual GPRA and PART reports evaluating these programs make no mention of competing or ambiguous goals.

In the context of either contradictory and/or ambiguous goals, the fixing of accountability is rendered more difficult and less effective (Heinrich 1999). Nowhere is this more evident than in the application of accountability expectations to FDA. FDA is expected to enable pharmaceutical and food products to reach market as quickly as possible while at the same time assuring the public that those pharmaceutical and food products are safe. Which of these two goals FDA maximizes is a function of many factors—with political forces in the congressional and executive branches of government and the influence of the pharmaceutical industry being the most important.

In the 1990s, the FDA, along with many other regulatory agencies, was part of a distinct shift in approaches toward federal regulation of businesses (Heimann 1998). This shift was away from traditional regulation toward deregulation and toward a logic of partnerships whereby regulators and regulated industries share regulatory roles. For FDA this meant that pharmaceutical and food companies paid fees for regulatory "services" and expected, in return, timely reviews. In addition, regulated industries such as the seafood industry, using agreed-upon standards and reporting protocols, were authorized to engage in self-regulation. But in 2004 and 2005, two drugs approved under these partnership arrangements, Vioxx and Bextra, were found to have undisclosed but known dangerous side effects and were taken off the market. As a result, Congress, FDA, and the administration are openly reconsidering a return to more traditional forms of food and drug regulations. Recently, the peanut-salmonella crisis again highlighted the FDA goal-conflict

issue. The tension between FDA's food and drug safety responsibilities and its food and drug promotion responsibilities is played out as political forces shift from one emphasis to the other and back again. Accountability at FDA is played out in the context of these contradictory and shifting goal emphases. Evidence of the influence of competing FDA goals is seen in the slowness of their GPRA performance measurement development and in the ongoing aggregation and disaggregation of FDA performance measures. The decision to shift between two contradictory emphases is a policy rather than a performance challenge and one which performance measurement is ill-suited to solve. This policy challenge leaves FDA in a position of attempting to manage its way around and through the problem of conflicting goals and leaves the matter of agency accountability unresolved.

MCR and MCD also have competing goals. Unlike FDA's regulation/promotion goals, MCR and MCD have financial service goals achieved through articulated networks of third parties—the states and their third-party contractors in the case of MCD and large-scale health finance contractors in the case of MCR. Clients are at the bottom of these vertical networks, clients entitled to MCR and MCD services from the doctors, pharmacies, hospitals, and nursing homes providing those services. MCR and MCD are expected to finance health care services to those who are eligible and at the same time control health care costs. Goal conflict and goal ambiguity are less evident at NIH, IHS, and HRSA.

Although goal conflict is evident in the operations of the FDA, MCR, and MCD, it appears that open consideration of conflicting goals is unacceptable. Accountability mechanisms such as GPRA and PART performance measurement do not attempt to reconcile these competing objectives. Performance measures and the data needed to make them operative are never described in terms of possible goal conflict. Agency and program goals are described as if they are compatible (Moynihan 2006). Based on GPRA and PART reports, it is as if agency goal conflict and competition is the problem that dare not speak its name. In interviews with agency officials, however, agency goal conflict is a regular theme. There is some evidence, particularly at FDA and CMS, that performance measures are attempts to pave over goal conflicts, as if to claim that really good agency performance will somehow demonstrate that agencies can be equally accountable to conflicting or competing goals. This would not be the first time that better management is thought to be the answer to vexing questions of goal conflict.

In third-party government, policy implementation is, by definition, formally extended beyond the agency's hierarchy. There are variations in several forms of agency centralization and decentralization, most particularly programmatic fragmentation and the geographic spread of policy implementation agents. Levels of centralization and decentralization of third-party policy implementation complicate agency accountability. Programs that use grants or contractors to provide services are decentralized by definition.

There appears to be an association between third-party government and the decentralization and fragmentation of the federal government. Patterns of policy implementation that are centrally funded and locally managed and delivered are now deeply established in contemporary federalism. In the health fields, the states and territories, as well as Indian tribes, constitute decentralized governance third parties.

In addition to geographic decentralization, federal program fragmentation is evident in the health fields; each state and territory operates dozens of federally funded programs. Much of the actual operation of these programs at the street level is contracted to private nonprofit or for-profit third parties by the states and territories. In every state and locality,

federally funded health contractors operate as financial intermediaries or carriers, as health researchers, or as providers of direct health and hospital services. Third-party government, decentralized federalism, and program fragmentation, it could be argued, are nearly the same thing.

How does the correlation between third-party government and decentralization influence accountability and vice versa? The study of GPRA implementation indicates, first, that federal agencies with state, territorial, or tribal third parties will build performance measurement accountability mechanisms into more general patterns of negotiated cooperative federalism or federal-state mutual adjustment. In this context, third parties participate in goal setting and performance measurement decisions and become partners in GPRA implementation. The most volatile examples are MCD and MCR, both exhibiting constantly changing and churning granting or contracting relationships between CMS and the states—a kind of marathon dance of accountability.

Most of the GPRA-related elements of agency-state relations tend to be among top executives at each level and among GPRA implementation specialists. The rank and file in the state agencies that implement federal programs, including the NIH-funded health research laboratories, know little of GPRA or PART and assume that accountability through performance measurement is someone else's responsibility. Research third parties will insulate themselves from formalized accountability mechanisms—both agency goal setting and performance measurement. NIH has entirely walled off their grant recipients from any participation in accountability-based performance measurement and has developed an effective GPRA compliance system in the context of a decentralized and fragmented medical research model.

This suggests that under loosely coupled arrangements GPRA is treated by the agencies as Washington-level bureaucratic requirements rather than important elements of their contract and grant management of third-party policy implementation. What may, in Washington, be thought of as essential features of accountability and good management tends not to be reflected down loosely articulated chains of third-party implementers. Real accountability, following the definition of accountability at the beginning of this article, therefore, may have less to do with formal systems of performance measurement such as GPRA and PART and more to do with grant and contract management and oversight at the operational level and execution and legislative branch politics.

Third, in the name of accountability, nonprofit or for-profit contractors will be obliged to build data and other performance measurement requirements into contracts. In ordinary contract circumstances, third parties will be required to keep certain records and aggregate certain data. Contractors may or may not know the connection between what they are obligated to provide as a part of contracts and how that data might be used in measuring agency performance. Such contract regimes are top down principal-agent models that assume it to be possible to formalize reliable and consistent data requirements and expectations through compliance language (Radin 2006). Despite rules and regulation, the gathering, organization, and analysis of reliable and consistent performance measurement, data have proven to be an illusion.

SOME CONCLUSIONS

Although the Government Performance and Results Act of 1993 was described as a major step forward in agency and program accountability, the act makes no mention of the fact

that the actual work of federal agencies is now mostly done by third parties. In the ensuing 15 years, it has become increasingly clear that of the many factors influencing agency and program accountability, it is federal policy implementation by third parties that is most important (Cooper 2003; Frederickson and Frederickson 2007; Handler 1996; Radin 2006). The failure of lawmakers in crafting GPRA to anticipate the salience of the state of agents explains in part why GPRA and its Bush administration application, PART, have only marginally influenced the accountability of federal agencies and programs.

To frame our conclusions, we return to the promises of accountability schema and put our primary claims into that schema. Table 2 summarizes the application of third-party government to the six promises of accountability.

Despite the fanfare associated with performance measurement in the federal government and the certitude of its adherents, the strongest forms of agency and program accountability continue to be based on promises of control (1A) and promises of ethical behavior resulting from rules and managerial procedures (2A). Independent of other factors, it appears that hierarchy, standardized procedure, policy directives, and the generally understood features of bureaucratic control are associated with good management and accountability to program goals. This being the case, the contemporary dominance of third-party implementation of federal public policy would seem to suggest that such arrangements bring with them serious accountability challenges. This we believe to be the case.

The accountability challenges associated with the state of agents vary widely depending on the characteristics of third parties and on the variations in their tasks. States as third parties, for examples, are very different from large for-profit health finance firms as third parties and therefore their forms and levels of accountability vary. States implementing federal programs, and in some cases cofunding them, are not just accountable to federal program purposes; that accountability competes with accountability to state interests and purposes. This form of third-party accountability is best understood as a form of dynamic negotiated fiscal federalism. Large health finance firms as federal contract third parties are accountable to federal program goals and also to corporate stockholders and the imperatives of commerce. Accountability in such settings is a highly dynamic iterative pattern of give and take, a kind of marathon dance of accountability.

In the state of agents, federal goal attenuation is a fact of life. Federal goal attenuation is greater as third-party policy implementation is geographically further away from the granting or contracting agency. Likewise, attenuation is greater under conditions of agency program fragmentation. Goal attenuation can be mitigated and agency and program accountability can be clarified and enhanced by the techniques of relational contracting as in the case of NIH and by the techniques of network management as in the cases of HRSA and MCR. Several issues of third-party performance are central to the issue of federal goal attenuation. But there is little evidence that the protocols of performance measurement enhance accountability.

Goal incongruence explains the accountability challenges faced by many hollowed-out agencies. The FDA, for example, is in a continuing state of accountability schizophrenia, bouncing back and forth between pressures to bring food and medicine efficiently to market and pressures to protect consumers from dangerous food and medicine. Both MCR and MCD also have competing goals with attendant accountability challenges. Third parties get caught in these challenges. In the cases of MCR and MCD, their third parties are asked to do the "expand services while containing costs" heavy lifting. In the state of

agents, it is often the case that agencies contract out making the tough choices or give grants to third parties to ration services or cut costs. Accountability is thus delegated down vertical chains of third parties.

Even if there is goal agreement, it is often difficult to find measures adequate to the determination of social outcomes. Performance measurement represents attempts to superimpose managerial logic and managerial processes on inherently political processes embedded in the separation of powers. In general terms, federal performance measurement regimes are forms of executive branch accountability. In lofty rhetoric, it is claimed that accountability through performance measurement will enhance the democratic qualities of government. There is little doubt that accountability to the constitution and the laws as well as to overarching regimes value is a means by which the "quality" of American democratic government can be evaluated. But there is little evidence that the protocols of performance measurement are important forces in accountability to democratic values. If anything, performance measurement is almost exclusively an executive branch instrument used to leverage the values and preferences of whoever is president and of his or her party. The language of accountability is a common feature of presidential campaign rhetoric, but in almost all cases it is framed as "accountability to the people." Although Congress passed GPRA in 1993, there does not appear to be a continuing congressional interest in performance measurement even when such measures are included in the president's budget request.

Finally, the sheer scope of just that part of the federal state of agents considered in this article is breathtaking. Not only is the era of big government not over, in the state of agents the era of VERY BIG *governance* is here to stay. In many ways, it is remarkable that such a jumble of principals, agencies, policies, and agents is accountable. Some elements of the state of agents, such as the NIH, HRSA, and even MCR on their best days, are largely fulfilling the promises of accountability and "working in practice." The challenge before us is to make them "work in theory."

REFERENCES

Aberbach, Joel D., and Bert A. Rockman. 2000. *In the web of politics: Three decades of the U.S. federal executive*. Washington, DC: Brookings Institution.

Axelrod, Robert. 1984. *The evolution of cooperation*. New York, NY: Basic Books.

Behn, Robert D. 2007. *What all mayors would like to know about Baltimore's Citistat performance strategy*. Washington, DC: IBM Center for the Business of Government.

Bertelli, Anthony M., and Craig R. Smith. 2005. *Relational contracting and managing networks: New public management and the search for gains from repeated interaction* Paper presented to the biannual meeting of the Public Management Research Association, Los Angeles, CA, October 1.

Bovens, Mark. 2005. Public accountability. In *The Oxford handbook of public management*, eds. E. Ferlie, L. E. Lynn, Jr., and C. Pollitt, 183–208. Oxford, UK: Oxford Univ. Press.

Cooper, Phillip J. 2003. *Governing by contract: Challenges and opportunities for public managers*. Washington, DC: CQ Press.

Dey, Ian. 1999. *Grounding grounded theory: Guidelines for qualitative inquiry*. New York, NY: Academic Press.

Dubnick, Melvin J. 2002a. *Seeking salvation for accountability.* Paper given at the annual meeting of the American Political Science Association, Boston.

———. 2002b. *The rhetoric and promises of accountability.* Remarks prepared for presentation at Leiden University, September 19.

———. 2005. Accountability and the promise of performance: In search of the mechanisms. *Public Performance & Management Review* 28:376–417.

Dull, Mathew. 2006. Why PART? The institutional politics of presidential budget reform. *Journal of Public Administration Research and Theory* 16:183–216.

Follet, Mary Parker. 1926. The giving of orders. Reprint, 1996. In *Classics of organization theory*, eds. J. M. Shafritz and S. J. Ott, 157–62. New York, NY: Harcourt Brace.

Frederickson, H George. 2005. Whatever happened to public administration? Governance, governance everywhere. In *The Oxford handbook of public management*, eds. E. Ferlie, L. E. Lynn, Jr., and C. Pollitt. Oxford, UK: Oxford Univ. Press.

Frederickson, David G., and H. George Frederickson. 2007. *Measuring the performance of the hollow state*. Washington, DC: Georgetown Univ. Press.

Frederickson, H George, Gary A. Johnson, and Curtis H. Wood. 2006. *The adapted city: Institutional dynamics and structural change*. Armonk, NY: M.E. Sharpe.

Gilmour, John B., and David E. Lewis. 2006. Assessing performance budgeting at OMB: The influence of politics, performance, and program size. *Journal of Public Administration Research and Theory* 16:169–87.

Grant, Ruth W., and Robert O. Keohane. 2005. Accountability and abuses of power in world politics. *American Political Science Review* 99 (1): 29–43.

Handler, Joel F. 1996. *Down from bureaucracy: The ambiguity of privatization and empowerment*. Princeton: Princeton University Press.

Heimann, C. F Larry. 1998. *Acceptable risks: Politics, policy, and risk technologies*. Ann Arbor, MI: Univ. of Michigan Press.

Heinrich, Carolyn J. 1999. Do government bureaucrats make effective use of performance management information? *Journal of Public Administration Research and Theory* 9:363–93.

Jennings, Edward T., and Meg Patrick Haist. 2002. *Does performance measurement perform?* Unpublished paper, Research Workshop of the Empirical Study of Governance, Texas A&M University, College Station, TX, February 21–23.

———. 2004. Putting performance measurement in context. In *The art of governance: Analyzing management and administration*, eds. P. W. Ingraham, L. E. Lynn, Jr, 173–94. Washington, DC: Georgetown Univ. Press.

Kuhn, Thomas S. 1996. *The structure of scientific revolutions*. Chicago: University of Chicago Press.

Light, Paul. 1999. *The true size of government*. Washington, DC: Brookings Institution.

Maxwell, Joseph A. 1996. *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage Publications.

Maynard-Moody, Steven, and Michael Musheno. 2003. *Streetwise workers: Enacting identities and making moral judgment*. Ann Arbor, MI: Univ. of Michigan Press.

McGregor, Douglas. 1960. *The human side of enterprise*. New York, NY: McGraw-Hill.

Mihm, Christopher. 2001. Implementing GPRA: Progress and challenges. In *Quicker, better, cheaper? Managing performance in American government*, ed. D. W. Forsethe, 101–11. Albany, NY: Rockefeller Institute Press.

Miles, Matthew B., and Michael A. Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*, 2nd ed. Thousand Oaks, CA: Sage Publications.

Milward, H. Brinton. 1996. Symposium on the hollow state: Capacity, control, and performance in interorganizational settings. *Journal of Public Administration Research and Theory* 6:193–95.

Milward, H. Brinton, Keith Provan, and Barbara Else. 1993. What does the hollow state look like? In *Public management: The state of the art*, ed. B. Bozeman, 309–22. San Francisco, CA: Jossey-Bass.

Moynihan, Donald P. 2006. What do we talk about when we talk about performance? Dialogue theory and performance budgeting. *Journal of Public Administration Research and Theory* 16:151–68.

O'Toole, Laurence J., and Kenneth J. Meier. 2004. Public management in intergovernmental networks: Matching structural networks and managerial networking. *Journal of Public Administration Research and Theory* 14:469–94.

Posner, Paul L. 2002. Accountability challenges of third-party government. In *The tools of government: A guide to the new governance*, ed. L. M. Salamon, 523–51. New York, NY: Oxford Univ. Press.

Prakash, Aseem, and Matthew Potoski. 2006. *The Voluntary Environmentalists: Green Clubs, ISO 14001, and Voluntary Environmental Regulations*. New York, NY: Cambridge Univ. Press.

Radin, Beryl A. 2006. *Challenging the performance movement: Accountability, complexity, and democratic values*. Washington, DC: Georgetown Univ. Press.

Romzek, Barbara S., and Melvin J. Dubnick. 1987. Accountability and the public sector: Lessons from the Challenger tragedy. *Public Administration Review* 47:227–38.

Romzek, Barbara S., and Patricia W. Ingraham. 2000. Cross pressures of accountability: Initiative, command, and failure in the Ron Brown plane crash. *Public Administration Review* 60:240–53.

Salamon, Lester M. 1989. Beyond privatization: The tools of government action. Washington, DC: Urban Institute Press.

Sclar, Elliott D. 2000. *You don't always get what you pay for: The economics of privatization*. Ithaca, NY: Cornell Univ. Press.

Sinclair, Amanda. 1995. The chameleon of accountability: Forms and discourses. *Accounting, Organizations and Society* 20:219–37.

Strauss, Anselm L. 1987. *Qualitative analysis for social scientists*. Cambridge, UK: Cambridge Univ. Press.

Strauss, Anselm L., and Juliet Corbin. 1990. *Basics of qualitative research: Grounded theory procedure and techniques*. Newbury Park, CA: Sage Publications.