

Robert M.

La Follette School of Public Affairs

at the University of Wisconsin-Madison

Working Paper Series

La Follette School Working Paper No. 2014-008

<http://www.lafollette.wisc.edu/publications/workingpapers>

Does Training Matter? Evidence from Performance Management Reforms

Alexander Kroll

Florida International University

akroll@fiu.edu

Donald P. Moynihan

La Follette School of Public Affairs, University of Wisconsin-Madison

dmoynihan@lafollette.wisc.edu

Forthcoming in *Public Administration Review*

November 2014



Robert M. La Follette
School of Public Affairs
UNIVERSITY OF WISCONSIN-MADISON

1225 Observatory Drive, Madison, Wisconsin 53706

608-262-3581 / www.lafollette.wisc.edu

The La Follette School takes no stand on policy issues; opinions expressed
in this paper reflect the views of individual researchers and authors.

Does Training Matter?

Evidence from Performance Management Reforms

Forthcoming in *Public Administration Review*

Alexander Kroll

Florida International University

akroll@fiu.edu

&

Donald P. Moynihan

University of Wisconsin-Madison

dmoynihan@lafollette.wisc.edu

Training is much discussed but rarely studied in public management. Using multiple waves of survey data, we examine the effects of training on the implementation of performance management reforms in the U.S. federal government, asking if those exposed to training are more likely to use performance data and strategic goals when making decisions. Training is positively associated with reform implementation, but there is little evidence that this association can be explained by the development of specific capacities to overcome performance management challenges. Our findings offer two implications for the practice and study of training. We propose that training is likely to succeed if it is designed and funded to close specific capacity gaps needed for successful reform implementation. But we also need to better understand alternative causal mechanisms by which training facilitates reform implementation, such as explaining and justifying reforms.

Training occupies a paradoxical place in management theory and practice. We ask it to do great things even as we are cynical about its capacity to do so. Training is a default solution to all manner of managerial challenges that range from ethical problems and rule violations to employees acquiring skills. It is credited with communicating organizational norms, new ideas, and reinforcing organizational culture. This optimistic account of training assumes that individual preferences and capacities are mutable and subject to external influence.

The pessimistic counter-narrative is that training does little. This view is held by anyone who has counted the minutes during a mandated training session, worried about taking time away from actual work and yearning to check email. Pessimists say training is just a symbolic or well-intended effort to deal with problems too fundamental to be solved by a daylong session with slides and easel boards. Managers may use training to ward off accusations of inattentiveness to undesirable employee behavior, or to punish employees under suspicion of wrong-doing. This kind includes training on racial sensitivity, sexual harassment, or ethics.

There is a middle-ground view, which is that training has high potential value but is offered too little or in a manner that rarely helps employees. U.S. federal employees rank attention to training lower than almost any other aspect of their work experience. One in two employees believe that their training needs had been assessed or were satisfied with the training they received (OPM 2014, 29, 32). As budget cuts squeeze spending, employee frustration grows. Among those who offer training for federal employees that we interviewed for this article, there was a uniform belief that resources available for training have declined as federal spending has come under pressure.¹ One trainer put it like this: “From a department perspective, there are real needs, but training is a want, not a need.”

In this article, we offer evidence on the value of training by examining whether the exposure to training is associated with reform implementation, specifically performance management reforms in the U.S. federal government. The variety of types and purposes of training makes it difficult to discern how it matters. The benefit of focusing on a single purpose of training (reform implementation) in the area of performance management is that it largely controls for differences in types of training. To better establish the validity of our analysis we examine the effects of training across three time periods, 2000, 2007, and 2013. Each time period correlates with a different reform: the Government Performance and Results Act (GRPA) passed in 1993 but not fully implemented until the late 1990s, the Program Assessment Rating Tool (PART) that was in place from 2002 to 2008, and the GPRA Modernization Act passed in 2010.

We use employee utilization of performance data and strategic goals as measures of reform implementation. This approach is consistent with explicit expectations in the language of these reforms that emphasized performance information use and a focus on agency-wide strategic objectives as central goals (GAO 2013a; Moynihan and Lavertu 2012). We examine whether training fosters the use performance data and strategic goals to make decisions by giving employees greater skills to overcome performance management problems, such as difficulties in measuring outcomes. Our quantitative analysis finds support for a direct positive effect of training, though we do not find that training is associated with the specific capacity gaps we identify. Qualitative interviews suggest performance management training spent a good deal of time explaining and justifying new reforms. This finding highlights alternate causal mechanisms for researchers to more systematically investigate. In short, while we find that although training matters, understanding how it matters remains a pressing practical and scholarly question.

This article unfolds as follows. First, we explain the context of our study – performance management reforms in the U.S. federal government and the role of training in these reforms. Second, we propose that training has a positive effect on reform implementation. Third, we examine a particular mechanism by which training may matter: the development of capacity to overcome implementation difficulties. Next, we explain the data (government surveys and additional interviews) and methods we use. Finally, we explain our findings, discuss them, and point out avenues for further research, before identifying implications for the practice of public administration.

Performance Management and Training in the U.S. Federal Government

In the last 20 years, the U.S. federal government experienced three waves of performance management reforms. GPRA required agencies to measure performance and undertake strategic planning. PART was an in-depth assessment of each federal program, ranking it on a scale from ineffective to effective (Moynihan 2013a). The GPRA Modernization Act of 2010 updated GPRA by requiring additional reporting and review processes, and institutionalizing new performance management staff positions and leadership expectations (Moynihan 2013b).

The policymakers who passed the original and updated GPRA identified a role for training. In GPRA, the Office of Personnel Management (OPM) was directed to “develop a strategic planning and performance measurement training component for its management training program and otherwise provide managers with an orientation on the development and use of strategic planning and program performance measurement” (Congress 1993, Sec. 9). The Modernization Act called upon the head of OPM and the Performance Improvement Council (a government-wide panel of performance improvement officers) to identify “the key

skills and competencies needed by Federal Government personnel for developing goals, evaluating programs, and analyzing and using performance information for the purpose of improving Government efficiency and effectiveness” (Congress 2010, Sec. 12a). Further, OPM was asked to work with agencies to incorporate these skills into training. Table 1 summarizes the performance management competencies identified by OPM (2012).

[Insert table 1 here]

The Modernization Act recognized that while OPM might identify key skills, much of the actual training would come via agencies and other parties. Agencies were directed to identify “the operation processes, training, skills and technology, and the human, capital, information, and other resources and strategies required” (Congress 2010, Sec. 3) to achieve performance goals. According to Government Accountability Office (GAO) surveys, in 2000, 62% of federal employees said they had some training on performance management in the last three years, 70% in 2007 and 66% in 2013. But it occurs largely through a patchwork of different providers: the OPM, agencies themselves, and non-profit or private providers. The relative influence of each group is difficult to determine, even as nonprofit and private providers have gained more prominent roles.

Research on Training and Reform Implementation

One way to conceptualize reform implementation is to treat it as a form of organizational change, thereby allowing us to draw insights from organizational development scholarship (e.g., see Gallos 2006). Theories of organizational change suggest that acceptance

of change requires a variety of types of information and arguments, hinting, though rarely explicitly identifying, at the role training might play. Beckhard and Harris's (1977) classic model identifies common challenges to change. They suggest that individuals must experience dissatisfaction with the status quo, observe the desirability and practical applicability of an alternative, and have specific doubts addressed. This model suggests a variety of roles for training to play. Through a mixture of evidence and normative persuasion, training can help to convey problems with the status quo and outline benefits of an alternative. Details on the content of a reform and competencies to overcome specific challenges can illustrate the practicality of the reform and overcome specific concerns about implementation.

Another influential model of organizational change identifies three distinct strategies: empirical–rational, power-coercive, and normative–reeducative (Chin and Benne 1969; Quinn and Sonenshein 2008). The empirical-rational strategy (“telling”) is based on the assumption that organizational members will adopt change if managers can justify the change and demonstrate its benefits. The power-coercive strategy (“forcing”) is about managers exercising coercion and using sanctions to ensure compliance with organizational changes. The normative–reeducative (“reeducating”) strategy primarily focuses on the social dimension of organizations. To adopt change, social norms, values, and habits need to be altered which requires mutual persuasion within collaborative relationships. Of these three change strategies, training could be paired with at least two. It can be used for “telling” and “reeducating,” though for the latter training would have to be about challenging assumptions rather than just learning about new practices.

The following two sections theorize about the effect of training on reform implementation in greater detail. The first section hypothesizes a positive training effect and elaborates on different explanatory mechanisms. We derive two mechanisms directly from

the organizational change literature: conveying information about a change needed and providing normative justification for behavioral change. We identify a third mechanism (the creation of capacity) that is less well-considered in organizational change research, but may facilitate reform implementation by helping to overcome specific objections (Beckhard and Harris 1977). Indeed, Fernandez and Rainey (2006) identify training as one means managers can use to overcome resistance to change. The second section theorizes more about capacity effects and develops hypotheses about specific performance management-related capacity.

The Effect of Training

One explanation of why training can foster reform implementation is that it has an information effect. Training conveys basic information about the reform and its requirements, and simply knowing more about a reform may make employees more likely to implement it. This type of declarative knowledge (information about what) is usually seen as inferior to procedural or tacit knowledge (information about how or which, when, and why) (Aguinis and Kraiger 2009). However, the provision of information on important reform concepts, changes to central processes and routines, and on expectations and rewards seem to be fundamental lower-order conditions for behavioral change. The dissemination of information through training may foster shared mental models, meaning that employees understand their own and their colleagues' role(s), behavior(s), and linkages as parts of a larger system. Though this has little to do with the development of specific competencies, research has shown that a better understanding of the "bigger picture" fosters coordination and successful collective action (Marks et al. 2002).

A second explanation of the training effect is that it may alter the trajectory of reform implementation to the extent it presents a reform as an appropriate norm to follow, thereby

raising employee support for and investment in the reform. This could be based on a legal requirement (“we are doing this because you are legally obliged to do so”), but also on a normative imperative (“this change is the right thing to do” or “all of our peers are doing it”). Institutional isomorphism theory, for example, argues that the potential for new organizational approaches to spread and embed themselves depends greatly upon training in educational or other professional settings. Such settings communicate these ideas as appropriate norms relevant to defining the identity of the group (DiMaggio and Powell 1983, 152). According to Feldman (1989, 399), “training has become the setting where perceptions and expectations about the organization are formed, where norms about social behavior are developed, where corporate values and ideology are communicated.”

Some studies of performance management emphasize that training can be important in changing normative beliefs about the value and appropriateness of the reform. Training can “reduce uncertainty, fear, and cynicism” (Yang and Hsieh 2006, 863) and help employees “understand, accept, and feel comfortable with the innovation” instead of “feeling pressured or overwhelmed” (Cavalluzzo and Ittner 2004, 249). Darnall and Kim (2012) suggest that technical changes designed to improve environmental performance become less likely to succeed in the absence of training that generates shared norms and beliefs.

A third explanation of a positive training effect, which we discuss in greater detail in the next section, is that it creates capacity to implement reforms. That is, the adoption of performance management reforms requires competencies with regard to performance measurement, analysis, and reporting, and training can be crucial for developing such capacity.

H₁: Performance management training improves the implementation of performance management systems.

Capacity Effects of Training

Reforms require employees to face new tasks or ways of thinking. Reform implementation is more likely to occur when training gives employees the skills to meet the new demands, what we refer to as a capacity effect. We borrow our definition of capacity from Huber and McCarty (2004, 481) who understand capacity as “skills and expertise” and “the ability to accomplish intended actions” (for overlapping definitions, see Ingraham, Joyce and Donahue 2003, 15; Ting 2011, 245).

The importance of capacity to governance elevates the value in understanding whether training can affect capacity. In models of governmental and third-party provider capacity, training is explicitly assumed to offer a means of raising capacity (Brown 2012; Ingraham, Joyce and Donahue 2003). Prominent scholars argue that capacity deserves attention equivalent to other concepts that animate formal theories of governance, such as information asymmetry or policy design (Huber and McCarty 2004; Ting 2011). The general treatment of training in public administration makes the assumption that training can generate specific capacities. When we speak of specific capacity we refer to skills and expertise that are related to a specific task, (in our case, performance management), consistent with Ting’s (2011, 248) distinction between generalist and specialist capacity. Examples of specialist capacity are the use of Federal Aviation Administration training to inculcate high reliability characteristics that reduce airplane crashes (O’Neill and Kriz 2013) or facilitating greater environmental performance as part of an environmental management system (Darnall and Kim 2012).

A longstanding tradition of research on training in psychology focuses on the effects on job performance or skills that serve as antecedents of job performance (Aguinis and Kraiger 2009). This research points to adaptive expertise as an important outcome variable

for training, studying whether and when employees show the capacity to match different skills to the appropriate task rather than simply learning a certain competency (Kozlowski et al. 2001). For the case of performance management this research implies that specific capacity is more than a checklist of additive competencies and that it is important to develop the ability to adapt measurement and evaluation skills to different situations and contexts.

Management studies research also assumes that training can be used to develop needed competencies to enable individuals to succeed in their specific organizational contexts (Burke and Hutchins 2007). Empirical research on the impact of training on skill development shows that this effect is stronger for technical training compared to more complex managerial training programs, and more significant for skills that can be segmented into step-by-step routines compared to, for example, soft skills (Hunt and Baruch 2003; Morrow, Jarrett and Rupinski 1997).

In the context of performance management, Yang and Hsieh (2007) suggest that technical training facilitated the adoption and effectiveness of performance systems in Taiwan because of the traditional lack of capacity to analyze performance data. The need for such analytical capacity has not lessened. For example, Behn's (2014) detailed account of performance reviews in the United States describes the central importance of analytical capacity.

Cumulatively then, there is a good deal of support for the idea that training matters by improving the capacity of individual employees. To assess whether training is associated with conveying capacity skills for our study, we identify four well-established performance management implementation challenges that demand capacity: (a) how to measure performance, (b) how to use discretion, (c) how to learn from performance data, and (d) how to use data for accountability purposes. Some of the capacity items we identify clearly map

onto performance management competencies identified by OPM in table 1. Most obviously, accountability skills relate to our measure of accountability context described below, and performance measurement skills relate to performance measurement challenges. In a high-accountability context, training could help to create more effective reports and enable managers to better explain the limitations of data to stakeholders. For measurement problems, training could facilitate the development of more accurate logic models and meaningful surrogates for difficult-to-measure outcomes (Hatry 2006).

For learning processes and discretion, the connections are less direct but relate to a wider array of skills listed in Table 1. Training could amplify the benefits of learning routines if regular reviews of performance trends are informed by performance analysis, problem solving, and oral communication skills (Behn 2014). Training can also better leverage discretion by conferring capacity to diagnose and redesign processes, or by facilitating the appropriate use of targets and incentives to motivate employees.

If training generates a capacity effect, we expect it to moderate the relationships between the variables we identified above and reform implementation. That is, those who are trained will be better able to use performance information and strategic goals in the face of measurement difficulties. We also hypothesize that more training will enable managers to better leverage the discretion granted to them, benefit more from learning routines and performance feedback, and better respond to contexts shaped by high accountability requirements.

H₂: Performance management training will **weaken** the negative effect of measurement problems for the implementation of performance management systems.

H₃: Performance management training will **strengthen** the positive effect of discretion for the implementation of performance management systems.

H4: Performance management training will **strengthen** the positive effect of a learning routine for the implementation of performance management systems.

H5: Performance management training will **strengthen** the positive effect of accountability for the implementation of performance management systems.

Data and Methods

Sample and Measures

This article uses data from three surveys conducted by the GAO in 2000, 2007, and 2013. The surveys were addressed to a random, nationwide sample of mid- and upper-level federal managers in the agencies covered by the Chief Financial Officers Act of 1990, stratified by agency and management level. The GAO used similar sampling frames across years and drew a new random sample of managers for each of the waves.² The overall response rate for each of the surveys was about 70% (for more information, see GAO 2008, 2013a). Since we do not hypothesize that the effect of training differs across reform initiatives, we do not run separate models for the years 2000, 2007, and 2013. Instead, we pool the data to increase the external validity of the study by examining whether the effects of training hold across different points in time and reform contexts. Other advantages of a pooled model are that it makes efficient use of the data structure since all three surveys employed identical questions and it estimates reliable standard errors because all t-values are based on a constant sample size.

While our primary goal with this article is to empirically test the hypotheses laid out above, we employed some qualitative methods to gain a contextual understanding of the implementation of training for performance management. We interviewed eight individuals who oversaw training in the GPRA, PART, and current era, including officials who worked

in agencies, for the OPM, Office of Management and Budget (OMB), and private and non-profit trainers. Because the population of such individuals is not easily identifiable, they were selected using a snowball sampling approach. They were asked questions on the nature, extensiveness, and content of performance management training at different time periods. We also examined actual training materials such as presentations and instructional manuals.

In our quantitative analysis, we measure reform implementation using factor scores of the variables “use of strategic goals” and “use of performance information” when setting program priorities; allocating resources; adopting new program approaches or changing work processes; and developing or refining program performance measures. Both proxies measure two important elements of performance-management-reform implementation: defining strategic priorities and collecting data on results and achievements as well as using both in managerial decision-making (Moynihan 2013a, 2013b). The variables are moderately positively correlated ($r=0.51$), but there still is a good portion of unshared variation worth studying. Indicators of the convergent (see Cronbach’s Alpha) and discriminant validity (see factor loadings) of both dependent variables as well as the exact measures of all variables and their scale reliability coefficients are reported in the appendix.

Training was measured as an additive index of six planning- and measurement-related activities, ranging from zero (no training) to six. We also included nine control variables, most of which were identified as relevant in a recent review of empirical studies on the implementation of performance management systems (Kroll 2014). The measures of these variables are widely established, and other studies that used survey data collected by the GAO utilized these measures (see, e.g., Cavalluzzo and Ittner 2004; Dull 2009; Moynihan and Lavertu 2012). More descriptive information on all variables and their bivariate correlation coefficients – which do not exceed a value of 0.55 – can be found in table 2.

[Table 2 here]

Modeling and Limitations

To model the different effects of training, we use the following strategy. Hypothesis 1 focuses on a general training effect. Here, we simply examine the direct effect of training on the two dependent variables, while controlling for ten variables (see table 3, models 1 and 2). The expectation is that more exposure to training will make it more likely that managers will pay attention to strategic goals and performance data when making decisions.

Hypotheses 2-5 are modeled as interaction effects because our assumption is that the relationships between the independent variables measurement problems, discretion, learning routine, and accountability, and the dependent variables performance information use and goal use are contingent on different levels of employee training. To test each interaction effect between all four variables and the training index, we regress each dependent variable on all of the eleven variables listed in table 3 and add each interaction term one at a time (see columns 3 through 10). This way, we can examine whether adding the interactions improves the model fit significantly or not. The interaction terms were generated as the multiplied scores of the mean-centered individual variables and the interpretation of their unstandardized regression coefficients is as follows: For every one-unit increase in the moderator variable (training), the slope relating the predictor to the outcome variable increases or decreases by the reported coefficient, while holding all other variables constant (Whisman and McClelland 2005).

Since our dependent variables are underlying factors rather than ordinal scales, we employ ordinary least squares regression analysis to estimate the direct and interactive effects of training on reform implementation. To account for the possibility that responses from the same agency might be related to one another, we cluster standard errors at the agency level. The usual shortcomings of cross-sectional survey research should be kept in mind when generalizing from our study. One is that our findings represent statistical associations rather than causal effects. Another is that the possibility that those who had higher interest and support for performance systems were more likely to select into training cannot be fully ruled out since our data do not come from a repeatedly surveyed panel of employees. Our interviews suggest, however, that there is not a clear and direct selection pattern that skews the results in a certain direction.³

As noted above, our test of specific capacity effects is not comprehensive. We examined four specific interactions (we also tested a fifth interaction with “missing link to action” which generated equivalent findings of the other interactions). It is plausible that further capacity measures worth examining exist but are not in our data set and some capacity effect occurs in the direct measure of training. That said, our efforts to measure specific capacity effects are more comprehensive than prior work, and the findings are consistent. We use ten statistical models to examine five interactions (see table 3) as well as another ten models featuring only the 2007 and 2013 samples that included data from additional eight agencies only identified in these waves (results not reported). Of these twenty models, only two found support for a significant capacity interaction between training and discretion, whereas eighteen showed null effects.⁴

Common source bias might generate artificial correlations between items, especially when items are highly perceptual and may have a strong social desirability bias. However, Meier and O’ Toole (2013) find that measures of performance information use seem less

susceptible to common source bias than measures with a higher social desirability bias, such as estimations of organizational performance. In addition, our primary independent variable, whether one participated in training, is a relatively straightforward query about an objective event that lacks a strong social desirability bias, making it less prone to common source bias. This follows Meier and O' Toole's (2013, 447) recommendation to focus on independent variables that are "more specific and less likely to generate spurious results."⁵

Results

Table 3 portrays the results of the statistical analysis, providing information on the direct effects of training (first two columns) as well as on all four interaction effects (following columns). For the direct effects, table 3 reports standardized regressions coefficients, mainly to provide a better overview of the relative magnitude of the training effects. For the moderated effects, unstandardized coefficients are shown which makes it possible to interpret them in a straightforward way, as we explain in the method section.

Given the number of variables in table 3, we will not examine each in detail, but instead, center our discussion on: (a) the direct effects of training; (b) findings related to the use of strategic goals, since this variable is less well-examined than forms of performance information use; and (c) the interaction effects between training and management challenges. For the other variables (missing link to action, leadership commitment, political conflict, senior-level position), it is sufficient to note that our findings are consistent with prior research (e.g., Dull 2009; Moynihan and Lavertu 2012).

[Table 3 here]

Across all models we find that training has a significant positive direct effect on the use of performance information and strategic goals. The magnitude of this effect seems to be slightly larger in the “PI use” compared to the “goals use” model, and it is of moderate strength in comparison to the slopes of other variables. A one standard-deviation increase in training can be associated with a 0.10 and a 0.08 standard-deviation increase in the use of performance data and strategic goals. Managers who experienced more training are also more likely to report paying attention to performance information and to strategic goals for decision-making, a finding that supports hypothesis 1.

Much of the recent literature on performance management has focused on the use of performance information, giving little attention to the role of strategic goals. But performance management systems generate strategic goals just as they do other types of performance information, such as measures or targets. There is, therefore, value in examining patterns associated with the use of strategic goals and comparing them to the patterns associated with performance information use. The results presented here suggest that these patterns are broadly similar – variables that predict performance information use also tend to predict the use of strategic goals.

One concern is that the strong parallels between the use of strategic goals and performance information might be driven by an inability of respondents to discern between the two. However, closer examination of two independent variables that do not produce equivalent findings (measurement problems, senior executive service) undercuts this concern. It is intuitively obvious why the variable measurement problems limits the use of performance data in decision-making, but there is little ex-ante reason to assume such problems would affect the use of strategic goals. It is also plausible that attention to broad

strategic goals might be greater for more senior managers relative to attention to performance data. The results follow these expected patterns – difficulties in measurement undermine the use of performance information but do not matter to the use of strategic goals, and more senior managers devote more attention to the use of strategic goals.

With respect to the interaction effects we theorized about in hypotheses 2, 4, and 5, little evidence supports the possibility that training provided capacity skills to overcome some of the challenges of performance management. The interactions of training with measurement problems, learning routines, and accountability turn out to be insignificant. Managers with more training are not better able to handle measurement problems and do not get more out of learning routines and accountability mechanisms.

The one exception is that training seems to help managers to use their discretion to facilitate performance information use (hypothesis 3). We can see that a one-unit increase on the training scale leads to a significant increase in the effect of discretion on performance information use and goals use (0.02 and 0.01, respectively). Put differently, the slope relating discretion to data and goals use will be 0.12 and 0.6 factor-score points steeper for managers who have experienced the maximum amount of training (25% of the respondents) when compared to those who did not receive any training at all (34% of the respondents). This pattern implies that the use of data and goals will be further stimulated when training complements discretion.

Our qualitative interview data also support the finding that training had a positive effect on reform implementation but did not develop performance management-related capacity. The officials we interviewed were skeptical that capacity-gaps could be filled in relatively short training sessions. They reported that training allotted more time to explaining and justifying reforms than to developing performance management-specific capacity.

Employees, even those who cared about performance metrics for their programs, were generally unfamiliar with the content of reforms and the broader federal performance framework. Although one might expect that familiarity with the GPRA Modernization Act would be greater since it built upon prior reforms, one trainer said: “Managers are not really familiar with Modernization Act – surprisingly so. They have heard of it, but actual application and use of it are pretty low.” This observation is in line with descriptive findings in our data: Though 84% of employees said they had heard of GPRA in 2000, only 64% had heard of the Modernization Act in 2013 and only 58% knew about PART in 2007.

Training materials also tended to convey normative claims about the virtues of performance management, building on the intuitive appeal of the idea of performance. PART-era training centered on conveying information about PART and on justifying it as a legal requirement rather than sharing normative arguments as to its benefits. “With PART, it was more about just education,” one trainer said. “We did not run it like the workshop for GPRA where we examined the details and challenges of the programs that individuals were managing...We mostly just made people aware of PART and what was being asked, so when OMB came to you, you were able to answer questions.” Agency-level training on PART was similar, with a stronger emphasis on how to answer the PART questions rather than use it as a tool to improve performance.

The most prominent reason for the limited attention to capacity gaps was that building competencies through training took time and resources that were not available. Enabling employees to develop better measures or make better use of performance feedback is more difficult to realize than broadly directing attention to changing requirements and new expectations. “To many people training confirmed that it [performance management] was a compliance exercise,” one interviewee said. “Training was not much more than ‘here is what GPRA requires.’” This perception was partly due to the difficulty of dealing with task-

specific challenges in contexts where trainees came from different backgrounds. A trainer at one non-profit company noted: “One limitation of training is time and the cross-agency nature of the group, which limits specific discussion of problems in individual agencies.”

Discussion

We studied performance management reforms in the U.S. federal government and find that training is positively associated with the implementation of these reforms. Managers who were exposed to more training reported higher levels of use of performance information and strategic goals when making decisions. Such a positive effect aligns with several theoretical explanations, but our interviews pointed out that most likely “information” and “justification” mechanisms were at play. That is, training mainly disseminated information about how performance management works and what is expected of managers, and it tried to explain why these efforts are necessary and why they can help to achieve better outcomes.

There are two possible interpretations of these findings. The positive one is that training can have a significant effect on reform implementation, even if it fails to develop reform-specific capacity. Explaining and justifying reforms, which are lower-order conditions for change, may improve implementation, suggesting that any training is better than no training. The more negative reading is that training could have done much more, if it had aimed at creating performance management-specific capacity. However, to better compare the effects of lower- and higher-order change conditions, and to disentangle the effects of information, justification, and capacity, we clearly need further research. The most robust design for this purpose would require a randomized quasi-experiment with three treatment groups: An organization offers three training programs, each of which only focuses on information, normative justification, or capacity. Managers are randomly selected into

programs, and reform implementation can be compared before and after the training programs as well as among groups, and compared to a control group.

Our results also did not provide strong evidence that training created performance management-related capacity to help managers better deal with known challenges. The interviews suggest that this deficit was due to the limited time and resources devoted to performance management training and the diversity of training groups, which limited discussion to generic issues common to multiple agencies. However, to better determine the extent to which resources affect training outcomes, or whether different training purposes require different levels of funding, we need further research.

One exception is that we found modest evidence that training helped managers to get more out of the discretion granted to them. While we are cautious about overstating this finding, since adding this interaction term only marginally improves model fit, it points toward an interesting implication for our understanding of the relationship between discretion and performance information use. Empirical studies suggest that performance management has often not met expectations because increasing output control was not complemented by the provision of greater operational discretion (Moynihan 2008; Nielsen 2014). Indeed, our results provide further support for the direct role of discretion in facilitating the use of performance data. But even if managers have discretion, they may still lack a strong sense of how to make use of performance information. Our findings offer some evidence that training may help managers to better understand how to utilize their discretion to use performance information and strategic goals.

Our research employed an indirect measure of capacity using interaction effects – we considered capacity to be created if managers who experienced more training were better able to handle known performance management-related challenges than those with less training.

Another way to learn more about capacity would be to measure it more directly based on surveys via supervisor ratings, or to observe it as a part of an experimental study before and after managers' participation in training programs. These methods would better examine the direct effect of training on capacity and sort out issues of causality with which correlational research often struggles.

For readers interested in performance management issues, this article is the first of which we are aware that compares the use of strategic goals and the use of performance information. Apart from some logical exceptions, there are strong parallels between the two behaviors. One plausible reason for these parallels is that a generation of federal managers has been told, since the implementation of GPRA, to consider strategic goals and performance information as connected parts of the same overall system.

Conclusion

This article has studied the role of training in reform implementation, using the three most recent performance management reform initiatives in the U.S. government as a case study. The findings contribute evidence on how training relates to public sector reform implementation. Despite the potential importance of training identified by scholars and practitioners, actual evidence about training is scarce in public sector settings in general, and on impacts on reforms in particular. We find that training is positively associated with reform implementation. Managers who received more training on performance management also paid more attention to performance data and strategic goals when making decisions. However, training did not lead to the development of specific capacities needed to cope with measurement or management challenges. We did not find differences in the effects of measurement problems, learning, and accountability on reform implementation when

comparing more and less trained managers – a significant interaction between training and discretion was the only exception.

Our findings fall between the negative and positive narratives of training noted in our introduction. Training directed attention to reform implementation, but it failed to equip managers to deal with implementation problems. Based on the case we studied, we can draw the following conclusions for the practice of public administration more broadly. First, training matters, but an understanding how training matters remains elusive. As organizations undertake large-scale change, their managers should outline what they want training to do, and how it will reach its goals. What is it that training is supposed to accomplish: inform employees, establish a new culture, develop reform-specific capacity, or a combination of all three? Based on the focus, training programs can be designed and have to be evaluated differently.

Second, higher-order needs, such as creating performance management-related capacity, require more resources than other training objectives. Though our quantitative models did not offer evidence on the effect of resources on training outcomes, our interviewees pointed out that not having enough funds and time was one of the major drawbacks for developing specific capacities. Interviewees also saw resources for training declining since the first training programs for GPRA were initiated. The Obama administration has acknowledged the issue, proposed more spending on training, and directed agencies to target their training needs (OMB 2014).

Third, the potential of training depends not just on adequate resources, but also on addressing specific capacity gaps. The types of performance management challenges we examine – dealing with measurement problems, making the most of performance feedback, and responding to a context with a strong accountability emphasis – are well-established in

prior research, but appear to be mostly unaddressed by training efforts. This disjunction between what research identifies as implementation challenges and the actual content of training intended to facilitate implementation suggests an opportunity for better connecting scholarship and practice. Research can identify implementation challenges. Training can address those challenges by filling specific capacity gaps. This article offers an initial effort to conceptualize training so it facilitates such dialogue between research and practice and offers guidance for training in the implementation of performance management reforms.

Endnotes

1 We turned to interviews in part because documentation on basics such as overall spending on training and how funds are divided among training providers is not available. Agencies have incentives to hide spending on training that would be vulnerable to cuts if made transparent, making it impossible to identify spending on training across government over time, let alone training on performance management.

2 The surveyed 24 executive branch agencies represent about 97% of the executive branch full-time workforce, excluding the Postal Service. All three randomly drawn samples are representative of the Chief Financial Officers Act agencies. Since the GAO lumped some of the agencies together into an “other” category when disseminating the 2000 data, we can only use 16 agency identifiers for our pooled analysis. An additional analysis features only the 2007 and 2013 samples and thus includes data from additional eight agencies only identified in these two waves.

3 Organizational units must pay for training for each individual, which means employees lack unfettered discretion to choose training. When units pay for training specifically for their employees, supervisors tend to encourage widespread participation to generate value from the training, which reduces the potential for systematic selection. In some cases, more ambitious employees might press for training opportunities; in others, participation in training can result from the identification of weaknesses or needs in supervisor-employee discussions. Overall, there is not a clear direction to potential selection effects.

4 There are certainly different ways how to model and test capacity effects. One alternative would be to use direct measures of capacity or human capital based on survey questions and then model the capacity effect as an indirect effect – more training leads to more capacity that fosters reform implementation. However, our data set does not provide such direct measures. Furthermore, using additional perceptual measures can increase common-method-bias issues, demonstrating why indirect measures of otherwise perceptual variables, which are based on interactions, can be superior to direct ones (Kristof 1996).

5 We also ran Harman’s single-factor test. Multiple factors emerged and the first factor only accounts for about 28% of the overall variation in all items; we count this as evidence against common-method bias.

References

- Aguinis, Herman, and Kurt Kraiger. 2009. Benefits of Training and Development for Individuals and Teams, Organizations, and Society. *Annual Review of Psychology* 60: 451–474.
- Behn, Robert. 2014. *The PerformanceStat Potential: A Leadership Strategy for Producing Results*. Washington D.C.: Brookings Institution.
- Beckhard, Richard, and Reuben Harris. 1977. *Organizational Transitions: Managing Complex Change*. Reading, MA: Addison-Wesley.
- Brown, Mitchell. Enhancing and Measuring Organizational Capacity: Assessing the Results of the U.S. Department of Justice Rural Pilot Program Evaluation. *Public Administration Review* 72(40): 506–515.
- Burke, Lisa, and Holly Hutchins. 2007. Training Transfer: An Integrative Literature Review. *Human Resource Development Review* 6(3): 263–296.
- Cavalluzzo, Ken, and Christopher Ittner. 2004. Implementing Performance Measurement Innovations: Evidence from Government. *Accounting, Organizations and Society* 29: 243–267.
- Chin, Robert, and Kenneth Benne. 1969. General Strategies for Effecting Changes in Human Systems. In *The Planning of Change*, edited by Robert Chin, Kenneth Benne, Warren Bennis, 32–59. New York: Holt, Rinehart & Winston.
- Darnall, Nicole, and Yongsung Kim. 2012. Which Types of Environmental Management Systems are Related to Greater Environmental Improvements? *Public Administration Review* 72(3): 351–365.
- DiMaggio, Paul, and Walter Powell. 1983. The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review* 48: 147–160.
- Dull, Matthew. 2009. Results-Model Reform Leadership: Questions of Credible Commitment. *Journal of Public Administration Research & Theory* 19(2): 255–284.
- Feldman, Martha. 1989. *Order without design: Information Production and Policy Making*. Stanford, CA: Stanford University Press.
- Fernandez, Sergio, and Hal Rainey. 2006. Managing Successful Organizational Change in the Public Sector. *Public Administration Review* 66(2): 168–176.
- Gallos, Joan. (Ed.). 2006. *Organizational Development*. San Francisco, CA: John Wiley & Sons.
- Hatry, Harry. 2006. *Performance Measurement: Getting Results*. Washington, D.C: Urban Institute Press.
- Huber, John, and Nolan McCarty. 2004. Bureaucratic Capacity, Delegation, and Political Reform. *American Political Science Review* 98: 481–494.

- Hunt, John, and Yehuda Baruch. 2003. Developing Top Managers: The Impact of Interpersonal Skills Training. *Journal of Management Development* 22(8): 729–752.
- Ingraham, Patricia, Joyce, Philip, and Amy Kneedler Donahue. 2003. *Government Performance: Why Management Matters*. Baltimore: The Johns Hopkins University Press.
- Kristof, Amy. 1996. Person–Organization Fit: An Integrative Review of its Conceptualizations, Measurements, and Implications. *Personnel Psychology* 49(1): 1–49.
- Kroll, Alexander. 2014. Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research. Forthcoming in *Public Performance & Management Review*.
- Kozlowski, Steve, Stanley Gully, Kenneth Brown, Eduardo Salas, Eleanor Smith, and Earl Nason. 2001. Effects of Training Goals and Goal Orientation Traits on Multidimensional Training Outcomes and Performance Adaptability. *Organizational Behavior and Human Decision Processes* 85(1): 1–31.
- Lavertu, Stephane, David Lewis, and Donald Moynihan. 2013. Government Reform, Ideology, and Bureaucratic Effort: Performance Management in the Bush Era. *Public Administration Review* 73(6): 845–856.
- Marks, Michelle, Mark Sabella, Shawn Burke, and Stephen Zaccaro. 2002. The Impact of Cross-Training on Team Effectiveness. *Journal of Applied Psychology* 87(1): 3–13.
- Meier, Kenneth, and Laurence O’ Toole. 2013. Subjective Organizational Performance and Measurement Error: Common Source Bias and Spurious Relationships. *Journal of Public Administration Research and Theory* 23(2): 429–456.
- Morrow, Charley, Quintin Jarrett, and Melvin Rupinski. 1997. An Investigation of the Effect and Economic Utility of Corporate-Wide Training. *Personnel Psychology* 50(1): 91–117.
- Moynihan, Donald. 2008. *The Dynamics of Performance Management: Constructing Information and Reform*. Washington, D.C: Georgetown Univ. Press.
- Moynihan, Donald. 2013a. Advancing the Empirical Study of Performance Management: What we Learned from the Program Assessment Rating Tool. *American Review of Public Administration* 43(5): 497–515.
- Moynihan, Donald. 2013b. *The New Federal Performance System: Implementing the GPRA Modernization Act*. IBM Business of Government Research Endowment.
- Moynihan, Donald, and Stephane Lavertu. 2012. Does Involvement in Performance Management Routines Encourage Performance Information Use? Evaluating GPRA and PART. *Public Administration Review* 72(4): 592–602.
- Nielsen, Poul. 2014. Performance Management, Managerial Authority, and Public Service Performance. *Journal of Public Administration Research and Theory*, 24 (2): 431–458.

- Quinn, Robert, and Scott Sonenshein. 2008. Four General Strategies for Changing Human Systems. In *Handbook of Organization Development*, edited by Thomas Cummings, 69–78. Los Angeles: Sage.
- Ting, Michael. 2011. Organizational Capacity. *The Journal of Law, Economics, & Organization*, 27: 245–271.
- U.S. Congress. 1993. Government Performance and Results Act. One Hundred Third Congress, 1st Session. Washington, D.C.: GPO.
- U.S. Congress. 2010. GPRA Modernization Act. One Hundred Eleventh Congress, 2nd Session. Washington, D.C.: GPO.
- U.S. Government Accountability Office [GAO]. 2008. *Government Performance: Lessons Learned for the Next Administration on Using Performance Information to Improve Results*. Washington, DC: U.S. Government Printing Office. GAO-08-1026T.
- U.S. Government Accountability Office. 2013a. *Managing for Results: Executive Branch Should More Fully Implement the GPRA Modernization Act to Address Pressing Governance Challenges*. Washington, DC: U.S. Government Printing Office. GAO-13-518.
- U.S. Government Accountability Office. 2013b. *Agencies Have Elevated Performance Management Leadership Roles, but Additional Training is Needed*. Washington, DC: U.S. Government Printing Office. GAO-13-356.
- U.S. Office of Management and Budget [OMB]. 2014. The President’s Budget for Fiscal Year 2015: Analytical Perspectives. <http://www.whitehouse.gov/omb/budget>
- U.S. Office of Personnel Management [OPM]. 2012. *Memorandum for Chief Human Capital Officers: Government Performance and Results Act Modernization Act of 2010 Functional Competencies*.
<https://www.chcoc.gov/transmittals/TransmittalDetails.aspx?TransmittalID=4517>
- U.S. Office of Personnel Management. 2014. Federal Employee Viewpoint Survey Results. http://www.fedview.opm.gov/2013FILES/2013_Governmentwide_Management_Report.pdf
- Whisman, Mark, and Gary McClelland. 2005. Designing, Testing, and Interpreting Interactions and Moderator Effects in Family Research. *Journal of Family Psychology* 19(1): 111–120.
- Yang, Kaifeng, and Jun Yi Hsieh. 2006. Managerial Effectiveness of Government Performance Measurement: Testing a Middle-Range Model. *Public Administration Review* 67(5): 861–879.

Table 1: Performance Management
Competencies identified by OPM

accountability
attention to detail
customer service
influencing/negotiating
information management
oral communication
organizational awareness
organizational performance analysis
partnering
performance measurement
planning and evaluating
problem solving
reasoning
technical competence
written communications

Table 2: Descriptive Information and Correlation Matrix

Variable	Range	Mean	SD	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1. Performance Information Use	-2.84 - 1.63	0.00	1.00	1.00									
2. Strategic Goal Use	-3.75 - 1.20	0.00	1.00	0.51	1.00								
3. Training	0-6	2.70	2.46	0.28	0.21	1.00							
4. Measurement Problems	1-5	2.91	0.95	-0.28	-0.11	-0.18	1.00						
5. Discretion	1-5	3.15	1.09	0.34	0.29	0.27	-0.22	1.00					
6. Learning Routine	1-5	3.44	1.18	0.36	0.27	0.29	-0.20	0.37	1.00				
7. Accountability	1-5	3.86	0.96	0.34	0.27	0.19	-0.16	0.30	0.37	1.00			
8. Missing Link To Action	1-5	2.48	1.13	-0.30	-0.17	-0.18	0.55	-0.21	-0.21	-0.17	1.00		
9. Leadership Commitment	1-5	3.75	1.07	0.39	0.34	0.28	-0.22	0.46	0.43	0.43	0.48	1.00	
10. Political Conflict	1-5	2.76	1.22	-0.14	-0.08	-0.14	0.55	-0.25	-0.16	-0.14	-0.16	0.37	1.00
11. Senior Executive Service	0-1	0.19	0.39	0.11	0.17	0.08	0.04	0.21	0.09	0.09	0.14	-0.05	0.17

Note: All information is based on the pooled data from 2000, 2007, and 2013.

Table 3: Regression Models

	Direct Effects		Moderated Effects							
	PI Use	Goal Use	PI Use	Goal Use	PI Use	Goal Use	PI Use	Goal Use	PI Use	Goal Use
Training	0.10** (8.27)	0.08** (4.35)	0.04** (0.01)	0.03** (0.01)	0.04** (0.01)	0.03** (0.01)	0.04** (0.01)	0.03** (0.01)	0.04** (0.01)	0.03** (0.01)
Measurement Problems	-0.14** (-6.24)	0.03 (1.15)	-0.15** (0.02)	0.03 (0.02)	-0.15** (0.02)	0.03 (0.02)	-0.15** (0.02)	0.03 (0.02)	-0.15** (0.02)	0.03 (0.02)
Discretion	0.12** (5.73)	0.11** (5.18)	0.11** (0.02)	0.10** (0.02)	0.11** (0.02)	0.10** (0.02)	0.11** (0.02)	0.10** (0.02)	0.11** (0.02)	0.10** (0.02)
Learning Routine	0.13** (8.67)	0.09** (4.94)	0.11** (0.01)	0.08** (0.02)	0.11** (0.01)	0.08** (0.02)	0.11** (0.01)	0.08** (0.02)	0.11** (0.01)	0.08** (0.02)
Accountability	0.14** (6.24)	0.12** (4.86)	0.15** (0.02)	0.12** (0.02)	0.15** (0.02)	0.12** (0.02)	0.15** (0.02)	0.12** (0.03)	0.15** (0.02)	0.13** (0.03)
Missing Link To Action	-0.14** (-8.28)	-0.09** (-5.79)	-0.13** (0.02)	-0.08** (0.01)	-0.13** (0.02)	-0.08** (0.01)	-0.13** (0.02)	-0.08** (0.01)	-0.13** (0.02)	-0.08** (0.01)
Leadership Commitment	0.16** (7.17)	0.17** (7.02)	0.15** (0.02)	0.16** (0.02)	0.16** (0.02)	0.16** (0.02)	0.15** (0.02)	0.16** (0.02)	0.15** (0.02)	0.16** (0.02)
Political Conflict	0.09** (4.36)	0.06** (3.04)	0.08** (0.02)	0.05** (0.02)	0.08** (0.02)	0.05** (0.02)	0.08** (0.02)	0.05** (0.02)	0.08** (0.02)	0.05** (0.02)
Senior Executive Service	0.03 (1.82)	0.09** (5.45)	0.07 (0.04)	0.21** (0.04)	0.06 (0.04)	0.21** (0.04)	0.07 (0.04)	0.21** (0.04)	0.07 (0.04)	0.21** (0.04)
Year 2000	0.01 (0.50)	0.01 (0.14)	0.02 (0.05)	0.01 (0.04)	-0.01 (0.05)	0.01 (0.04)	0.03 (0.05)	0.01 (0.04)	0.03 (0.05)	0.01 (0.04)
Year 2013	0.01 (0.69)	0.04* (2.39)	0.02 (0.02)	0.08* (0.04)	-0.02 (0.05)	0.08* (0.03)	0.02 (0.02)	0.09* (0.04)	0.02 (0.03)	0.09* (0.04)
Training X Measurement Problems			0.01 (0.01)	-0.01 (0.01)						
Training X Discretion					0.02** (0.01)	0.01* (0.00)				
Training X Learning Routine							0.01 (0.01)	0.01 (0.01)		
Training X Accountability									0.01 (0.01)	0.01 (0.01)
Observations	3774	3774	3774	3774	3774	3774	3774	3774	3774	3774
Adj. R ²	0.303	0.186	0.303	0.187	0.305	0.187	0.303	0.187	0.304	0.187

* p<0.05, ** p<0.01 (two-tailed tests). Standardized coefficients and t statistics are reported for the direct-effect models, whereas unstandardized coefficients and standard errors, which are adjusted for 16 agency clusters, are reported for the moderated-effects models. This analysis is based on the pooled data from 2000, 2007, and 2013.

Appendix: Measures

Variable	Operationalization		
Performance Information Use ($\alpha = 0.90$)	For those program(s)/operation(s)/project(s) that you are involved with, to what extent, if at all, do you use the information obtained from performance measurement when participating in the following activities?	Goal Use	PI Use
	Setting program priorities	-0.01	0.90
	Allocating resources	0.04	0.86
	Adopting new program approaches or changing work processes	0.02	0.87
	Developing or refining program performance measures	0.03	0.81
Strategic Goal Use ($\alpha = 0.91$)	For those program(s)/operation(s)/project(s) that you are involved with, to what extent, if at all, do you consider your agency's strategic goals when participating in the following activities?		
	Setting program priorities	0.90	-0.01
	Allocating resources	0.89	0.01
	Adopting new program approaches or changing work processes	0.86	0.03
	Developing or refining program performance measures	0.81	0.08
Training ($\alpha = 0.91$)	During the past 3 years, has your agency provided, arranged, or paid for training that would help you to accomplish the following tasks?		
	Conduct strategic planning		
	Set program performance goals		
	Develop program performance measures		
	Assess the quality of performance data		
	Use program performance information to make decisions		
	Link the performance of program(s)/operation(s)/project(s) to the achievement of agency strategic goals		
	(0 = No; 1 = Yes)		
Measurement Problems ($\alpha = 0.85$)	Based on your experience with the program(s)/operation(s)/project(s) that you are involved with, to what extent, if at all, have the following factors <u>hindered</u> measuring performance or using the performance information?		
	Difficulty determining meaningful measures		
	Different parties are using different definitions to measure performance		
	Difficulty obtaining valid or reliable data		
	Difficulty obtaining data in time to be useful		

Discretion	Difficulty distinguishing between the results produced by the program and results caused by other factors Agency managers/supervisors at my level have the decision making authority they need to help the agency accomplish its strategic goals.
Learning Routine	The individual I report to periodically reviews with me the results or outcomes of the program(s)/operation(s)/project(s) that I am responsible for.
Accountability	Agency managers/supervisors at my level are held accountable for the results of the program(s)/operation(s)/project(s).
Missing Link To Action	Difficulty determining how to use performance information to improve the program
Leadership Commitment	My agency's top leadership demonstrates a strong commitment to achieving results.
Political Conflict	Difficulty resolving conflicting interests of stakeholders, either internal or external.
Senior Executive Service	What is your current grade level? (0 = Others; 1 = Senior Executive Service or equivalent)

Unless otherwise stated, agreement with the items is measured using the following Likert scale: 1 = To no extent; 5 = To a very great extent.

Author Biographies

Alexander Kroll is Assistant Professor of Public Administration at Florida International University. His research interest is in studying organizational effectiveness, employee behavior, and particularly the roles of performance information, strategy, leadership, and motivation.

Donald P. Moynihan is professor in the La Follette School of Public Affairs, University of Wisconsin–Madison. He is fellow of the National Academy of Public Administration, author of *The Dynamics of Performance Management: Constructing Information and Reform* (Georgetown University Press, 2008), and winner of the ASPA/NASPAA Distinguished Research Award.