AARHUS UNIVERSITY

Coversheet

This is the accepted manuscript (post-print version) of the article.

Contentwise, the post-print version is identical to the final published version, but there may be differences in typography and layout.

How to cite this publication

Please cite the final published version:

Holm, J. M. (2018). Successful Problem Solvers? Managerial Performance Information Use to Improve Low Organizational Performance. *Journal of Public Administration Research and Theory*, 28(3), 303-320. <https://doi.org/10.1093/jopart/muy017>

Publication metadata

Title:	Successful Problem Solvers? Managerial Performance Information Use to Improve Low Organizational Performance
Author(s):	Jakob Majlund Holm
Journal:	<i>Journal of Public Administration Research and Theory</i> , 28(3), 303-320
DOI/Link:	https://doi.org/10.1093/jopart/muy017
Document version:	Accepted manuscript (post-print)

General Rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Successful Problem Solvers?

**Managerial Performance Information Use to
Improve Low Performance**

ABSTRACT

Performance management is increasingly the norm for public organizations. While there is an emergent literature on performance information use, we still know little on how managers engage in functional performance management practices. At the same time, growing evidence suggests that managers face pressure to improve low performance because of a negativity bias in the political environment. However, in managerial performance information use, the negativity bias might be reconsidered as a prioritization heuristic with positive performance attributes, directing attention to organizational goals with a favorable return of investment. I test this argument with data from public schools. A fixed-effect estimation is used to analyze how principals prioritize when they are provided with performance information on a number of different educational goals. Furthermore, a difference-in-differences model tests whether the prioritizations of certain goals have performance-enhancing effects over time. The analysis shows that principals prioritize goals with low performance and that prioritizations result in performance increase. The improvements primarily occur for goals that have a low performance level and that are repeatedly prioritized.

INTRODUCTION

During the last decades, one of the most widespread trends in the governing of public organizations has been the introduction of performance management. The core of performance management systems is performance information, which should inform managers' decisions and thereby improve organizational performance (Andersen and Moynihan 2016; Heinrich 1999; Nielsen 2014a; Kroll 2015a). However, evidence of success is scarce. This is partly because managers show reluctance to integrate performance information in their routines and decisions (Kroll 2013; Melkers and Willoughby 2005; Moynihan and Lavertu 2012) and partly due to the inherent difficulty in identifying examples of purposeful use. Taken together, these circumstances leave us without much empirical and theoretical understanding of functional performance management practices (for an exception see Kelman and Friedman 2009).

So, how could managers use performance information to create organizational improvements? While the literature on performance management points to informed decision-making as the key to this process (e.g. Moynihan 2008; Taylor 2009; Vakkuri 2010), it offers little guidance on how data could be transformed into meaningful messages that guide organizational change. This study addresses this shortcoming by presenting two ways managers could utilize information: Either by reacting to signals of failure and initiating problem solving (Cyert and March 1963; Meier, Favero and Zhu 2015), or by exploiting successes to develop expertise and specialize (March 1981).

To provide insights into which strategy managers pursue, this study examines how performance on different organizational goals influences managers' willingness to prioritize certain goals. While previous research has provided important insights into the ways to increase performance information use (Moynihan et al. 2011b; Yang and Hsieh 2007), most studies rely on self-reports from managers to evaluate their use. By testing the influence of

information on a tangible decision outcome, this study offers one of the first accounts of performance information use in actual decision-making.

A key argument in the paper is that prioritizations do not only concern rationales for improvements, signaling effects to the political environment also play a part (Bourdeaux and Chikoto 2008; Lavertu, Lewis and Moynihan 2013). The study extends previous work in this area by testing the idea that the strong emphasis on failures (i.e., a negativity bias) in the political environment (Hood 2011; Nielsen and Moynihan 2017; Olsen 2017) encourages managers to prioritize goals with low performance.

The distinction between internal and external rationales is not only important for understanding a manager's choice of priorities; it also has implications for the outcome of this decision. Priorities that are chosen solely for their signaling value to political principals might end up being nothing more than symbolic (Feldman and March 1981). However, even when the decision initiates functional organizational change, success is no guarantee. I argue that prioritizations influence outcomes with a diminishing return, meaning that they primarily lead to performance increases for the most problematic goals. In this way, managers' tendency to focus on performance failures may actually have positive attributes, directing scarce attention to goals with the highest potential for improvement.

Other than being one of the first studies to address these questions, the paper contributes to the literature on performance management in three general ways. First, I provide an empirical example of how managers engage in functional performance management practices, namely by using performance information to solve problems. The context for this finding is Danish public schools where principals prioritize the educational goals (for instance grades, truancy, wellbeing, parental satisfaction, and student overweight) where their school is performing the worst. After a goal have been made a priority for several years, the performance level increases. While it is not possible to attribute these

improvements directly to the use of information, the analysis shows that performance primarily increases for goals with a low initial performance level, thereby indicating that the underlying mechanism is managerial problem solving. On this basis, the study provides an important perspective to the many studies that tell a pessimistic story on performance management in the public sector (e.g. Andersen 2008; Gerrish 2016; Heinrich 2010).

In addition to the empirical account, the paper contributes theoretically to the debate on how managers could use performance information to make decisions that improve organizational performance (Askim, Johnsen, and Christophersen 2008; Kroll 2015a; Moynihan 2009). More specifically, I develop a comprehensive theoretical framework that theorizes on the rationales for utilizing performance information in strategic decision-making but that also considers the pitfalls when these decisions unfold as attempts to improve performance.

Finally, a valid test of the research questions requires addressing the methodological challenges in the literature. To do this, I operationalize managers' use of performance information in a novel way. Since the principals make an explicit prioritization, it is possible to couple their decisions with the performance level of each goal. This way of structuring data avoids the problem of common source bias (Meier and O'Toole 2013) and allows the use of a fixed-effect model and a difference-in-differences setup that can address the problem of omitted variable bias more exhaustingly than previous research.

PURPOSEFUL PERFORMANCE INFORMATION USE

The theory consists of the following sections. The first section outlines the conceptual framework for understanding performance information use in strategic decision-making. In the next section, I present the internal and external managerial considerations for prioritizing goals with low or high performance levels. The following sections concern the effects of goal

prioritizations on performance. Here, I consider situations with symbolic prioritizations, as well as instances where prioritizations initiate functional organization change. Expectations for the latter situation are developed by theorizing on the production process, including the possibility of diminishing returns, the importance of sustained attention, and performance trade-offs between goals.

Strategic performance management

This paper focuses on managerial use of performance information in strategic decision-making, which broadly concerns a manager's choice of strategies and actions to ensure goal achievement (Poister 2010). To theorize on this process, two shortcomings in the literature on performance management need consideration. The first relates to goal multiplicity. Public organizations often pursue multiple goals, and performance indicators at least partly reflect this multiplicity (Moynihan et al. 2011a). However, studies of performance information use rarely consider this point, neither theoretically nor empirically. Performance is either treated as an abstract concept (in studies based on self-reports), or only one dimension of performance is examined (for example students' final grades). Second, research on performance management needs a more precise conceptual understanding of what "use" entails. Previous research relies on self-reports from managers to solve this issue, thereby leaving what constitute use to be a subjective matter. This study takes a more independent approach and conceptualizes use as the process in which performance information - consciously or unconsciously - influences managers and thereby contributes to a specific thought, a decision or action (Rich 1997: 15).

In relation to these points, performance information use in strategic decision-making concerns the relationship between performance levels across organizational goals and a manager's willingness to make certain goals a priority. Simply put, this means that managers

can use the information to direct attention to goals with a (relatively) low performance level or goals with a (relatively) high performance level.

However, there is also a possibility that managers do not let their prioritizations influence by the results at all. Such a tendency would reflect the many studies that find managerial reluctance to inform decisions and practices with performance information (Kroll, 2015b; Moynihan and Lavertu 2012; Moynihan et al. 2011b; Taylor 2009). Even though the studies do not examine use in relation to strategic prioritizations, the findings suggest that managers' decisions in performance management systems often reflect other factors than the content of the information.

Rationales for using performance information to prioritize goals

In the situation where managers actually consult performance information when deciding on organizational goals, where do they turn their attention? An important distinction in answering this question is between internal and external considerations (O'Toole and Meier 2011). The internal dimension concerns management in relation to processes that take place within the organization with the objective of creating performance improvements (Meier, Favero and Zhu 2015). While external management could also concern improvements, this aspect of public management has a more intermediate objective, namely to maintain autonomy and support for the organization (Carpenter 2001).

Internal considerations: Problem-solving and specialization

An obvious starting point for the internal consideration is the research tradition that originates from the behavioral model developed in the work by Cyert and March (1963). A key idea in this model is that managers improve performance through three sequential phases, namely 1) identification of problems, 2) search for solutions, and 3) implementation of the most

appropriate solution (Cyert and March 1963: 34). The identification phase starts with a stimulus that indicates a problem and a need for managerial action (Mintzberg, Raisinghani, and Théorêt 1976). Performance information is a key component in this process, quantifying the results of activities and thereby easing the identification of problematic areas in the organization's work.

The main theoretical developments and empirical tests of the problem-solving model has taken place in the private sector (for a review see Shinkle 2012). However, recently, a number of studies have used the framework to explain public managers' response to performance data. For example, Salge (2011) finds that performance shortfalls lead public organizations to increase search for innovative solutions, and Nielsen (2014b) and Rutherford and Meier (2015) show that managers redirect attention to various organizational outcomes in response to performance declines and negative social feedback. While these studies have provided vital first steps in adapting the behavioral model to the public sector, their conceptualization of performance shortfalls do not capture the complexity in modern public management completely. The underlying assumption in the studies is that managers evaluate performance shortfalls and successes either along a single performance dimension or in serial processing between multiple dimensions, comparing absolute results to an aspiration level. Thus, the performance hierarchy between goals is not taken into consideration, which leaves out an important aspect of how managers form opinions about what constitutes low and high performance.

On a theoretical level, the idea of a problem-solving manager has also been subject to discussion. On the one side is the line of research that draws on the psychological literature on self-affirmation and points to self-enhancement as a motive when managers process performance information (Jordan and Audia 2012). Empirical support for this idea is found in the studies that show that managers adapt various strategies to cope with performance shortfalls, for instance by choosing social and historical reference points that portrait

performance in a positive light (Audia et al. 2015) or shifting focus to favorable performance indicators (Audia and Brion 2007). The last finding is particularly interesting as it connects to another research tradition that focuses on how managers create improvements from successes (e.g. Baum and Dahlin 2007; KC, Staats and Gino 2013).

A key point in this literature is that an organization could build competence and gain improvement through exploitation and utilization of tacit knowledge about functional activities (March 1981). Knowledge on favorable production factors could be employed in two different ways to further improvements. One way is to dismantle information on the functional activities throughout the organization in an attempt to create best-practice learning (Askim, Johnsen, and Christophersen 2008). Another way is to strengthen the work on the successful goal dimension. Here, the knowledge would enter as competence building in an effort to specialize activities, for example by reorganizing resources to enhance the practices that have proven functional.

External considerations: Buffering and reputation building

Managerial reactions to performance information do not take place in an environmental vacuum; rather, they are a vital part of hierarchical accountability relations. This point is evident in the many studies that show that an engaging political environment significantly increases the likelihood of managers responding to the information (Bourdeaux and Chikoto 2008; Lavertu, Lewis and Moynihan 2013). While the findings suggest that managers pay attention to their environment in performance information use, there is little theoretical basis for understanding how political considerations influence the decisions made in the process.

One important thing to note about the decision to prioritize certain organizational goals is the signaling effect to the environment (Spence 2002). This signaling effect occurs is particularly relevant when priorities have a clear connection to a performance result (Poister

2010). For goals with low performance, a prioritization signalizes responsiveness to organizational problems; for goals with high performance, a prioritization signalizes competence building. Both signals could be used strategically by managers in their quest for political support and autonomy.

Prioritization of goals with low performance is a managerial strategy to buffer the organization from political interventions (Meier and O'Toole 2008). Politicians have a strong electorate incentive to avoid blame for poor performance results (Krause 2003), which could lead them to disclaim responsibility and instead attribute control over outcomes to managers (Nielsen and Moynihan 2017). In addition, low performance could offset a political reaction, for instance in terms reduced autonomy (MacDonald and Franko 2007) and funding (Gilmour and Lewis 2006). Managers can use their prioritizations to avoid such a reaction by signaling to the political environment that the problem is recognized and action is being taken.

Prioritization of goals with high performance is a strategy to improve the organization's reputation. Since a reputation describes the stakeholders' perceptions of the unique capacities, roles and obligations of the organization, it is a valuable political asset (Carpenter 2010). For this reason, managers spend time creating, maintaining and protecting their organization's reputation (Maor, Gilad and Bloom 2013). One way to achieve this objective is to highlight the organization's contribution to the public good by appearing successful and distinct from other organizations (Busuick and Lodge 2017). A prioritization of a goal with high performance can serve to fulfill this purpose because the signal should draw political attention to well-functioning areas of the organizations work, thereby highlighting the organization's impact on society.

How do managers weight the different considerations?

The previous sections have identified both internal and external rationales for how managers could integrate performance information in their prioritizations of organizational goals. So, how should we expect managers to weight these different considerations? One aspect that could factor into this decision is accountability relations. In this regard, an important feature of political environment is a negativity bias, which creates an environment where politicians and citizens evaluate performance information with a strong emphasis on failures (Hood 2011; Nielsen and Moynihan 2017; Olsen 2017). Because of these tendencies, managers might perceive the problem solving and buffering strategies as particularly beneficial to them. This expectation is tested in the first hypothesis.

H₁ *Public managers will react to performance information on multiple goals by prioritizing goals with the lowest performance level.*

Effects of goal prioritization

Symbolic decisions

The distinction between external and internal rationales for prioritizing certain organizational goals is particularly important when considering the potential performance effects of the decision. The reason is that prioritizations based solely on external considerations risk being symbolic use of performance information (Field and March 1981). Symbolic use occurs in a situation where managers decide in a way that reflects the content of the information, but they do nothing more than making the decision publicly available. Such perverse engagement in performance management systems has been found in a number of studies, which show that public servants perceive the systems as irrelevant for their work (Lavertu, Lewis, and Moynihan 2013; Soss et al. 2011) and as a structural obstacle to game (Bevan and Hood

2006). If this type of behavior also characterizes managers' goal prioritizations, we should not expect any organizational change and therefore no performance improvements. This scenario is tested in hypothesis 2.

H₂ *Public managers' prioritization of goals with low performance is symbolic and does not lead to performance improvements.*

Performance potentials and diminishing returns

While the hypothesis above expects managers to engage in performance management practices with a minimum of effort, a prioritization could also entail a real effort to create performance improvements. In this situation, a prioritization initiates a process that adds resources to the particular goal dimension. The resource allocation should be understood broadly, encompassing both material resources (for instance new equipment, better facilities, or supporting staff) and activation of mental resources (for instance motivation and abilities).

No matter the specific content of the prioritization, the manifestation should follow the internal management rationales for making the prioritization. This means that prioritizations of goals with low performance start a process where the organization uses resources to analyze the reasons for the shortcomings and correct the problems (Cyert and March 1963). For goals with high performance, resources add to a process where successes are scrutinized to create knowledge that build competences and commence specialization efforts. One thing to note about both processes is the risk that the investment does not yield performance payoff. Research on both financial (Hanushek 2003; Andersen and Mortensen 2010) and human resources (e.g. Banerjee et al. 2012) have consistently shown that the mere addition of resources to achieve a goal does not equal success. Instead, there are good reasons to believe the marginal return of resources should be taken into consideration.

To understand why, we need to consult the production function that describes how resources transform into outcomes on a goal dimension. This relationship is often characterized using the Cobb-Douglas production function. Here, resources have a diminishing return, meaning that addition of resources to the production increases outputs at a diminishing rate. The reason is that the benefits of the capital stock are being exploited as production increases, thereby lowering the product of each additional worker. Because of this exploration, the marginal return of resources is highest when few resources are used in the production and the corresponding output level is low.

While resources' diminishing return typically is used to explain how companies can maximize their production of material outputs, the idea has also been applied to describe production processes in the public sector, concerning service delivery and the creation of humane outcomes (Heckman 2000). The theoretical argument for making this application is that production processes in public organizations also risk exploration of the capital stock, thereby resembling those in the private sector (Pritchett and Filmer 1999: 223–4). Substantially, the argument provides an explanation for why resources do not always yield a performance payoff. The reason is that the marginal product trends towards zero, or even becomes negative, at high production levels because the many resources cause inefficiency.

While the argument for resources' diminishing return focuses on the input side of a production, the equation could also be turned around, implying that output or outcome levels could be used to assess a production's potential for improvement. Following this idea, goals with low performance have a great potential because the production is at a low level and accordingly, the marginal product of adding resources should be high. This expectation is tested in the following hypothesis.

H₃ *The prioritization of a goal will improve performance, primarily for goals with an initial low level of performance.*

The importance of sustained attention

A standard way to define performance management is as an ongoing system that is bound by performance information creation, decision-making, and outcomes (Moynihan 2008). A system approach, by definition, includes a temporal component, as repeated measurements produces new information over time (van Helden, Johnsen, and Vakkuri 2012). However, this temporal aspect is a rather unexplored aspect of performance management, especially in relation to managerial decision-making.

The cyclic aspect of performance systems means that an initial decision (in the first cycle) can be revisited in later cycles (Pollitt 2013). In the situation where managers prioritize resources, they face a choice in the second cycle of either reprioritizing the same goals as in the first cycle or choosing to prioritize a new set of goals. To understand the implications of this decision, it is necessary to open the black box, which has been labeled the “production process” so far.

Here, a relevant point of departure is the classical input-output model (Boyne 2002). In this model, organizational activities are the key mechanism that transforms resources into performance (Boyne 2003). A prioritization could aim at improving existing activities or developing new activities. An example is the principal that prioritizes the goal “student health.” At the school, a certain amount of resources is already used on student health, such as the provision of gym classes and green areas with playgrounds. When the goal is made a priority, the principal can increase the number of gym classes or develop a new initiative. Neither solutions quick fixes because they cannot be implemented and provide a return within a day, week, or month. In this way, the example highlights two general traits of

production processes in public organizations, namely that (1) they consist of a long chain of complex sub-processes and (2) there could be a substantial time lag before addition of input to a goal actually has an influence on the outcome (Ostrom et al. 1978).

When these considerations are related to the cyclic aspect of performance management, it becomes evident why managers are facing an important decision in later cycles of the system. In a situation where a goal is only prioritized in one cycle, it is unlikely that the prioritization will lead to improvements, simply because the input to production processes lack the sufficient time to transform into improvements. If the manager chooses to prioritize a completely new set of goals, there is a risk that the initial prioritization will only have superficial attention without any real improvement as a result. This argument is tested in hypothesis 4.

H₄ *The longer a goal is prioritized, the higher the probability that the prioritization will lead to improved performance.*

Performance trade-offs

An important aspect of hypothesis 2.1 and 2.2 is what happens to the performance level for non-prioritized goals. Part of the answer to this question depends on the level of available resources in the organization. In one scenario, the organization finds itself in a situation where resources are used for activities that do not benefit the users (i.e., slack) (Migué and Bélange, 1974). While such a situation could be the result of a deliberate managerial choice, improper use of resources could also originate at lower levels of the organization. For this reason, a problem-solving approach often starts with a managerial search for available resources that can be used to identify and implement solutions (Salge 2011). If this search ends successfully, goals with low performance could be prioritized with no consequence for activities and performance levels on non-prioritized goals.

In the other scenario, all resources are used to produce outcomes on the various goal dimensions, meaning that the allocation of resources to one goal reduces the amount of inputs available to produce other goals. Performance trade-offs are likely to occur in such a situation. However, they do not necessarily cancel each other out. A production-possibility frontier illustrates why by plotting the possible outcome combinations in the production of two goals. The frontier connects the goals in a concave relationship, reflecting the increasing opportunity cost of moving resources from one goal to the other. Substantially, the relationship reflects the fact that the most general resources are relocated at first, while specialized and less effective resources are being relocated in the movement along the frontier.

FIGURE 1 AROUND HERE

Figure 1 shows a frontier for two goals, one with a high performance level (y_2) and one with a low performance level (y_1) in the starting point A. When resources are allocated to the goal with low performance, the organization moves to scenario B. $\Delta y_{2A} - y_{2B}$ expresses the performance loss from this reallocation, and $\Delta y_{1A} - y_{1B}$ expresses the performance gain. As seen by the figure, the concave relationship leads the allocation of resources to y_1 to exceed the performance loss for y_2 .

In sum, the two scenarios lead to an identical expectation for how prioritizing resources to goals with low performance will influence the organization's performance level overall, namely that the reallocation will end in a surplus. This expectation is tested in hypothesis 5.

H₅ *The overall performance changes from prioritizing goals with low performance will end in a surplus for the organization.*

RESEARCH DESIGN AND DATA

This section presents the main elements of the research design, namely the empirical setting, data, operationalization, identification strategy, and model specifications.

Setting and data: Performance management in Danish public schools

During the last 20 years, a number of reforms have sought to change the Danish public schools in line with the ideas of new public management (Nielsen 2014a). In the education system, the municipalities oversee the quality of the public schools and accordingly, the implementation of the initiatives were organized here. This flexibility created distinctive performance management regimes across municipalities, varying considerably in their expression and scope. In this study, the data comes from schools in a municipality with one of the most comprehensive systems.

The system consists of two-year cycles. A cycle begins with the measurement of school performance on a large number of goal dimensions. Then, the schools receive a quality report that summarizes their achievements. After the release of the quality report, the principal meets with the stakeholders of the school (representatives of the parents, municipality, and teachers) to discuss the results. The meeting is organized as a learning forum (Moynihan 2005) in which each participant can express his or her interpretation of the results. At the end of the meeting, the principal decides on a number of priorities for the school. These priorities are areas of the school's work that should be given special attention until the beginning of the next cycle. A school has from 12 to 18 months to work on the priorities before the measurement of performance is repeated. Figure 2 illustrates the cyclic nature of the system.

FIGURE 2 ABOUT HERE

The system has three attractive features that makes it suitable for testing the hypotheses. First, the large number of performance indicators creates a situation where principals receive multiple information signals on how their school performs. Second, the principals' decisions has a manifest expression because the priorities are written down and made publicly available. These features make it possible to connect information and decisions, thereby enabling a test of how managers react to differences in performance levels. Finally, and as shown by table 1 below, the schools in a municipality are - on average - quite similar to the rest of Denmark on a number of parameters. These parameters are school size, budget per student, the share of students in special classes, teaching time, student/teacher ratio, and competence-coverage. Thus, any functional aspect of performance management in this setting is more likely to be caused by the system than the characteristics of the schools.

TABLE 1 AROUND HERE.

The data covers three cycles that begin in 2009, 2011, and 2013, respectively. The unit of analysis is goals with available performance information. Since each school is measured on a number of goals, the data has a multilevel structure with goals nested within a school. Over the years, the system has increased the number of performance indicators (and goals) from 33 in 2009 to 40 in 2013.¹ There are 51–48 schools in the municipality in the timespan from 2009 to 2013, and therefore, the number of observations varies from 1500 to 1900.

¹ See table A1 in the appendix for a complete list of goals and descriptive statistics.

Operationalization

A goal's relative performance

In line with the theoretical expectations, when managers face variation in performance information across different organizational goals, it should lead to decisions that give priority to the goals with the lowest level of performance. Accordingly, a key feature in the empirical test of this theoretical argument is a variable that contains information on how well a school is performing across the various educational goals.

This standardized measure of relative performance is created in the following way. First, the maximum level of performance for the goal is determined. This maximum is either 12 on the grading scale or 0/100 on a percentage scale. For example, for the goal of “parental satisfaction”, the maximum is 100 percent, while it is 0 percent for the goal of “student smoking.” With the maximum level of performance in place, the following formula calculates the *relative performance* of a goal:

$$p_{ist} = \frac{pab_{ist}}{pmax_i} * 100$$

where p_{ist} is the measure of relative performance on goal i for school s in time t . pab_{ist} and $pmax_i$ are measures of the school's absolute level of performance and the maximum possible level of performance on the goal, respectively. Thus, higher values on the variable indicate better performance.

An alternative way to create the variable for relative performance is to use the empirical maximum value in the sample instead of the theoretical maximum. This variable has been used as a robustness check to the results from model 4 in table A2. Results from the alternative specification are almost identical (the coefficient is -0.0066 and the t-statistic is -6.48).

Goal prioritization

The variable *prioritization* is a dichotomous variable that contains information on whether a principal (for a given school at a given time) has chosen to prioritize a goal or not. The variable was created by matching a school's priorities- those selected by the principal after meeting with stakeholders - with the performance indicators. A goal was coded 1 if there was a match to one (or more) of the school's priorities and the value 0 if none of the priorities related to the goal.

Since priorities are chosen in an open process, there are no specific demands on the number of priorities chosen, how specific they should be, or how closely they should relate to metrics. This way of organizing the process has two implications for the operationalization. First, some priorities do not match metrics. Instead, these priorities focus on themes such as technology, community collaboration, or the school's economy. Second, there is an element of interpretation in the matching of priorities with the relevant goals, and therefore, there was a need for creating coding rules that ensure that the matching occurred in a valid and reliable manner.

Some priorities were easily connected to a goal. For example "Improve reading in pre-preparatory classes" matches the goal "Reading test: third grade." Other priorities were not directly related to one specific goal but, rather, to a theme (such as learning, parental collaboration, student well-being, or health) with many underlying goals. Finally, some priorities were indirectly connected to either one or several goals. In the coding of these priorities, it was necessary to include additional information that could help uncover the meaning of the priority. For the majority of schools, this information was available in the school's strategy plan and in the following quality report in which the school describes and reflects on the work with the priorities. Since it was rather consistent what principals meant

by their indirect priorities, this meaning was used to code schools with indirect priorities but with no information available.

Lastly, three points should be mentioned. First, even though the aim of the coding strategy was to connect priorities with intended goals, there will be instances where the match does not reflect the principal's intention. Since there is no reason to expect mismatches to be systematic, the problem is a matter of noise in the measurement of the variable, which may have the disadvantage that it makes it more difficult to detect relationships in the estimations. Second, in relation to the validity of the variable, the coding has been reliability tested by having a research assistant code 10% of the goals. A comparison of the two codings gives a Krippendorffs alpha of 0.83, which is considered reliable. Third is the question of what kind of organizational changes a prioritization entails. Empirically, principals add resources (through their choice of strategy) to a goal in many different ways. For example, a strategy for improving parental satisfaction could be a newsletter, increasing the number of parent-teacher meetings, or by highlighting the importance of a good relationship with the parents. Thus, a prioritization could initiate change by providing material resources or by activating mental ones. Principals describe their strategy in a development plan (which is a document with the overall strategy for the school), and superiors in the administration approve the plan. Thus, while it is not possible to create a measure of how many and what kind of resources is entailed in a prioritization, the accountability mechanisms should ensure that improvements come about through functional processes.

Identification strategies and model specifications

The influence of relative performance on principals' prioritizations

Estimating the relationship between relative performance and principals' prioritizations risks endogeneity problems because factors related to the school or the principal could influence both variables. To handle this problem, I use a fixed-effects model. Formally, the model is estimated as follows:

$$Pr_{ist} = \beta_0 + \beta_1 P_{ist} + \beta_2' X_{ist} + \gamma_s + \phi_t + \varepsilon_{ist} \quad (1)$$

Pr_{ist} is the variable that indicates whether goal i is prioritized or not at school s at time t . P_{ist} is the measure of relative performance on a goal i for a school s in time t . X_{ist} is a vector of control variables, γ_s is school fixed effects, ϕ_t is time fixed effects, β_0 the constant, and ε_{ist} is the error term.

In model, time invariant factors at the school level are handled by including the fixed-effects term for the schools (Hesketh and Skrondal 2012: 95–7, 159). However, the model still needs to include factors at the goal level that could influence performance levels and the probability of a prioritization (included in the vector X_{ist}). Two of these control variables relate to fundamental characteristics of a goal, namely, the number of years with a metric and an indicator variable that shows whether performance is measured through a survey.

Furthermore, the political attention of a goal could influence principals' prioritizations of resources and the daily management of the school (thereby influencing performance). To account for such attention, the model includes two indicator variables, measuring whether there is a political aspiration level (decided by the city council) on a goal, and whether this aspiration level has been changed throughout the three cycles. Finally, the model includes

five dummies for goals that have been given special political attention at the national level (GPA, transition to further education, truancy, inclusion, and bullying).

Even though the fixed-effects model makes it possible to handle part of the omitted variable problem, it cannot account for a problem related to the cyclic nature of the data. Within one cycle, there is a sequential relationship between performance information on relative performance (t_0), prioritizations ($t_{1/4}$), and relative performance (t_1). To address this issue, the main analysis is supplemented with analysis split by cycle (in table A2), which allows an estimation of the effect of relative performance on principals' prioritizations independent of the possible performance effect these prioritizations might have.

Another possible confounding factor is principal turnover. Principal turnover could confound the estimates if the change is somehow systematically related to both relative performance and goal prioritization. To account for this problem, the analysis has been robustness checked by splitting the sample for schools with and without principal turnover. This analysis yields substantial identical results to the ones presented in model 4 in table A2 (the coefficient and t-statistic are -0.0037 and -4.62 for schools with no turnover and -0.0069 and -5.59 for schools with turnover).²

Equation 1 has a binary outcome, which makes the models appropriate for both a linear probability model and a logit model (Wooldridge 2009: 246–50). Because the model includes school fixed effects, a substantial number of observations drops out of the estimation if there is no variation in the dependent variable within schools. Therefore, the models are estimated

² Another possible confounding factor is incrementalism in principals' decisions (Lindblom 1959) such that decisions from the first cycle influence decisions in the following cycles. However, a lagged dependent variable in a fixed effects model is by nature endogenous and likely to cause biased estimates (Wooldridge 2002). Lagged dependent variable models are therefore used as a robustness check to the main model. The model does not point to any concern compared the results in model 4 in table A2 (the effect is -0.0047 and the t-statistic - 6.92).

with a linear probability model.³ All models are estimated with cluster-robust standard errors to account for heteroscedasticity in the error term (Angrist and Pischke 2009: 47).

The effect of goal prioritizations on performance

The model that estimates whether prioritized goals influence performance in subsequent years needs to address two issues. First, different factors might influence both a goal's likelihood of being prioritized and the performance level. Second, initial performance differences between non-prioritized and prioritized goals (as expected in H1) would blur the estimation in an ordinary regression model that does not properly capture changes over time.

To handle these issues, I use a difference-in-differences model. The difference-in-differences setup compares developments in the performance trends for non-prioritized and prioritized goals, thereby taking the initial performance differences between the two types of goals into account. Following the difference-in-differences design, non-prioritized goals act as a control group that captures time trends, and prioritized goals receive a treatment in the form of a the prioritization. Formally, the model is given by:

$$P_{ist} = \beta_0 + \beta_1 T_t + \delta(Pr_{ist} \times T_t) + \beta_2' X_{ist} + \gamma_{is} + \varepsilon_{ist} \quad (2)$$

P_{ist} is the outcome, which is the relative performance level for goal i at school s at time t .

Pr_{ist} is the variable that indicates whether a goal is prioritized (i.e., the treatment variable), T_t is the time measure, X_{ist} is a vector of control variables.⁴ This vector includes two indicator variables. One is a variable indicating whether a political aspiration level was introduced on

³ The results from a logit model are reported in table A3 in the appendix.

⁴ The influence of principal turnover has also been tested for this model. Compared to the effect in model 3 in table A4, the effect is 3.83 (t-statistic 3.57) for schools with no turnover and 4.51 (t-statistic 3.51) for schools with turnover.

the performance dimension, the other indicates a in the aspiration level from the previous cycle. γ_{is} is goal fixed effects (that also capture time invariant factors at the school level), β_0 the constant, and ε_{ist} is the error term. δ is the difference-in-differences estimate, consisting of the following terms in a two-period setup:

$$\delta = (P_{prioritized,t1} - P_{prioritized,t0}) - (P_{non-prioritized,t1} - P_{non-prioritized,t0})$$

The basic model is expanded in two ways. First, treatment is normally given at one point in time (Angrist and Pischke 2009). However, the data in this study allows for repeated treatment because the performance management system entails two decision moments. Figure 3 illustrates this point.

FIGURE 3 ABOUT HERE

As shown by the figure, a goal could be treated in either late 2009, late 2011, or in both moments. This feature of the data is handled by creating two treatment groups, namely, goals with prioritization in one cycle (in either 2009 or 2011) and goals with prioritizations in two cycles. The two treatment groups are not collapsed in the first cycle because at the first decision moment principals might decide on a long-term strategy for the goals that are prioritized in two cycles, meaning that these goals experience a different treatment and a unique performance path.

Second, the design rests on the assumption that the control group and the treatment group share a common trend prior to the introduction of the treatment (Wooldridge 2009: 453–4). To test this assumption, the data must include at least two pre-treatment data points for the outcome variable. Thus, the time variable should consist of at least three categories,

namely, two pre-treatment categories and one post-treatment category (Betrand, Duflo and Mullainathan 2004). To accommodate this requirement, I include outcome data for 2007, meaning that t in equation 2 = {2007; 2009; 2010; 2013}.⁵

In a classic difference-in-differences design, the pre-treatment values reflect a situation where all units of analysis are untreated. Such a scenario does not hold in the present setting as performance management reforms were introduced in the Danish educational system already in the beginning of the 2000s, and accordingly, the schools could have worked with goal priorities before the first decision moment in 2009. This aspect of the setting poses a challenge to the design on two levels.

First, it could induce bias in the estimations if a systematic component characterizes pre-2009 prioritizations. The most concerning scenario is if the group of goals that were prioritized in 2009 includes a higher share of pre-2009 prioritized goals. Such a situation would make it difficult to estimate treatment effects after 2009 because the prioritized goals were already on a positive trend (due to the pre-2009 prioritization). One way to assess the problem is the test for common trends as this test indicates whether there is a systematic difference in performance trends between prioritized and non-prioritized goals leading up to the first decision moment in 2009. As shown across the models in table A4, the common trend test does not point to any concerns, showing no significant pre-treatment differences.

To supplement this test, I have run two additional robustness checks. A number of goals did not have a performance indicator before 2009, and thus had a low probability of being prioritized prior to this year. By running the analysis solely on these goals, it is possible to obtain an estimate that is less influenced by the decisions that took place before 2009. Although effects are somewhat larger (compared to model 3 in table A4, the effect is 7.94

⁵ Performance data is not available for all goals in 2007. Data is available for nine goals that are related to *learning* in terms of grades and tests scores, parents' perceptions of *student wellbeing*, and other aspects of *parental satisfaction* and *collaboration* with the school.

and the t-statistic 5.09) for goals with no data prior to 2009, the overall conclusion is the same. The second robustness check is a generalized difference-in-differences model.⁶ This model includes unit and time fixed effects and estimates how changes in treatment status (i.e., whether the goal is prioritized or not) within a unit influence outcomes over time. Since this model does not rely on variation across goals that were either prioritized or non-prioritized in 2009, it is not sensitive to systematic differences between these groups prior to 2009. The alternative difference-in-differences specification yields substantial identical results to the ones presented in model 3 table A4 with an effect of 4.13 (t-statistic 5.51).

While the three robustness checks do not point to concerns with regard to the estimates, the pre-treatment issue is still relevant when interpreting treatment effects. The reason is that effects could partly reflect the work that took place before 2009, which substantially means that positive performance increases are the result of more than one or two prioritizations. Naturally, this point is worth keeping in mind in the analysis.

Finally, three aspects of the model specification and analytic strategy should be noted. First, the analysis consists on three main models (shown in table A4 in the appendix). In each model, non-prioritized goals act as the control group (i.e. the reference category). Models 1 and 3 test the effect of prioritizing a goal in one cycle (2009 and 2011 respectively), and model 2 tests the effect of repeated prioritizations. Second, in addition to these models, the analysis includes two models that are split by goals within a school that have either low or high performance (divided by the median). Third, the difference-in-differences models are estimated with the full set of observations. This model is robust to a completely balanced panel-model that only includes goals with data from all four time periods (compared to model 3 in table A4, the effect is 3.54 and t-statistic 2.85).

⁶ Formally described as $P_{it} = \beta_0 + \beta_1 \text{prioritization}_{it} + \beta_2' X_{ist} + \gamma_i + \phi_t + \varepsilon_{ist}$

ANALYSIS

Figure 4 illustrates the relationship between relative performance and principals' tendency to prioritize a goal. The β_1 in the figure is the estimated regression coefficient from the model specified in equation 1.

FIGURE 4 ABOUT HERE

Based on figure 4, it is possible to evaluate the expectation in H₁ that public managers will use performance information to prioritize goals with the lowest level of performance. As evident in the figure, this expectation is confirmed. β_1 is significantly negative, which means that goals with lower performance are more likely to be prioritized. In substantial terms, a performance difference of 10 percentage points between two goals will lead to a 5 percent increase in the probability that a principal prioritizes the goal with the lowest level of performance. For a standardized outcome, the effect is 0.17 standard deviations. This estimate underlines a substantial important aspect of the analysis, namely that even though relative performance influences principals' prioritizations, the performance differences need to be quite substantial before the information offsets a managerial reaction. Such differences are rather common for the principals as they, on average, experience a difference of 63 percentage points between the highest and lowest performing goals in their school. Across such large performance differences, the probability of prioritizing the goals with low performance is 40 percent higher.

The last part of the analysis concerns the influence of priorities on performance in subsequent years. First is the question of whether prioritizations are merely symbolic (H₂), or they actually create improvements but primarily for goals with a low initial performance level (H₃). The analysis also includes a test of H₄, which claims that the longer a goal is prioritized,

the higher the probability of performance improvements. These hypotheses are tested in difference-in-differences models, and the results are summarized in figure 5 (the full models are shown in table A4 in the appendix).

FIGURE 5 ABOUT HERE

The results confirm the expectations in H₃ and H₄. As evident by the figure, goals that have been prioritized in two cycles experience a performance increase in the second cycle. The effect is around four scale points on the performance scale. In relation to H₃, the analysis split for goals with low and high performance levels confirms the expectation that improvements occur for goals with a low initial level of performance (the result is statistically significant, $p < 0.01$, in a three-way interaction). Figure 5 also show that short term prioritizations do not create improvements. Goals that are prioritized in one cycle experience a small (insignificant) increase in performance (0.56 for goals that are prioritized in 2009 and 0.58 for goals that are prioritized in 2011) in the first cycle; however, they do not gain the most substantial performance increase because they are not reprioritized.

At first sight, the effect of around four scale points might not seem overwhelming in substantial terms as the scale of relative performance ranges from 19 to 100. However, when standardizing the coefficient, the effect of 0.25 standard deviations presents itself as quite substantial. In relation to other studies that focus on interventions in public schools, Rockoff and colleagues (2012) find that providing principals with reports on teachers' performance increases students' math achievements by 0.053 standard deviations. Andersen and colleagues (2016) show that increasing instruction time in schools improves students' reading by 0.15 standard deviations. In relation to these findings, it is worth noting that the effect covers different strategies for creating improvements. In some instances, the schools have

made major changes in their organization, facilities, and curriculum. However, improvements also reflect minor initiatives that aim specifically at the particular goal, for example to create a newsletter to the parents.

The last part of the analysis concerns H_5 and the expectation that a prioritization of goals with low performance will end in a performance surplus for the organization over the years. Performance trends for non-prioritized goals are estimated by the time variable (i.e., the effect of time when treatment is at the reference category) in models 1, 2 and 3 in table A4. When summarizing the coefficients from these models and comparing them to the treatment effects, the overall positive result confirm the expectation that principals are able to improve performance on goals with low performance without substantial performance trade-offs as performance for non-prioritized goals increase over time.

CONCLUSION

This article examines how the performance on various organizational goals influence managers' tendency to prioritize the goals and whether the prioritization of certain goals improves performance in subsequent years. The results show that principals in a municipality in Denmark prioritize the educational goals where their schools are performing the worst. However, the information has the most profound influence on prioritizations when there is a substantial discrepancy between goals with low and high performance. After a goal has been made a priority for several years, the performance level increases. As non-prioritized goals do not experience a drop in performance, the schools end in an overall performance surplus over the years, suggesting minimal performance trade-offs. These results are noteworthy in themselves, but they also have broader implications for our understanding of how performance management fits in the public sector.

First, as the results show that public managers are able to engage in functional performance management practices, they naturally raise the question of what drives the purposeful use of performance information. Understanding when performance management efforts make a difference is essential in order to move beyond the simplistic claims that performance management does or does not work (Gerrish 2016). One apparent contextual factor in the present setting is the learning forums where managers discuss the information with stakeholders. While previous research suggests that such forums support performance information use (Moynihan 2005), there is not much knowledge on what makes them functional. The forums in this study are interesting because they strike a balance between outside involvement and managerial autonomy. The deliberation between stakeholders might help overcome some of the biases that has been shown in individual processing of performance information (Nielsen and Moynihan 2017; Olsen 2017) while at the same time allowing managers to feel ownership of the process.

This point also relates to a general trait of the system, namely, the balance between autonomy and coercion. On the one side, the system includes various checkpoints for principals, forcing them to make a decision and develop a strategy for improvements. However, they are also given autonomy in their choice of priorities and strategy. In this way, the findings point to an important balance between finding the right means to engage managers but leaving sufficient room for management (Nielsen 2014a). Finally, it is also worth mentioning that the context is just as interesting from the perspective of what is not present; there are no financial incentives at stake for the principals. Thus, the results suggest that it is possible to create a functional management process in the absence of incentives but with the presence of deliberative learning forums.

A second contribution is the idea of managers as biased problem solvers. More specifically, I not only document that the level of performance affects managers' decisions; I

also document that such decisions still facilitate problem-solving by directing attention to goals where the greatest returns to investment may occur. The principals' emphasis on negative information (low performance) is identified for one specific decision, namely, the prioritization of organizational resources and attention. It remains to be seen how managers react to the content of performance information in other settings, for example, when the information is used in the management of employees (Behn, 2003).

Third, an important result from this study is that performance gains require ongoing persistence. This result has two practical implications. First, there are no quick fixes in public organizations. However, a political environment may demand new initiatives each year and grow impatient if a strategy does not yield a fast payoff. Such rapid changes are likely to undercut long-term thinking and strategies that need time to bear fruit. Second, there is an inherent risk in performance management systems that attempts at improving performance are ruined by too little and too superficial attention. As the results show that one cycle (one and a half year) is probably not enough to improve performance, too frequent shifts in priorities will most likely result in wasted resources with no performance gains. If performance strategies are to pay off, they need to be deployed with a medium- and long-term perspective.

REFERENCES

- Andersen Simon Calmar, Maria K. Humlum and Anne B. Nandrup. 2016. Increasing instruction time in school does increase learning. *Proceedings of the National Academy of Sciences* 113(27):7481–7484.
- Andersen, Simon Calmar and Donald P. Moynihan. 2016. How leaders respond to diversity: The moderating role of organizational culture on performance information use. *Journal of Public Administration Research and Theory* 26(3):448–460.
- Andersen, Simon Calmar. 2008. The impact of public management reforms on student performance in Danish schools. *Public Administration* 86(2):541–58.
- Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly harmless econometrics: An empiricists companion*. Princeton: Princeton University Press.
- Askim, Jostein, Åge Johnsen, and Knut-Andreas Christophersen. 2008. Factors behind organizational learning from benchmarking: Experiences from Norwegian municipal benchmarking networks. *Journal of Public Administration Research and Theory*, 18 (2): 297–320.
- Audia Pino G., Sebastian Brion and Heinrich R. Greve. 2015. Self-assessment, self-enhancement, and the choice of comparison organizations for evaluating organizational performance. *Advances in Strategic Management: Cognition and Strategy* 32: 89–118.
- Audia, Pino G. and Sebastian Brion. 2007. Reluctant to change: Self-enhancing responses to diverging performance measures. *Organizational Behavior and Human Decision Processes* 102(2): 255-269.

- Banerjee, Abhijit, Raghavendra Chattopadhyay, Esther Duflo, Daniel Keniston and Nina Singh. 2012. Can institutions be reformed from within? Evidence from a randomized experiment with the Rajasthan Police. *Working paper*.
- Baum, J. A. and K. B. Dahlin. 2007. 'Aspiration Performance and Railroads' Patterns of Learning From Train Wrecks and Crashes', *Organization Science*, 18, 3, 368-385.
- Behn, Robert D. 2003. Why measure performance? Different purposes require different measures. *Public Administration Review* 63(5):586–606.
- Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics* 119(1):249–75.
- Bevan, Gwyn, and Christopher Hood. 2006. What's measured is what matters: Targets and gaming in the English public healthcare system. *Public Administration* 84(3):517–38.
- Bourdeaux, Carolyn and Grace Chikoto. 2008. Legislative influences on performance management reform. *Public Administration Review* 68(2): 253-65.
- Boyne, George A. 2002. Concepts and Indicators of Local Authority Performance: An Evaluation of the Statutory Frameworks in England and Wales. *Public Money and Management* 22(2):17–24.
- _____. 2003. What is Public Service Improvement?. *Public Administration* 81 (2):211-227.
- Boyne, George A., Oliver James, Peter John, and Nicolai Petrovsky. 2010. Does public service performance affect top management turnover? *Journal of Public Administration Research and Theory* 20(2):261-279.

- Busuioc, Madalina and Martin Lodge. 2017. Reputation and accountability relationships: Managing accountability expectations through reputation. *Public Administration Review* 77: 91–100.
- Carpenter, Daniel. 2001. *The forging of bureaucratic autonomy: Reputations, networks, and policy innovation in executive agencies*. Princeton, NJ: Princeton University Press
- Carpenter, Daniel. 2010. *Reputation and Power: Organizational Image and Pharmaceutical Regulation at the Fda*: Princeton Univ Press
- Cyert, Ricard M., and James G. March. 1963. *A Behavioral Theory of the Firm*. Englewood Cliffs, N.J: Prentice-Hall.
- Feldman, Martha S. and James G. March. 1981. Information in organizations as signal and symbol. *Administrative Science Quarterly* 26(2): 171–186.
- Gerrish, Ed. 2016. The impact of performance management on performance in public organizations: A meta-analysis. *Public Administration Review* 76(1):48–66.
- Gilmour, John B., and David E. Lewis. 2006. Does performance budgeting work? An examination of the office of management and budget's PART scores. *Public Administration Review* 66(5):742–52.
- Hanushek, Eric A. 2003. The failure of input-based schooling policies. *The Economic Journal* 113: 64–98.
- Heckman, James J. 2000. Policies to foster human capital. *Research in Economics* 54(1):3–56.

- Heinrich, Carolyn J. 1999. Do government bureaucrats make effective use of performance management information? *Journal of Public Administration Research and Theory* 9(3):363–93.
- _____. 2010. Third-party governance under no child left behind: Accountability and performance management challenges in supplemental educational services provision. *Journal of Public Administration Research and Theory* 20(1):59–80.
- Hesketh, Sophia Rabe, and Anders Skrondal. 2012. *Multilevel and longitudinal modeling using Stata*. 3. edition. College Station, Texas: Stata Press Publication.
- Hood, Christopher. 2011. *The blame game: Spin, bureaucracy, and self-preservation in government*. Princeton, NJ: Princeton University Press.
- Jacob, Brian A., and Stephen D. Levitt. 2003. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics* 118(3):843–877.
- Jordan, A. H. and P. G. Audia. 2012. Self-enhancement and learning from performance feedback. *Academy of Management Review* 37(2): 211–231.
- KC, Diwas, Bradley R. Staats and Francesca Gino. 2013. Learning from my success and from others' failure: Evidence from minimally invasive cardiac surgery. *Management Science* 59: 2435-2449
- Kelman, Steven, and John N. Friedman. 2009. Performance improvement and performance dysfunction: An empirical examination of distortionary impacts of the emergency room wait-time target in the English National Health Service. *Journal of Public Administration Research and Theory* 19:917–46.

- Krause, George A. 2003. Coping with Uncertainty: Analyzing Risk Propensities of SEC Budgetary Decisions, 1949-1997. *American Political Science Review* 97(1):171-188.
- Kroll, Alexander. 2013. The other type of performance information: Nonroutine feedback, its relevance and use. *Public Administration Review* 73(2):265–76.
- _____. 2015a. Exploring the link between performance information use and organizational performance: A contingency approach. *Public Performance and Management Review* 39(1):7–32.
- _____. 2015b. Explaining the use of performance information by public managers: A planned-behavior approach. *American Review of Public Administration* 45(2):201–15.
- Lavertu, Stéphane, David E. Lewis, and Donald P. Moynihan. 2013. Government reform, political ideology, and administrative burden: The case of performance management in the Bush administration. *Public Administration Review* 73(6): 845–57.
- Levinthal, Daniel A., and James G. March. 1981. A model of adaptive organizational search. *Journal of Economic Behavior and Organization* 2(4):307–33.
- Lindblom, Charles E. 1959. The science of “muddling through”. *Public Administration Review* 19 (2): 79–88.
- MacDonald, Jason A. and William W. Franko. 2007. Bureaucratic Capacity and Bureaucratic Discretion: Does Congress Tie Policy Authority to Performance?. *American Politics Research* 35(6): 790-807.
- Maor, Moshe, Sharon Gilad, and Pazit Ben-Nun Bloom. 2013. Organizational reputation, regulatory talk, and strategic silence. *Journal of Public Administration Research and Theory* 23(3): 581–608.

- March, James G. 1991. Exploration and exploitation in organizational learning. *Organization Science*, 2(1):71-87.
- Meier, Kenneth J., and Laurence J. O'Toole Jr. 2013. Subjective organizational performance and measurement error: Common source bias and spurious relationships. *Journal of Public Administration Research and Theory* 23(2):429–56.
- _____. 2008. Management Theory and Occam's Razor: How Public Organizations Buffer the Environment. *Administration and Society* 39(8):931–58
- Meier, Kenneth J., Nathan Favero, and Ling Zhu. 2015. Performance gaps and managerial decisions: a Bayesian decision theory of managerial action. *Journal of Public Administration Research and Theory* 25(1):1221-1246.
- Melkers, Julia, and Kathrine Willoughby. 2005. Models of performance-measurement use in local governments: Understanding budgeting, communication, and lasting effects. *Public Administration Review* 65(2):180–90.
- Migue, Jean-Luc, Gérard Belanger and William A. Niskanen. 1974. Towards a general theory of managerial discretion. *Public Choice* 17: 27-51.
- Mintzberg, Henry, Duru Raisinghani, and Andre Théorêt. 1976. The structure of “unstructured” decision processes. *Administrative Science Quarterly* 21:246–75.
- Moynihan, Donald P. 2005. Goal-based learning and the future of performance management. *Public Administration Review* 65:203–16.
- _____. 2008. *Dynamics of performance management: Constructing information and reform*. Georgetown University Press.

- Moynihan, Donald P. 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.
- Moynihan, Donald P., Sanjay K. Pandey, and Bradley E. Wright. 2011b. Setting the table: How transformational leadership fosters performance information use. *Journal of Public Administration Research and Theory* 22(1):143–64.
- Moynihan, Donald P., Sergio Fernandez, Soonhee Kim, Kelly LeRoux, Suzanne J. Piotrowski, Bradley Wright, and Kaifeng Yang. 2011a. Performance Regimes amidst Governance Complexity. *Journal of Public Administration Research and Theory* 21(s1):i141-155.
- Nielsen, Poul A. 2014a. Performance management, managerial authority, and public service performance. *Journal of Public Administration Research and Theory* 24(2):431–58.
- Nielsen, Poul A. 2014b. Learning From Performance Feedback: Performance Information, Aspiration Levels, and Managerial Priorities. *Public Administration* 92(1):142–60.
- Nielsen, Poul A., and Donald P. Moynihan. 2017. How do politicians attribute bureaucratic responsibility for performance? Negativity bias and interest group advocacy. Forthcoming in *Journal of Public Administration Research and Theory*.
- O'Toole, Laurence, and Kenneth J. Meier. 2011. *Public management: Organizations, governance, and performance*. Cambridge: Cambridge Univ. Press.
- Olsen, Asmus. 2017. Compared to what? Experimental evidence on social and historical reference points in performance evaluation. Forthcoming in *Journal of Public Administration Research and Theory*.

- Ostrom Elinor, Roger B. Parks, Gordon P. Whitaker, and Stephen L. Percy. 1978. The public service production process: a framework for analyzing police services. *Policy Studies Journal* 7(1):381–89.
- Percy, Stephen. 1983. Citizen coproduction: Prospects for improving service delivery. *Journal of Urban Affairs* 5(3):203–10.
- Poister, Theodore H. 2010. The Future of Strategic Planning in the Public Sector: Linking Strategic Management and Performance. *Public Administration Review* 70 (1): 246-254
- Pollitt, Christopher. 2013. The logics of performance management. *Evaluation* 19(4):346-363.
- Pritchett, Lant, and Deon Filmer. 1999. What education production functions really show: A positive theory of education expenditure. *Economics of Education Review* 18(2):223–39.
- Rich, Robert F. 1997. Measuring knowledge utilization: Process and outcomes. *Knowledge and Policy* 10(3):11–24.
- Rockoff, Jonah E., Douglas O. Staiger, Thomas J. Kane, and Eric S. Taylor. 2012. Information and employee evaluation: Evidence from a randomized intervention in public schools. *American Economic Review* 102: 3184–3213.
- Rutherford, Amanda, and Kenneth J. Meier. 2015. Managerial Goals in a Performance-Driven System: Theory and Empirical Tests in Higher Education. *Public Administration* 93(1): 17–33.

- Salge, Torsten O. 2011. A Behavioral Model of Innovative Search: Evidence from Public Hospital Services. *Journal of Public Administration Research and Theory* 21(1):181–210.
- Soss, Joe, Richard Fording, and Sanford F. Schram. 2011. The organization of discipline: From performance management to perversity and punishment. *Journal of Public Administration Research and Theory* 21(2):203–32.
- Spence, Michael. 2002. Signaling in retrospect and the informational structure of markets. *American Economic Review* 92(3): 434-459.
- Taylor, Jeannette. 2009. Strengthening the link between performance measurement and decision making. *Public Administration* 87(4):853–71.
- Vakkuri, Jarmo. 2010. Struggling with ambiguity: Public managers as users of NPM-oriented management instruments. *Public Administration* 88(4):999–1024.
- van Helden, Jan G., Åge Johnsen, and Jarmo Vakkuri. 2012. The life cycle approach to performance management: Implications for public management and evaluation. *Evaluation* 18(2): 159–175.
- Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.
- Wooldridge, Jeffrey M. 2009. *Introductory econometrics – a modern approach*. South Western College Pub.
- Yang, Kaifeng, and Jun Yi Hsieh. 2007. Managerial effectiveness of government performance measurement: Testing a middle-range model. *Public Administration Review* 67(5):861–79.

TABLES AND FIGURES

Figure 1: The production-possibility frontier for goals with high and low performance

Figure 2: The cyclic nature of the system

Figure 3: Time sequence in the difference-in-differences setup

Figure 4: How relative performance influences principals' prioritization of goals

Note: The full model with control variables is shown in table A2 in the appendix.

Figure 5: How prioritization of goals influences performance in subsequent years

Note: The figure is based on models 1 and 3 from table A4 in the appendix.

Table 1: Characteristics of the municipality - the school year 2012/2013.

	Municipality			Denmark
	Min	Max	Mean	Mean
School size (number of students)	124	1021	602	428
Budget pr. student (in DKK)	-	-	62746	61065
Share of students in special classes	5%	7%	4,3%	5,2%
Share of bilingual students	2,7%	99,7%	25,0%	~10%
Share of teachers' time used for teaching	25,2%	38,7%	32,3%	33,7%
Student/teacher ratio	9	18	14	13
Competence-coverage ¹	33%	100%	70,0%	80,0%

Note: The table is based on data from the quality reports and the Danish ministry of education.¹ The share of classes being conducted from a teacher with specialization in the subject.

APPENDIX

A1: Descriptive statistics for goals.

Goal [min; max]	Mean ¹	Std. deviation ¹	Prioritized ²
School GPA [19; 71]	54.81	8.63	0.35
School GPA – lowest fourth [59; 83]	70.73	5.37	0.54
School GPA – highest fourth [48; 98]	83.42	9.05	0.47
Test – numbers (3rd grade) [44; 100]	88.71	10.30	0.29
Test – reading (3rd grade) [71; 100]	94.23	5.7	0.35
Test – reading (8th grade) [38; 100]	90.71	8.75	0.35
Transition: 3 months [45; 100]	93.62	6.43	0.21
Transition: 15 months [30; 100]	82.61	14.84	0.21
P ³ . satisfaction: transition [39; 84]	61.44	9.50	0.20
S ⁴ . satisfaction: boredom [55; 96]	89.42	4.00	0.41
P ³ . satisfaction: academic challenge [35; 100]	58.85	11.04	0.39
S ⁴ . satisfaction: friends [79; 100]	96.35	4.44	0.16
S ⁴ . satisfaction: co-determination [54; 89]	72.11	6.23	0.16
S ⁴ . satisfaction: victim of bullying [63; 100]	93.98	4.22	0.21
P ³ . satisfaction: ability to stop bullying [47;81]	69.26	8.98	0.21
P ³ . satisfaction: fellowship [37; 84]	70.95	8.83	0.16
P ³ . satisfaction: equal opportunities [35; 84]	51.79	10.54	0.24
S ⁴ . satisfaction: like the class [47; 100]	94.50	8.11	0.23
P ³ . satisfaction: bullying occurs [63; 97]	88.67	5.96	0.20
Truancy [53; 96]	90.03	7.81	0.21
Worrisome truancy [50; 100]	85.26	11.25	0.24
S ⁴ . satisfaction: happiness in general [47; 100]	96.26	6.00	0.25
S ⁴ . satisfaction: happy at the time [47; 100]	98.10	3.54	0.27

S ⁴ . satisfaction: like the school [70; 100]	91.41	3.58	0.27
S ⁴ . satisfaction: recognition [81; 100]	93.32	2.39	0.25
P ³ . satisfaction: child's well-being [70; 92]	83.70	4.60	0.24
Student exercise [27; 100]	75.83	22.61	0.29
Student overweight [51; 100]	84.76	7.30	0.30
Student smoking [67; 100]	95.65	4.35	0.14
Student drunkenness [64; 100]	92.29	5.60	0.16
Quality of teeth: filling [0; 79.8]	56.16	11.78	0.17
Quality of teeth: cavity [50; 100]	92.22	7.10	0.17
P ³ . satisfaction: collaboration [52; 92]	75.86	6.45	0.52
P ³ . satisfaction: own contribution [60; 88]	75.31	7.51	0.51
P ³ . satisfaction: expectations [32; 69]	51.70	6.66	0.51
P ³ . satisfaction: school-satisfaction [51; 97]	77.47	8.39	0.51
P ³ . satisfaction: youth-center-satisfaction[44; 98]	78.03	10.76	0.53
P ³ . satisfaction: daily contact [43; 90]	65.73	8.73	0.51
P ³ . satisfaction: involvement [23; 86]	49.15	10.19	0.51
Total	82.25	16.13	0.31

Note: ¹= of relative performance; ²= the relative number of prioritizations, across all years; ³= parental; ⁴ = student

A2: How relative performance influences principals' prioritization of goals. OLS – schools fixed effects. Clustered standard errors.

Model	1	2	3	4
Cycle	2009	2011	2013	All
Relative performance	-0.0035* (-2.66)	-0.0055*** (-4.92)	-0.0043*** (-4.48)	-0.0050*** (-7.06)
Survey indicator	0.086* (2.17)	-0.030 (-0.67)	-0.12** (-3.18)	-0.013 (-0.48)
Number of years measured	0.030** (2.80)	0.026** (2.91)	0.012 (1.75)	0.029*** (5.29)
Political aspiration level	0.0021 (0.09)	-0.14* (-2.44)	-0.12** (-3.01)	-0.069* (-2.07)
Change in aspiration level		0.028 (0.61)	-0.029 (-0.82)	-0.046* (-2.51)
Goal: Inclusion	-0.22*** (-4.01)	0.073 (1.06)	-0.039 (-0.77)	-0.078* (-2.36)
Goal: GPA	-0.096 (-1.37)	-0.18*** (-3.86)	-0.057 (-0.91)	-0.13** (-3.00)
Goal: Transition to further education	-0.23*** (-4.19)	-0.0082 (-0.09)	-0.13 (-1.81)	-0.14** (-3.16)
Goal: Bullying	-0.33*** (-4.41)	-0.12 (-1.49)	-0.19*** (-3.82)	-0.18*** (-5.20)
Goal: Truancy	0.039 (0.50)	-0.21** (-3.12)	-0.19* (-2.13)	-0.10* (-2.04)
Cycle: 2009				Reference
Cycle: 2011				0.10** (2.80)

Cycle: 2013				0.081
				(1.94)
Constant	0.43***	0.83***	0.83***	0.64***
	(4.46)	(7.96)	(7.52)	(10.12)
<i>N</i> (schools)	49	50	47	53
<i>N</i> (goals)	1519	1839	1763	5121
F-test	8.76***	8.12***	6.24***	11.81***
Adjusted R^2	0.062	0.054	0.033	0.036

Note: *=p<0.05; **=p<0.01; ***=p<0.001 (two-sided test). t-statistics in parenthesis.

A3: How relative performance influence principals' prioritization of goals. Logistic regression – schools fixed effects.

Model	1	2	3	4
Cycle	2009	2011	2013	All
Relative performance	-0.019*** (-4.04)	-0.029*** (-6.36)	-0.022*** (-4.50)	-0.024*** (-9.44)
Survey indicator	0.50** (3.11)	-0.16 (-1.06)	-0.63*** (-4.31)	-0.061 (-0.75)
Number of years measured	0.18** (2.73)	0.14** (2.65)	0.061 (1.24)	0.15*** (4.94)
Political aspiration level	-0.027 (-0.10)	-0.71*** (-4.50)	-0.60*** (-3.40)	-0.31** (-3.15)
Change in aspiration level		0.16 (0.60)	-0.14 (-0.80)	-0.23* (-2.55)
Goal: inclusion	-2.25** (-3.02)	0.39 (1.07)	-0.32 (-0.76)	-0.53* (-2.20)
Goal: GPA	-0.45 (-0.98)	-0.93* (-2.29)	-0.33 (-0.80)	-0.65** (-2.73)
Goal: transition to further education	-2.21*** (-4.05)	-0.028 (-0.10)	-0.70* (-2.36)	-0.77*** (-4.35)
Goal: bullying	-2.07*** (-4.03)	-0.64 (-1.61)	-1.15* (-2.55)	-0.96*** (-4.18)
Goal: truancy	0.24 (0.83)	-1.35*** (-3.75)	-1.13** (-3.28)	-0.56** (-3.06)
Cycle: 2009				Reference
Cycle: 2011				0.49*** (4.68)

Cycle: 2013

0.39***

(3.54)

<i>N</i> (schools)	45	49	45	52
<i>N</i> (goals)	1402	1801	1687	5083
Log likelihood	-637.39	-884.47	-852.23	-2770.02
Likelihood ratio	112.76***	108.44***	67.48***	194.53***

Note: *=p<0.05; **=p<0.01; ***=p<0.001 (two-sided test). t-statistics in parenthesis.

**Table A4: How prioritization of goals influences performance in subsequent years.
Difference-in-differences estimates – OLS (goal fixed effects).**

Model	1	2	3	3.1	3.2
Goals are prioritized in	One cycle (09-11)	One cycle (11-13)	Two cycles (09-13)	Two cycles (09-13)	Two cycles (09-13)
Sample	All goals	All goals	All goals	Low performance goals	High performance goals
<hr/>					
Time					
2007	-0.72 (-1.17)	-1.97** (-3.02)	-0.66 (-1.14)	-2.24** (-2.85)	-4.21*** (-9.62)
2009	Ref. (.)	-1.27*** (-3.33)	Ref. (.)	Ref. (.)	Ref. (.)
2011	1.36*** (3.58)	Ref. (.)	1.14** (3.17)	2.43*** (5.03)	-0.23 (-0.85)
2013	0.21 (0.57)	-1.03* (-2.53)	0.28 (0.80)	4.42*** (8.16)	-0.31 (-1.35)
<hr/>					
Diff-in-diffs estimates					
Prioritization * 2007	-1.83 (-1.46)	0.52 (0.47)	-1.19 (-1.11)	0.46 (0.37)	1.07 (1.07)
Prioritization * 2009	Ref. (.)	0.78 (1.16)	Ref. (.)	Ref. (.)	Ref. (.)
Prioritization * 2011	0.56 (0.75)	Ref. (.)	1.10 (1.38)	1.34 (1.47)	0.012 (0.02)
Prioritization * 2013	0.21 (0.27)	0.51 (0.78)	4.07*** (4.97)	3.10** (3.25)	1.07 (1.48)
Constant	81.9*** (349.23)	82.3*** (270.62)	80.6*** (353.92)	66.1*** (191.91)	93.6*** (615.52)
Controls	Yes	Yes	Yes	Yes	Yes
N	4165	4491	4028	1958	2070
F-test	5.04***	2.56**	10.38***	29.04***	15.27***
Adjusted R ²	0.014	0.0040	0.006	0.014	0.11

Note: *= $p < 0.05$; **= $p < 0.01$; ***= $p < 0.001$ (two-sided test). t-statistics in parenthesis. Coefficients that test for a common trend are in *cursive*. The models includes controls for whether a political aspiration level was introduced on the performance dimension and whether this level changed.