

Drivers of Performance Information Use: Systematic Literature Review and Directions for Future Research

Forthcoming in Public Performance & Management Review

Alexander Kroll

Assistant Professor of Public Administration

Florida International University

akroll@fiu.edu

Abstract

The use of performance information in decision-making is a management behavior that has received much attention in public administration research and practice. This article seeks to contribute to a better understanding of this behavior. It conducts a systematic review of 25 recently published empirical studies which have examined drivers of performance information use. Analyzing these studies, which were selected on the basis of their definition of purposeful data use, this article identifies factors that have repeatedly shown a positive impact – measurement system maturity, stakeholder involvement, leadership support, support capacity, an innovative culture, and goal clarity. This systematic analysis also uncovers less conclusive variables; findings which are highly relevant for future studies. Based on the review, the article suggests directions for further research endeavors, including theoretical and methodological propositions.

Introduction

Performance measurement practices have become widespread across the OECD world (Bouckaert & Halligan, 2008; OECD, 2007; Pollitt and Bouckaert, 2011). In recent years, a growing number of initiatives to hold public administration accountable have been carried out in various countries across several level of governments (Bouckaert & Halligan, 2008; Van de Walle & Bovaird, 2007). The logic behind these reforms is that governments have to justify their expenditures by giving account of what has been achieved with it in order to ensure that tax money is spent efficiently and effectively. To do so, public administration has to collect information on the impact of its services and report it. However, just collecting data does not automatically lead to improvements, which is why practitioners and researchers declared that it is essential to understand what happens to performance information after it is has been reported and evaluated (Hatry, 2006; Van Dooren & Van de Walle, 2007). Only if performance information is used in day-to-day decision-making, can performance-oriented reforms be assessed as being successful. When public managers deliberately utilize this information, they display not only compliance with the formal routines of data production but also show that they buy-in to the original reform idea that more systematic feedback about the work of public administration is needed to enable better-informed decisions (Moynihan, 2008; Van Dooren, Bouckaert, & Halligan, 2010).

Studying the use and non-use of performance data in decision-making has become a highly relevant and fast growing research area (Ammons & Rivenbark, 2008; de Lancer Julnes & Holzer, 2001; Dull, 2009; Kroll, 2013; Melkers & Willoughby, 2005; Moynihan & Pandey, 2010; Moynihan & Hawes, 2012; Taylor, 2011). Studies have aimed at determining

which factors¹ foster or constrain performance information use in order to theorize about this behavior as well as to facilitate it. For a research field to mature, it needs reviews summarizing the current state of inquiry, which further research can be built on. Though every empirical study also provides a narrative review of the literature, systematic reviews are very rare. A systematic review only includes articles which share a very similar understanding of the empirical phenomenon under study (Merton, 1968) – in this case, performance information use by public managers. Instead of providing a broad overview of an extensive literature, a systematic review focuses on a manageable sample of comparable studies and analyzes them in depth. This includes findings for control variables which usually receive little attention but are highly informative since one study's control might be another's variable of main interest (Cooper, 2010).

This article provides a systematic review of the literature on performance data use by managers in public administration, and its contribution is three-fold. First of all, the review aims to be more transparent than common narrative reviews. The article selection is systematic, all studies are closely examined, summarized, and every single one of their findings is operationalized, counted, and weighted. Secondly, the article differentiates between impact factors that vary in their influence; which is essential for a research field to mature. Its systematic character makes the review more than just the sum of its parts and its conclusions more easily generalizable. Thirdly, the article provides theoretical and methodological suggestions for future research directions. Before conducting this review, the article specifies its empirical phenomenon of interest – the use of performance data. This is followed by an explanation of the review method and a discussion of the review results, existing research gaps, and potential directions for future research.

¹ In this article, the terms factors, variables, determinants, and drivers are used interchangeably. For the purpose of this paper it was not necessary to further differentiate among these terms.

Purposeful Performance Information Use

This section will explain how the studies, included in this review, operationalize the concept of performance information use. When speaking of performance data, the studies refer to feedback information on outputs and outcomes of the public service as well as its efficiency (a ratio of outputs and inputs) and its effectiveness (a ratio of outcomes and outputs). This is in line with established definitions which highlight that performance information is more than just financial data and has a specific focus on the results and achievements of public administration (Hatry, 2006; Pollitt & Bouckaert, 2011). A second characteristic is that performance information is in a quantitative, aggregated format, and it is made transparent through reports or data bases. Its collection is not ad-hoc but follows a systematic control-cycle logic where indicators for goal achievement are defined, and performance information is collected and analyzed, and it is supposed to be used for future decision-making (see figure 1; for further explanations, see de Lancer Julnes, Berry, Aristigueta, & Yang, 2007; Kroll, 2013; Van Dooren et al., 2010).

[Figure 1]

Though performance data could be used to serve many functions, this review will focus on a type of information use that has been labelled “purposeful” (Moynihan, 2009). This includes using performance information to improve services through better informed decisions, goal-based learning, or sanctioning and rewarding. Similarly, Van Dooren and colleagues (2010) identify learning (which is mostly concerned with future improvements) as

well as steering and controlling (keeping track of present activities) as purposeful functions. Most of Behn's (2003) more disaggregated activities also fall in this category – using performance data to evaluate, control, budget, motivate, celebrate, learn, and improve (see table 1).

There are two reasons why this review will focus on antecedents of purposeful use. Thus far, most studies have focused on this type of data use, which is why reviewing empirical findings on purposeful use seems to be the most promising path to follow in order to conduct a systematic meta-evaluation. The second reason is that this type of use has been recognized as desirable, and it is consistent with the normative expectations promoted by performance-oriented reforms in public administration.

Though purposeful use is the behavior that has received the most attention in the conceptual and empirical literature, research has also shown that managers could respond differently when confronted with performance measurement. Managers might just use performance data passively (Moynihan, 2009) while answering to stakeholders (Van Dooren et al., 2010) without drawing conclusions within their own zone of discretion (for an empirical study, see Moynihan & Lavertu, 2012). Or, they use performance information politically (Moynihan, 2009) to “promote” (Behn, 2003) their interest in turf wars, budget negotiations, or political arguments where its interpretation is highly subjective, controversial, and role-induced (for an empirical study, see Moynihan, Pandey, & Wright 2012a). Another type of use has been labelled dysfunctional or perverse (Moynihan, 2009). It includes behaviors, such as gaming, cheating, and cherry-picking (for empirical studies, see Bevan & Hood, 2006; Hood, 2006; Kelman & Friedman, 2009). Table 1 provides an overview of this typology.

[Table 1]

Data and Method

Article Selection

A systematic review needs criteria based on which articles are included in or excluded from the analysis. To do so, the review follows Merton's (1968) understanding of how to build middle-range theories. He argues that the most promising approach to advance research in the social sciences is to aim for theories which "lie between the minor but necessary working hypotheses that evolve in the abundance during day-to-day research and the all-inclusive systematic efforts to develop a unified theory that will explain all the observed uniformities of social behavior, social organization, and social change" (Merton, 1968, p. 39). That is, researchers will make a valuable contribution to the generation of knowledge in their fields if they study clearly defined and specifically delimited concepts that are empirically observable (Pinder & Moore, 1979). Using this approach for a literature review has advantages and disadvantages. On the one hand, it means excluding all studies which are only broadly tied to the concept under investigation. On the other hand, this approach ensures that only comparable studies, which share very similar operationalizations, will eventually be compared with each other, and it solves the apples-and-oranges problem.

A middle-range perspective suggests avoiding overly abstract concepts ("human responses to feedback information") and, instead, focusing on a more specific one, such as the use of performance information by decision-makers. However, even the latter is not really exclusive, as it would basically include a great deal of the research that has been conducted in the organizational sciences and all related areas, following the seminal work by March, Guetzkow, and Simon (1958). To deal with this issue and further delimit the outcome

variable, the review will only include research on “purposefully used” “performance data” and thus exclude information and uses that deviate from the understanding put forward in the previous section. It will also focus on data use by managers and consequently exclude other user groups like politicians (Askim, 2009; Ter Bogt, 2004) or citizens (Ho, 2008; James, 2011) as well as studies from the areas of business and accounting (Broadbent & Laughlin, 2009; Marchand, Kettinger, & Rollins, 2002) or broader public sector research on the topic of information systems, such as Bozeman and Bretschneider (1984) or Bretschneider (1990).

Again, this procedure might result in a smaller sample of studies to review, but it will ensure that the concepts and findings in the selected articles will actually be comparable which is important for an in-depth analysis. To identify all relevant empirical articles, all Public Management and Public Administration journals accessible through the data bases *WEB OF KNOWLEDGE*, *JSTOR*, and *EBSCOhost* were searched. In addition to peer-reviewed published articles, papers that were referenced on the website of the *Performance Information Project*² and already accepted for publication in a *SSCI*-listed journal were also considered. Applying this method of analysis, the following 25 studies were identified (see table 2).³ Most of our knowledge about the use of performance information comes from research originated in the United States (16 studies, 64%). Three Studies are based on data from Germany, two from Australia, one from Norway, one from Sweden, one from Wales, and one from Taiwan. The majority of the articles (14) studied the local level, seven articles focused on the state, two on the federal level, and two had a mixed data set. The mean R^2 of

² <http://www.lafollette.wisc.edu/publicservice/performance/index.html> (retrieved on March 3, 2013)

³ In a first step, 28 studies were selected. Due to the missing disclosure of all used operationalizations or the use of rather idiosyncratic explanation factors, in a second step three studies were excluded from the sample. In order to avoid an overrepresentation of singular data sets, only one regression model per publication (if several had been presented) were included in the review. The review always focused on the model that took the most control variables into account and where the dependent variable came the closest to the presented understanding of purposeful information use.

the models predicting performance information use is 39%.⁴ This indicates that the studies were on average able to explain a good deal of the observed variation and that the factors tested thus far had reasonable effect sizes.

[Table 2]

Review Method

The advantage of a systematic review over a narrative review is that the former does not only include the significant results of empirical studies but also the insignificant ones. Since published articles, the basis for a literature review, are biased towards the presentation of significant findings (Rosenthal 1979; Sutton 2009), reviewers will usually read more pages about the validation of hypotheses than about their rejections or other insignificant observations. A systematic review aims to balance this bias to a certain extent by taking into account the findings for control variables which usually receive little attention but are highly informative since one study's control might be another's variable of main interest. Afterwards, a simple "vote count" (comparing the number of positive, negative, and insignificant "votes") can be used to assess the relative explanatory significance of every potential influence factor. Such an analysis is highly informative as it makes a big difference as to whether a variable that has been found to be significant in three studies at the same time turned out to be insignificant in another three or even five articles (for a similar approach, see Boyne & Walker, 2010).

⁴ Since some studies only reported either the R^2 or the Adjusted R^2 , my estimation took both indicators into account. In cases where both measures were reported, their mean was used.

A systematic review or “vote count analysis” is different from a statistical meta-analysis. The latter does not count the number of differently directed effects but estimates average effect sizes of the variables weighted by their sample sizes. A meta-analysis is mainly applicable if the studies of interest have reached a high quantity, if they report standardized correlational information for all used variables, and if they have applied similar research designs (Cooper, 2010). Particularly the first two requirements are not fulfilled by the studies under investigation. Out of a relatively small sample of 25 studies, only nine articles reported standardized correlational information on all the variables that were used which makes further statistical meta-analyses of the variables’ effect sizes not feasible. Counting positive, negative, and insignificant studies is a straightforward procedure. However, results from studies which are based on a large number of observations are more trustworthy than from a small-N sample which is why it makes sense to weight the vote counts.⁵ Every finding which was significant at a p level below 0.1 was considered and categorized. Results from case studies were counted as simple directional votes (positive or negative) and not further qualified in terms of their significance. The weighting assigned lesser importance to those studies.

Results and Discussion

The results of the systematic review of antecedents of purposeful performance information use can be found in table 3. The first column lists all factors that were identified, and it is divided in the three categories of environmental, organizational (performance-measurement related and unrelated), and individual variables which is in line with previous research (Kroll & Vogel, 2013; Melkers & Willoughby, 2005; Moynihan & Pandey, 2010).

⁵ All vote counts were weighted using the following formula: $\text{VoteCount}_{\text{weighted}} = 1 - \frac{1}{\sqrt{n}}$

The second column displays the number of studies which have tested each impact factor. The next five columns show the unweighted vote counts for negative, insignificant and positive effects (at a 0.1 and 0.05 level of significance). The last three columns display the weighted counts, however, here the positive and negative effects are not divided according to their significance but are all at least significant at a 0.1 level. The first insight from table 3 is that the weighted results do not differ a lot from the unweighted ones, suggesting that most of the findings are based on reliable samples sizes and that case studies and small samples are the exception.⁶ Though table 3 is arranged according to the established categories of environmental, organizational, and individual factors, the review will proceed by focusing on the “Important Drivers of Data Use”, followed by the “Promising Impact Factors” and the “Insignificant Variables”. This way, the discussion will be structured in a more pointed manner and focus on the most critical differentiations without losing focus due to the attempt to evaluate every variable from every category individually.

[Table 3]

Important Drivers of Data Use

This category includes all variables which have been found to be essential triggers of data use. That is, several studies tested these determinants (four was used as the threshold), a clear majority (three) found evidence for a positive effect, while the rest did not entirely

⁶ For example, regarding the first variable “stakeholder involvement” the table shows nine counts (large-N samples and single-case studies counted equally) in support of a positive effect at a significance level of at least 0.1 (two plus seven counts). Similarly, the last column shows that there are also 8.43 positive counts if one assigns weights to every study’s sample.

contradict this finding. Table 4 lists these variables, provides examples of how they have been measured and cites studies which found support for their impact.

The two most prominent factors in this category are measurement system maturity and stakeholder involvement. More sophisticated measurement systems, which go beyond the simple production of raw data, make information use more likely. These systems provide a good range of different data, align the reporting to the demands of the addressees, link information to goals and strategic plans, and offer benchmarks. Though all the studies listed in table 4 find corroborating evidence for this effect, one study also points out that more complex and sophisticated measurement systems – though they foster managerial data use – might also make political control more difficult if they reinforce information asymmetries between bureaucrats and politicians (Kroll & Proeller, 2013).

The involvement of external stakeholders also matters a great deal. Stakeholders can encourage managers to take performance information seriously, and they can also help in making sense of numbers or in identifying meaningful indicators. Knowing that stakeholders care about performance data adds “political weight” (Ho, 2006) and signals to managers that they should be on top of their department’s data, performance trends, and explanations of outliers (Moynihan & Ingraham, 2004). Since there were enough studies that tested and reported the correlations between data use and system maturity as well as stakeholder involvement, it was possible to compare their mean Pearson’s r values. It turns out that on average system maturity (0.47) has a stronger bivariate effect than stakeholder involvement (0.23).

[Table 4]

Another important factor is leadership support for performance measurement practices. Top-level support is a good measure for line managers to evaluate whether reforms or initiatives are taken seriously in organizations and whether they really need take on the effort to participate and devote their scarce resources accordingly. Studying leadership behavior and its effects during the implementation of the Government Performance and Results Act, Dull (2009, p. 273) concludes the following: “If managers suspect agency leaders are prone or simply will not be around to make good on reform commitments, their attention and effort will gravitate to other problems and priorities.” Boyne et al. (2004) make a similar point when they summarize that feedback information loses importance in local authorities when senior managers take a hands-off approach to performance measurement and management practices.

Similarly critical is the support capacity for performance management practices which can be defined as the resources, capabilities, and technology that are available to make performance measurement work. This finding clearly indicates that the success of performance management is highly dependent on how these systems are adopted and that early investments in their support can pay off later on. De Lancer Julnes and Holzer (2001) find that capacity is an important factor which impacts the successful adoption and use of performance measurement systems. Berman and Wang (2000) also present corroborating evidence for the positive effect of an adequate technical infrastructure while examining different dimensions of capacity.

Several studies reported that performance information is used more intensely in public organizations which lean towards an innovative or developmental culture. Such a culture embraces openness to change and values learning from mistakes in order to improve for the future. Performance information seems to be appreciated here because it facilitates learning and improvement by providing additional feedback. Furthermore, an innovative culture can foster data use because it emphasizes dialogue and discussion rather than reward and

punishment, which is why performance information will appear less threatening to employees (Moynihan et al., 2012b). A strong goal orientation also shows a positive effect. If organizations have clearly stated goals, it is also more likely that the achievement of these objectives will be regularly discussed and evaluated which adds weight to the organizations' performance information.

Promising Impact Factors

This sub-section will discuss variables that were only examined by two or three studies but found positively related to data use in all of them. One of these factors is the establishment of learning forums or routines which offer mechanisms for discussing and collectively making sense of this information. They create opportunities for managers and employees to reflect on their core processes and related outcomes, and provide communication channels to make solutions to problems on the individual level usable for the whole organization (Moynihan, 2005).

[Table 5]

Promising factors on the individual level are the managers' attitude, prosocial motivation, and networking behavior. The finding that managers' enthusiasm (Ammons & Rivenbark, 2008), interest (Ho, 2006) or positive attitude (Taylor, 2011) matter, clearly shows that the existence of organizational routines is often not sufficient. Instead, for performance information to become relevant in decision-making, managers have to be convinced that these data will help them to improve the outcomes of their work. Another important factor is the

managers' motivation. If managers are driven by a strong other-orientation, they are also willing to take on the extra effort that the analysis, reporting, and use of performance information require (Kroll & Vogel, 2013) because these data will help them to track and increase their organizations' social impact (Moynihan et al., 2012a). A final favorable individual characteristic is the managers' networking behavior. Individuals who are more generally open to feedback from outside the organization (professional networks, see Kroll, 2013, as well as relevant stakeholders, see Moynihan & Hawes, 2012) were also found to be frequent users of performance data.

Two promising environmental variables are political support and the fragmentation or heterogeneity of the community. Political support is critical because it gives public organizations the certainty that problematic performance data will not automatically lead to budget cuts or the organizations' delegitimization and thus not trigger blame-avoidance behavior when confronted with performance feedback (Yang & Hsieh, 2006). In fragmented environments, performance data are used more intensely because managers need them to justify their decision to a heterogeneous field of competing stakeholders. These data might also be seen as a necessary additional information source to successfully maneuver in an environment that requires tailored solutions for diverse problems (Bourdeaux & Chikoto, 2008).

Insignificant or Inconclusive Variables

There are also factors that have been tested several times but were found to be mainly insignificant. One of them is the size of an organization. Though one could hypothesize that larger organizations are more professionalized and can devote more resources to their performance management practices, it cannot be concluded that they show higher levels of

data use once other variables have been controlled for (Bourdeaux & Chikoto, 2008; Melkers & Willoughby, 2005; Moynihan & Ingraham, 2004; Taylor, 2011). One explanation is that the overall size does not matter after more specific measure of resources like “support capacity” have been taken into account. Another explanation might be that one should not mix up the variables adoption and implementation of performance measurement. Though larger organizations might be more likely to adopt more sophisticated systems, this will not automatically increase data use when other important conditions, such as stakeholder involvement, leadership support, or an innovative culture, are not present (Johansson & Siverbo, 2009).

Similar findings exist for the financial situation that public organizations find themselves in. Most studies reported that this variable does not really make a difference (Askim, Johnsen, & Christophersen, 2008; Berman & Wang, 2000). Again, financial distress might put pressure on public administrations to adopt performance measurement, but it does not necessarily foster data use (Johansson & Siverbo, 2009).

[Table 6]

The findings for a politically competitive environment turn out to be mostly inconclusive. Though research indicates that it does not have a negative effect on purposeful data use, there are as many positive as insignificant results. There is scattered evidence that in an environment with dynamic government-opposition constellations performance information can become important ammunition for defending or attacking political positions (Askim, Johnsen, & Christophersen, 2008; Bourdeaux & Chikoto). However, this effect was also found to be insignificant when other factors were taken into account (Johansson & Siverbo,

2009; Moynihan & Lavertu, 2012). Similarly inconclusive are the results for managers' familiarity with performance measures. There seems to be a positive baseline effect which, however, vanishes when models become further specified (Dull, 2009; Melkers & Willoughby, 2005).

Most of the socio-demographic variables, such as job experience (Dull, 2009; Melkers & Willoughby, 2005; Taylor, 2011), hierarchical position (de Lancer Julnes & Holzer, 2001; Taylor, 2011), or educational level (Moynihan & Ingraham, 2004; Moynihan & Hawes, 2012) turned out to be insignificant or inconclusive. Though differences on the individual level can matter, simple socio-demographic characteristics do not seem to be able to pick up on variations in data use.

Limitations

There are some limitations which need to be acknowledged when generalizing this study's results. One limitation is that the vote count analysis focused on the significance of the effects but not on their sizes (for a discussion, see Rosenthal & DiMatteo, 2001). That is, effects have been equally counted, no matter whether their regression coefficients were 0.1, 0.3, or 0.6, as long as they were significant at a 10% or 5% level. This study might have even been slightly biased towards effects of smaller magnitudes because these are more likely to achieve statistical significance in larger samples, and this study assigned larger weights to those samples. Since only very few of the articles under investigation provided the standardized information necessary for further statistical examination of the effect sizes, such a meta-analytic approach was not feasible. It is also worth noting that the results of statistical meta-analyses are not always more informative than the findings from simpler methods.

Though it would be interesting to know whether there are variables which have an impact on

data use twice as high as other factors, the informative value of comparing average effect sizes such as, for example, 0.14, 0.2, and 0.22 is limited.

The review is based on a systematic selection of relevant studies, and it documented all positive, negative and insignificant effects, but it is still possible that an article or a variable were missed. However, these blind spots, if existent, were not generated because of any systematic error but are due to chance. Furthermore, the review is based on 25 carefully selected studies. This means that many studies from a large area of performance management were not considered when their measures deviated too far from those used in this article.

The Achilles' heel of every literature review is publication bias (Togerson, 2006). Articles reporting insignificant findings are less likely to ever get published which leads to a bias towards significant counts. However, the review included each study's control variables and reported, as shown in table 3, a good number of insignificant findings. Another limitation comes with the design of the literature review that focused only on direct effects on data use but did not account for mechanisms which are more indirect or implicit. This point will be taken up in the next section that discusses directions and approaches for future research.

Future Research Directions

One important function of the review was to systematically analyze the selected studies in order to contribute to a better understanding of purposeful use by distinguishing between important drivers of data use, promising impact factors, and insignificant variables. Before drawing conclusions from these findings, the current section is devoted to the gaps of the existing literature. It will discuss potential directions for further research and provide theoretical and methodological propositions.

Proposition 1: More research is needed on indirect and contingency effects.

For a research field to mature, investigators will need to consider indirect and contingency effects. Studying indirect effects means explaining variation in a dependent variable not only by using a list of independent factors, but also by including mediators which will provide a better (stepwise) understanding of the mechanisms behind direct effects. This is in line with Merton's (1968) approach of developing midrange theories, as he emphasizes that it is critical to examine the interplay of several variables in order to be able to grasp complex social phenomena. Contingency theories are another approach to better account for real world's complexity and in particular for the fact that some effects will occur in certain environments but not in others (Donaldson, 2001). In other words, the impact of a variable on another variable could be contingent on the existence or variation of a third one. This is why context can matter and why certain managerial interventions only work in certain situations (Otley, 1980; Pollitt, 2013).

In the case of purposeful information use, an exemplary study of an indirect effect is the one by Moynihan et al. (2012b). It found that transformational leadership does not have a significant direct effect on data use, but it has an indirect impact: Transformational leadership facilitates goal clarity and an innovative culture which both foster managerial data use. Another study detected a contingency effect (Kroll & Vogel, 2013). The effect of a prosocial motivation on the use of performance information is moderated by the transformational leadership behavior of managers' supervisors.

Proposition 2: More research is needed on the role of the potential information users.

The literature review showed that very little is known with regard to effects on the individual level. Even though most of the socio-demographic variables used so far have yielded insignificant or inconclusive results, variation among information users could still matter. Drawing on other research, one could speculate that different learning styles might make a difference, as it could be assumed that different preferences for how to learn – which means to grasp and transform information – might also explain how managers make use of performance data (Kolb & Kolb, 2009; Riding & Cheema, 1991).

Another interesting variable could be the ownership of performance information shown by public managers. That is, the purposeful use of data might be a function of the involvement of managers in the measurement process far before performance reports have reached their desks. One might hypothesize that data use becomes more likely when managers participate in the selection and customization of measures and when they show credible commitment to a performance-based steering philosophy (Behn, 1994; Dull, 2009). What might also matter is the role identity of the managers. Particularly in countries with a traditionally strong legalistic public service ethos, purposeful performance information use might require switching identities from passively implementing state servants to “shaping”, risk-taking managers (Pratchett & Wingfield, 1996; Meyer & Hammerschmid, 2006).

Proposition 3: More research is needed on the role of different information types.

All the studies which were included in the review focused on one particular type of performance information – quantitative, aggregated data on efficiency and effectiveness that were systematically collected, following a control-cycle logic. Though this type of information was strongly promoted as the essential component of a management-for-results

movement, it is questionable whether this is the most important feedback source for decision-makers within the politico-administrative system. Mintzberg (1975) has put forward the argument that business managers receive much of their most valuable feedback through calls, meetings, and observational tours rather than backwards-oriented information systems. Studying politicians' information preferences, ter Bogt (2004) and Van Dooren (2004) found that these preferences are situational and can differ among policy sectors. Kroll (2013) has argued that research might miss important facets of managers' receptiveness to performance feedback if investigators only study the use of quantitative data. Instead, one should also pay attention to nonroutine performance information, which includes qualitative feedback that is often collected on an ad-hoc basis or passively received instead of actively pursued through an established management routine. It might be particularly interesting how the uses of different types of performance information are related to each other, and whether there is substitution or complementarity.

Proposition 4: More research is needed on the connection of purposeful data use to actual performance improvements.

Purposeful information use is not an end in itself. More research is required on the effects of managerial data use. Or in other words, do agencies do better when their managers are enthusiastic users of performance reports? Though there are a couple of very recent studies which have examined the effects of performance measurement on organizational performance (Hvidman & Andersen, 2013; Nielsen, 2013; Poister, Pasha, & Edwards, 2013; Sun & Van Ryzin, 2013), there are barely any studies which have had a main focus on the connection of data use to performance.

Another related research gap is the measurement of the performance information use variable. When examining this behavior, most studies conceptualized it as being individual rather than collective. That is, individual managers were the units of analysis instead of organizational routines which could only be reliably observed using multiple informant surveys for each organization. Though focusing on individuals is absolutely legitimate, treating purposeful use as an organizational routine might be the more promising approach when studying effects on organizational performance (Argyris & Schön, 1996; Levitt & March, 1988).

Proposition 5: More experimental research is needed.

Most of the articles in this review have used survey-based approaches which are usually cross-sectional (that is, it is often difficult to determine which variable causes which), pick up on people's attitudes rather than observing actual behavior, and are prone to an omitted variable bias (it remains uncertain whether all relevant alternative explanations have been controlled). To avoid these pitfalls, experiments might be a promising research method. Experiments allow one to observe behavior or real-life decision-making. Here, researchers actively manipulate a predictor variable of interest and can observe whether this affects an outcome, which makes it easy to determine causality. To deal with omitted variable bias, this manipulation is random which means that, for example, out of two randomly divided groups one receives a treatment whereas the other serves as a control group. Because of this randomization, experiments automatically control for all potential confounding variables which are equally present in both groups and therefore cannot explain differences in outcomes (Shadish, Cook, & Campbell, 2002). There are exemplary studies which examined the effects of positive and negative performance feedback on citizens' attitudes towards local

government (James, 2011) and on politicians' spending decisions (Nielsen & Baekgaard, 2013). Similar designs could also be used to study the behavior of public managers.

Quasi-experiments are another interesting approach. Here, the independent variables of interest are usually not manipulated by the researcher which can be an advantage, because using "natural interventions" makes it possible to conduct experiments in the field which increases external validity (Grant & Wall, 2009). However, this can also be a disadvantage because in these cases it is almost impossible to use randomization as a way to control for alternative explanations. Every situation where a public organization decides to adopt (changes to) a performance measurement system as a pilot project ("natural intervention") and does not roll this out over the entire organization is an opportunity to realize a quasi-experiment. In these cases, it will be possible to examine whether the "pilot units" ("treatment group") show unique patterns of data use or decision-making.

Conclusion

This article offered a systematic literature review of factors influencing purposeful performance information use, analyzing 25 recently published empirical studies. Based on this review, it distinguished between important drivers, promising factors, and insignificant variables. Such a differentiation is essential because theoretically all the variables under investigation could be related to purposeful data use, but the review showed that empirically only some of them are. This section will draw the major conclusions from this article's findings.

First of all, there are the important drivers of data use which have shown significant positive effects study after study. These variables – measurement system maturity, stakeholder involvement, leadership support, support capacity, an innovative culture, and goal

clarity – need to be included in future models of purposeful use, at least as important controls. Though this review is probably more useful to researchers than to practitioners, it does suggest that, in order to facilitate purposeful use among managers, trying to establish these factors in an organization is the best bet for public administrators. It goes without saying that in practice some of these factors (leadership support, support capacity) are easier to establish than others (stakeholder involvement, innovative culture).

Secondly, there are promising impact factors, including learning routines, the users' (prosocial) motivation, networking behavior, and political support, which will need more attention in future research. They were found to be relevant, but corroborating evidence as well as more in-depth examinations of the mechanisms behind these effects are still required. In particular, it will be interesting to see whether these factors will still be significant when the variables from the first category ("important drivers") have been taken into account.

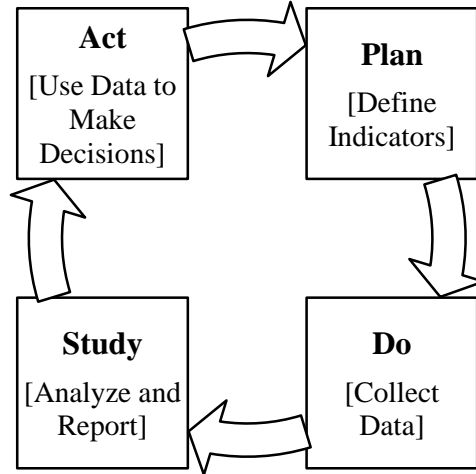
Thirdly, the review identified insignificant and inconclusive variables, such as an organizations' size and financial situation or managers' socio-demographics. One interesting implication from these findings certainly is that these variables do not form obstacles for information use. It seems fair to say that larger organizations without financial problems and certain demographics are not more likely to observe enthusiastic data use among their managers than any other organizations. Though these factors were mainly found to be irrelevant, the review does not suggest dismissing them entirely from further research. They should receive less attention in future public administration survey projects, particularly when space is scarce, but there is still very little research on the role of these variables as potential contingency factors.

Finally, based on the review, research gaps were identified and theoretical as well as methodological propositions formulated. It was suggested that future research needs to include indirect and contingency effects, focus on the role of the information users and

different information types, connect the research on purposeful data use to the literature on organizational performance, and make use of experimental designs. Ways how to do this were outlined and examples provided.

Tables and Figures

Figure 1: Performance Management Cycle



Source: Deming’s (1994, p. 132) PDSA cycle adapted to the case of performance management

Table 1: Different Uses of Performance Information

Dimensions of Use				Author(s)
Purposeful	Passive	Political	Perverse	Moynihan (2009)
Learn	Give Account			Van Dooren et al. (2010)
Control		Promote		Behn (2003)
Evaluate				
Control				
Budget				
Motivate				
Celebrate				
Learn				
Improve				

Table 2: Overview of Selected Studies

Author(s) (Year)	Study Title (abbreviated)*	Journal Abbrev. **	Country	Level of Governm.	Metho- dology (Model)	Sample	Summary
Ammons & Rivenbark (2008)	“Factors influencing the use of performance data to improve municipal services”	PAR	USA	local	case studies	15 cases	Study on the use of benchmarking data which identifies system maturity, data quality and managers’ attitudes as critical factors.
Askim et al. (2008)	“Factors behind organizational learning from benchmarking”	JPART	Norway	local	quantitative (“Solution Model”)	n=138	Study on the use of benchmarking data to change policies or routines which focuses on network and administrative characteristics as well as political factors.
Berman & Wang (2000)	“Performance measurement in U.S. counties”	PAR	USA	local	quantitative	n=155	Study examines the outcomes of multiple purposeful uses which points to stakeholder involvement and support capacity as critical factors.
Bourdeaux & Chikoto (2008)	“Legislative influences on performance management reform”	PAR	USA	state	quantitative (Extent of Use Model 2)	n=124	Study on the extent of effective PI use which finds that stakeholder involvement and political competition are important.
Boyne et al. (2004)	“Toward the self-evaluating organization?”	PAR	Wales	local	quantitative	n=79	Study on the use of PI to self-evaluate an authority’s services which focuses on leadership support and employee involvement as critical variables.
de Lancer Julnes & Holzer (2001)	“Promoting the utilization of performance measures in public organizations”	PAR	USA	state/local	quantitative (Including Adoption Model)	n=363	Study on the use of PI for multiple management purposes which points to the critical role of political and cultural factors.
Dull (2009)	“Results-model reform leadership”	JPART	USA	federal	quantitative (Outcome 2000 Model)	n=1,860	Study on the use of PI across a variety of management functions which identifies leadership commitment as a critical variable.
Folz et al. (2009)	“The adoption, use, and impacts of performance measures in medium-size cities”	PPMR	USA	local	quantitative	n=157	Study examines the perceived usefulness of PI which finds that employee involvement and the familiarity with the measures matter.
Ho (2006)	“Accounting for the value of performance measurement	JPART	USA	local	quantitative (Second-	n=250	Study examines the perceived impact of PI which suggests that stakeholder involvement

Johansson & Siverbo (2009)	from the perspective of Midwestern mayors” “Explaining the utilization of relative performance evaluation in local government”	FAM	Sweden	local	Stage Model) quantitative (OLS Model)	n=210	and managers’ attitudes and involvement are critical. Study on the use of benchmarking data which concludes that cultural and political factors are essential.
Kroll (2013)	“The other type of performance information”	PAR	Germany	local	quantitative	n=195	Study on several uses of PI which identifies, among other things, managers’ networking as a crucial variable.
Kroll & Proeller (2013)	“Controlling the control system”	IJPSM	Germany	local	case studies	2 cases	Study on the use of PI which finds that more sophisticated, complex measurement systems increase the use by administrators but can also make political control more difficult.
Kroll & Vogel (2013)	“The PSM-leadership fit”	PA	Germany	local	quantitative	n=200	Study on the use of PI for several management functions which examines the motivational basis of this behavior and concludes that prosocial motivation is important.
Melkers & Willoughby (2005)	“Models of performance-measurement use in local governments”	PAR	USA	local	quantitative (Lasting Effects Model)	n=197	Study on the use of performance information to create purposeful lasting effects which suggests that the maturity of the measurement system and employee involvement are relevant factors.
Moynihan (2005)	“Goal-based learning and the future of performance management”	PAR	USA	state	case studies	3 cases	Study on the use of PI for organizational learning which finds that learning forums and organizational culture are fundamental determinants.
Moynihan & Hawes (2012)	“Responsiveness to reform values”	PAR	USA	state	quantitative	n=516	Study on PI use for several management functions which concludes that managers’ general openness to the environment can foster internal data use.
Moynihan & Ingraham (2004)	“Integrative leadership in the public sector”	A&S	USA	state	quantitative (Agency Model)	n=42	Study on the use of PI in decision-making which concludes that commitment by the direct superior is important.
Moynihan & Landuyt	“How do public organizations learn?”	JPART	USA	state	quantitative	n=24,614	Study on the use of feedback information for organizational learning which suggests that

(2009)								both structural and cultural factors matter and in particular learning forums.
Moynihan & Lavertu (2012)	“Does involvement in performance management routines encourage performance information use?”	PAR	USA	federal	quantitative	n=1,422		Study on purposeful and passive use of PI (from GPRA and PART) which concludes that, among other things, leadership support and learning forums are critical impact factors on purposeful use.
Moynihan & Pandey (2010)	“The big question for performance management”	JPART	USA	local	quantitative	n=1,132		Study on the use of PI (one-item measure) which finds that, among other things, information availability, culture, and motivation matter.
Moynihan et al. (2012a).	“Prosocial values and performance management theory”	Govern.	USA	local/NPO	quantitative	n=183		Study on the use of PI for several management functions which examines the motivational basis of this behavior and concludes that prosocial motivation is important.
Moynihan et al. (2012b)	“Setting the table”	JPART	USA	local	quantitative	n=1,538		Study on the use of PI which suggests that transformational leadership has a positive influence, as it creates an innovative culture and goals clarity.
Taylor (2009)	“Strengthening the link between performance measurement and decision making”	PA	Australia	state	qualitative + quantitative (Internal Use Correlates)	12 cases		Study which finds that internal (purposeful) and external (passive) use are disconnected and that internal use is triggered by more sophisticated systems.
Taylor (2011)	“Factors influencing the use of performance information for decision making in Australian state agencies”	PA	Australia	state	quantitative (Combined Model)	n=53		Study on PI use which reports that the measurement system, stakeholder support, organizational culture, and the external environment are essential.
Yang & Hsieh (2006)	“Managerial effectiveness of government performance measurement”	PAR	Taiwan	local	quantitative	n=684		Study on the effectiveness of PI which focuses on the political environment and stakeholder participation as crucial factors.

* Full titles can be found in bibliography.

** Journal abbreviations: A&S = Administration & Society; Govern. = Governance; FAM = Financial Accountability and Management; IJPSM = International Journal of Public Sector Management; JPART = Journal of Public Administration Research & Theory; PA = Public Administration; PAR = Public Administration Review; PPMR = Public Performance & Management Review

Table 3: Systematic Overview of Impact Factors and Results

Factor	# Tests	Unweighted					Weighted		
		neg **	* *	ns	* *	pos **	neg */**	ns	pos */**
Environmental									
Stakeholder Involvement	13	0	0	4	2	7	0	3.68	8.43
Political Competition	7	0	0	4	1	2	0	3.73	2.78
General Political Support	2	0	0	0	0	2	0	0	1.89
Political (Mis-)Use Risk	2	0	0	2	0	0	0	1.80	0
Heterogeneity/Fragmentation	2	0	0	0	0	2	0	0	1.87
Organizational									
<i>PM-related</i>									
Measurement System Maturity	18	0	0	5	1	12	0	4.50	11.21
Leadership Support	7	1	0	1	0	5	0.93	0.91	4.65
Support Capacity	6	0	0	2	0	4	0	1.87	3.79
Employee Involvement	4	0	1	0	0	3	0.95	0	2.74
Data Quality/Usability	3	0	0	1	1	1	0	0.85	1.72
Information Availability	2	0	0	0	1	1	0	0	1.80
<i>PM-unrelated</i>									
Organization Size	9	0	0	7	1	1	0	6.45	1.90
Innovative Culture	8	0	0	3	2	3	0	2.81	4.21
Financial Distress	6	1	0	5	0	0	0.99	4.66	0
Goal Orientation/Clarity	4	0	0	1	0	3	0	0.9	2.9
Internal Task/Service	3	2	0	1	0	0	1.90	0.93	0
Learning Forums/Routines	3	0	0	0	1	2	0	0	2.38
Flexibility/Decision Discretion	3	0	0	1	1	1	0	0.93	1.94
Transf. Leadership Supervisor	2	0	0	2	0	0	0	1.90	0
Hierarchical Culture	2	0	0	2	0	0	0	1.86	0
Group Culture	2	0	0	2	0	0	0	1.86	0
% Unionized	2	0	2	0	0	0	1.87	0	0
Individual									
Job Experience	11	3	0	8	0	0	2.88	7.54	0
Familiarity with Perf. Measures	6	0	0	3	1	2	0	2.88	2.77
Upper-Level Position	6	1	0	3	0	2	0.97	2.85	1.85
Higher Educational Level	4	1	0	3	0	0	0.99	2.74	0
Gender (female)	3	0	0	1	0	2	0	0.97	1.92
Attitude towards Perf. Mngmt.	3	0	0	0	1	2	0	0	2.54
Prosocial Motivation	3	0	0	0	1	2	0	0	2.83
Networking Behavior	2	0	0	0	0	2	0	0	1.89
Age	2	0	0	1	0	1	0	0.86	0.99

“neg” = negative effect, “ns” = not significant, “pos” = positive effect, * significant at a 0.1 level, ** significant at a 0.05 level, */** significant at least at a 0.1 level, two-tailed-test logic was applied

Table 4: Important Drivers of Data Use

Factor	Exemplary Measures	Exemplary Studies
Measurement System Maturity	“Performance indicators are easy to access, [...], have stretching but achievable performance targets, [...] and meet performance information needs.”; “Managers are involved in a PI-based benchmarking with other cities.”	Ammons and Rivenbark (2008); Berman and Wang (2000); de Lancer Julnes and Holzer (2001); Ho (2006); Kroll and Proeller (2013); Melkers and Willoughby (2005); Moynihan and Pandey (2010); Taylor (2009); Yang and Hsieh (2006)
Stakeholder Involvement	“The extent of support for performance measures by elected officials and/or citizens.”; “Citizens help this organization evaluate performance.”	Berman and Wang (2000); Bourdeaux and Chikoto (2008); Moynihan and Ingraham (2004); Ho (2006); Moynihan and Hawes (2012); Moynihan and Pandey (2010); Yang and Hsieh (2006)
Leadership Support	“Agency’s top leadership demonstrates a strong commitment to achieving results.”; “Top managers emphasize and care about the process of performance management.”	Boyne et al. (2004); Dull (2009); Moynihan and Ingraham (2004); Moynihan and Lavertu (2012); Yang and Hsieh (2006)
Support Capacity	“The extent to what the organization has committed resources (time, people, money) to be used in measurement of program performance”; “Most departments in our jurisdiction have adequate information technology for performance measurement.”	Berman and Wang (2000); de Lancer Julnes and Holzer (2001); Moynihan and Hawes (2012); Yang and Hsieh (2006)
Innovative Culture	“My department is a very dynamic and entrepreneurial place. People are willing to stick their necks out and take risks.”; “The glue that holds my department together is a commitment to innovation and development. There is an emphasis on being best.”	Folz, Abdelrazek, and Chung (2009); Johansson and Siverbo (2009); Moynihan (2005); Moynihan and Pandey (2010); Moynihan, Pandey, and Wright (2012b)
Goal Clarity	“It is easy to explain the goals of this organization to outsiders.”; “This organization’s mission is clear to those who work here.”	Moynihan and Landuyt (2009); Moynihan, Pandey, and Wright (2012a; b)

Table 5: Promising Impact Factors

Factor	Exemplary Measures	Exemplary Studies
Learning Forums/Routines	“Work groups are actively involved in making work processes more effective.”; “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.”	Moynihan (2005); Moynihan and Landuyt (2009); Moynihan and Lavertu (2012)
Attitudes towards Performance Measures	“Enthusiasm about performance measurement”; “Performance measurement has brought more advantages than disadvantages to my unit/agency.”	Ammons and Rivenbark (2008); Ho (2006); Taylor (2011)
Prosocial Motivation	“Making a difference in society means more to me than personal achievements.”; “I am prepared to make sacrifices for the good of society.”	Kroll and Vogel (2013); Moynihan and Pandey (2010); Moynihan, Pandey, and Wright (2012a)
Networking Behavior	“Assessment of how frequently managers interact with different relevant stakeholders.”; “I maintain a regular exchange with colleagues from other cities who work in the same field.”	Kroll (2013b); Moynihan and Hawes (2012)
General Political Support	“Most elected officials trust our organization,”; “Most elected officials believe that our organization is effective.”	Moynihan, Pandey, and Wright (2012a); Yang and Hsieh (2006)
Fragmented Environment	“A measure which captures the extent to which the population is heterogeneous across the dimensions of race, rural or urban residence, education, and homeownership.”; “The stakeholders in this school district fulfill in general their agreements with one another. (reversed).”	Bourdeaux and Chikoto (2008); Moynihan and Hawes, (2012)

Table 6: Insignificant or Inconclusive Variables

Factor	Exemplary Measures	Exemplary Studies
Organization Size	“State/city population”; “Number of employees”	Bourdeaux and Chikoto (2008); Johansson and Siverbo (2009); Kroll (2013); Melkers and Willoughby (2005); Moynihan and Ingraham, (2004); Taylor (2011)
Financial Distress	“Net operating result per capita, average of the last years” (reversed); “Fiscal stress is the mean of the financial results (per inhabitant) from the two fiscal years before the study”	Askim, Johnsen, and Christophersen (2008); Berman and Wang (2000); Johansson and Siverbo (2009); Kroll (2013); Moynihan and Pandey (2010);
Political Competition	“Party concentration in the municipal council indicated by 1-Herfindahl Index”; “A dummy variable capturing whether different parties control the House and Senate”	Askim, Johnsen, and Christophersen (2008); Bourdeaux and Chikoto (2008); Moynihan and Hawes (2012)
Familiarity with Performance Measures	“Reported number of years working with Performance measures”; “Awareness and knowledge of GPRA and its requirements”	Askim, Johnsen, and Christophersen (2008); Dull (2009); Melkers and Willoughby (2005)
Job Experience	“Years in present position”; “Length of service”	Dull (2009); Melkers and Willoughby (2005); Moynihan and Pandey (2010); Moynihan et al. (2012b); Taylor (2011)
Hierarchical Position	“Non-Supervisor = 1; Team Leader = 2; First-Line Supervisor = 3; Middle Manager = 4; Top Manager = 5”; “1 = middle manager; 0 = senior executive”	de Lancer Julnes and Holzer (2001); Moynihan et al. (2012a); Taylor (2011)
Educational Level	“High School/GED (or less) = 1; Some College = 2; Bachelor’s Degree = 3; Graduate Degree = 4”	Moynihan and Ingraham (2004); Moynihan et al. (2012a); Moynihan and Hawes (2012)

References

- Ammons, D., & Rivenbark, W. (2008). Factors influencing the use of performance data to improve municipal services: Evidence from the North Carolina Benchmarking Project. *Public Administration Review*, 68(2), 304–31.
- Argyris, C. & Schön, D. 1996. *Organizational learning: Theory, method, and practice*. Reading, Mass.: Addison-Wesley Pub.
- Askim, J., Johnsen, A., & Christophersen, K.-A. (2008). Factors behind organizational learning from benchmarking: Experiences from Norwegian municipal benchmarking networks. *Journal of Public Administration Research and Theory*, 18(2), 297–320.
- Askim, J. (2009). The Demand side of performance measurement: Explaining councillors' utilization of performance information in policymaking. *International Public Management Journal*, 12(1), 24–47.
- Behn, R. (1994). *Leadership counts: Lessons for public managers from the Massachusetts welfare, training and employment program*. Cambridge, Mass.: Harvard Univ. Press.
- Behn, R. (2003). Why measure performance? *Public Administration Review*, 63(5), 586–607.
- Berman, E., & Wang, X. (2000). Performance measurement in U.S. counties: Capacity for reform. *Public Administration Review*, 60(5), 409–420.
- Bevan, G., & Hood, C. (2006). What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, 84(3), 517–38.
- Bouckaert, G., & Halligan, J. (2008). *Managing performance: International comparisons*. London: Routledge.
- Bourdeaux, C., & Chikoto, G. (2008). Legislative influences on performance management reform. *Public Administration Review*, 68(2), 253–65.
- Boyne, G., Gould-Williams, J., Law, J., & Walker, R. (2004). Toward the self-evaluating organization? An empirical test of the Wildavsky model. *Public Administration Review*, 64(4), 463–73.
- Boyne, G., & Walker, R. (2010). Strategic management and public service performance: The way ahead. *Public Administration Review* 70(s1): s185–92.
- Bozeman, B., & Bretschneider, S. (1986). Public management information systems: Theory and prescription. *Public Administration Review* 46 (Special Issue: Public Management Information Systems): pp. 475–87.
- Bretschneider, S. (1990): Management information systems in public and private organizations: An empirical test. *Public Administration Review*, 50(5), 536–45.
- Broadbent, J., & Laughlin, R. (2009): Performance management systems: A conceptual model. *Management Accounting Research*, 20, 283-95.
- Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach*. Los Angeles: Sage.
- de Lancer Julnes, P., Berry F., Aristigueta, M., & Yang, K. (Eds.). (2007). *International handbook of practice-based performance management*. Los Angeles, Calif.: Sage.

- de Lancer Julnes, P., & Holzer, M. (2001). Promoting the utilization of performance measures in public organizations: An empirical study of factors affecting adoption and implementation. *Public Administration Review*, 61(6), 693–708.
- Deming, W. (1994). *The new economics for industry, government, education*. Cambridge, MA: Massachusetts Institute of Technology.
- Donaldson, L. (2001): *The contingency theory of organizations*. Thousand Oaks, CA: Sage.
- Dull, M. (2009). Results-model reform leadership: Questions of credible commitment. *Journal of Public Administration Research & Theory*, 19(2), 255–84.
- Folz, D., Abdelrazek, R., & Chung, Y. (2009). The adoption, use, and impacts of performance measures in medium-size cities. *Public Performance & Management Review*, 33(1), 63–87.
- Grant, A., & Wall, T. (2009). The Neglected science and art of quasi-experimentation: Why-to, when-to, and how-to advice for organizational researchers. *Organizational Research Methods* 12(4), 653–86.
- Hatry, H. (2006). *Performance measurement: Getting results*. Washington, D.C: Urban Institute Press.
- Ho, A. (2006). Accounting for the value of performance measurement from the perspective of Midwestern mayors. *Journal of Public Administration Research and Theory*, 16(2), 217–37.
- Ho, A. (2008). Reporting public performance information: The promise and challenges of citizen involvement. W. Van Dooren and S. Van de Walle (Eds.). *Performance information in the public sector. How it is used* (pp. 192-210). Hampshire, UK: Palgrave MacMillan.
- Hood, C. (2006). Gaming in targetworld: The targets approach to managing British public services. *Public Administration Review*, 66(4), 515–21.
- Hvidman, U., & Andersen, S. 2013. The impact of performance management in public and private organizations. Forthcoming in *Journal of Public Administration Research and Theory*, doi:10.1093/jopart/mut019.
- James, O. (2011). Performance measures and democracy: Information effects on citizens in field and laboratory experiments. *Journal of Public Administration Research and Theory*, 21(3), 399–418.
- Johansson, T., & Siverbo, S. (2009). Explaining the utilization of relative performance evaluation in local government: A multi-theoretical study using data from Sweden. *Financial Accountability & Management*, 25(2), 197–224.
- Kelman, S., & Friedman, J. (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) 4, 917–46.
- Kolb, A., & Kolb, D. (2009). Experiential learning theory: A dynamic, holistic approach to management learning, education and development. S. Armstrong & C. Fukami (Eds.). *The SAGE handbook of management learning, education and development* (pp. 42–68). Los Angeles: SAGE.
- Kroll, A. (2013). The other type of performance information: Non-routine feedback, its relevance and use. *Public Administration Review*, 73(2), 265–276.

- Kroll, A., & Proeller, I. (2013). Controlling the control system: Performance information in the German childcare administration. *International Journal of Public Sector Management*, 26 (1), 74–85.
- Kroll, A., & Vogel, D. (2013). The PSM-leadership fit: A model of performance information use. Forthcoming in *Public Administration*.
- Levitt, B., & March, J. (1988). Organizational learning. *Annual Review of Sociology* 14: 319–338.
- Marchand, D., Kettinger W., & Rollins, J. (2002). *Information orientation: The link to business performance*. Oxford: Oxford Univ. Press.
- Melkers, J., & Willoughby, K. (2005). Models of performance-measurement use in local governments. *Public Administration Review*, 65(2), 180–90.
- Merton, R. (1968). *Social theory and social structure*. New York, NY: Free Press.
- Meyer, R., & Hammerschmid, G. (2006). Changing institutional logics and executive identities: A managerial challenge to public administration in Austria. *American Behavioral Scientist*, 49(7), 1000–14.
- Moynihan, D. (2005). Goal-based learning and the future of performance management. *Public Administration Review*, 65(2), 203–16.
- Moynihan, D. (2008). *The dynamics of performance management: Constructing information and reform*. Washington, D.C: Georgetown Univ. Press.
- Moynihan, D. (2009). Through a glass, darkly: Understanding the effects of performance regimes. *Public Performance & Management Review*, 32(4), 592–603.
- Moynihan, D., & Hawes, D. (2012). Responsiveness to reform values: The influence of the environment on performance information use. *Public Administration Review*, 72(s1), 95–105.
- Moynihan, D., & Ingraham, P. (2004). Integrative leadership in the public sector: A model of performance-information use. *Administration & Society*, 36(4), 427–53.
- Moynihan, D., & Landuyt, N. (2009). How do public organizations learn? Bridging cultural and structural perspectives. *Public Administration Review*, 69(6), 1097–105.
- Moynihan, D., & Lavertu. (2012). Does involvement in performance management routines encourage performance information use? Evaluating GPRA and PART. *Public Administration Review*, 72(4), 592–602.
- Moynihan, D., & Pandey, S. (2010). The big question for performance management: Why do managers use performance information? *Journal of Public Administration Research and Theory*, 20(4), 849–66.
- Moynihan, D., Pandey, S., & Wright, B. (2012a). Prosocial values and performance management theory: The link between perceived social impact and performance information use. *Governance*, 25(3), 463–83.
- Moynihan, D., Pandey, S., & Wright, B. (2012b). Setting the table: How transformational leadership fosters performance information use. *Journal of Public Administration Research and Theory*, 22(1), 143–64.
- Nielsen, P. (2013). Performance management, managerial authority, and public service performance. Forthcoming in *Journal of Public Administration Research and Theory*, doi:10.1093/jopart/mut025.

- Nielsen, P., & Baekgaard, M. (2013). Performance information, blame avoidance, and politicians' attitudes to spending and reform: Evidence from an experiment. *Journal of Public Administration Research and Theory*, doi:10.1093/jopart/mut051.
- OECD [Organization for Economic Co-operation and Development]. (2007). *Performance budgeting in OECD countries*. Paris: OECD.
- Otley, D. (1980). The contingency theory of management accounting: Achievement and prognosis. *Accounting, Organization, and Society*, 5(4), 413–28.
- Pinder, C., & Moore, L. (1979). The resurrection of taxonomy to aid the development of middle range theories of organizational behavior. *Administrative Science Quarterly*, 24(1), 99–118
- Poister, T., Pasha, O., & Edwards, L. (2013). Does performance management lead to better outcomes? Evidence from the U.S. public transit industry. *Public Administration Review* 73(4) 625–36.
- Pollitt, C. (2013, Ed.). *Context in public policy and management: The missing link?* Cheltenham, UK: Edward Elgar.
- Pollitt, C., & Bouckaert, G. (2011). *Public management reform: A comparative analysis*. Oxford: Oxford Univ. Press.
- Pratchett, L., & Wingfield, M. (1996). Petty bureaucracy and woollyminded liberalisms? The changing ethos of local government officers. *Public Administration* 74(4), 639–656.
- Riding, R., & Cheema, I. (1991). Cognitive Styles: An overview and integration. *Educational Psychology*, 11(3), 193–215.
- Rosenthal, R. (1979). The "File Drawer Problem" and Tolerance for Null Results. *Psychological Bulletin*, 86(3), 638–641.
- Rosenthal, R., & DiMatteo, M. (2001). Meta-Analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, 52, 59–82.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston, Mass.: Houghton Mifflin.
- March, J., Guetzkow H., & Simon, H. (1958). *Organizations*. New York: Wiley.
- Sun, R., & Van Ryzin, G. (2013). Are performance management practices associated with better public outcomes? Empirical evidence from New York public schools. *American Review of Public Administration*, doi: 10.1177/0275074012468058
- Sutton, A. (2009). Publication Bias. In H. Cooper, L. Hedges, & J. valentine (Eds.). *The handbook of research synthesis and meta-analysis* (pp. 435–52). New York, NY: Russell Sage Foundation.
- Taylor, J. (2009). Strengthening the link between performance measurement and decision making. *Public Administration*, 87(4), 853–71.
- Taylor, J. (2011). Factors influencing the use of performance information for decision making in Australian state agencies. *Public Administration*, 89(4), 1316–34.
- ter Bogt, H. (2004). Politicians in search of performance information? Survey research on Dutch aldermen's use of performance information. *Financial Accountability and Management*, 20(3), 221–52.
- Togerson, C. (2006) Publication bias: The Achilles' heel of systematic reviews? *British Journal of Educational Studies*, 54(1), 89–102.

- Van de Walle, S., & Bovaird, T. (2007). *Making better use of information to drive improvement in local public services: A report for the Audit Commission*. Project Report. INLOGOV.
- Van Dooren, W. (2004). Supply and demand of policy indicators. *Public Management Review*, 6(4), 511–30.
- Van Dooren, W., Bouckaert, G., & Halligan, J. (2010). *Performance management in the public sector*. London: Routledge.
- Van Dooren, W., & Van de Walle, S. (Eds.) (2008). *Performance information in the public sector. How it is used*. Hampshire: Palgrave Macmillan.
- Yang, K., & Hsieh, J. (2007). Managerial effectiveness of government performance measurement: Testing a middle-range model. *Public Administration Review*, 67(5), 861–79.