

Performance in Public Organizations Clarifying the Conceptual Space

Andersen, Lotte Bøgh; Boesen, Andreas; Holm Pedersen, Lene

Document Version

Accepted author manuscript

Published in:

Public Administration Review

DOI:

[10.1111/puar.12578](https://doi.org/10.1111/puar.12578)

Publication date:

2016

License

Unspecified

Citation for published version (APA):

Andersen, L. B., Boesen, A., & Holm Pedersen, L. (2016). Performance in Public Organizations: Clarifying the Conceptual Space. *Public Administration Review*, 76(6), 852–862. <https://doi.org/10.1111/puar.12578>

[Link to publication in CBS Research Portal](#)

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 recognise and abide by the legal requirements associated with these rights.

Take down policy

If you believe that this document breaches copyright please contact us (research.lib@cbs.dk) providing details, and we will remove access to the work immediately and investigate your claim.

Download date: 04. Jul. 2025



Performance in Public Organizations: Clarifying the Conceptual Space

Lotte Bøgh Andersen, Andreas Boesen, and Lene Holm Pedersen

Journal article (Accepted version)

This is the peer reviewed version of the following article: Performance in Public Organizations : Clarifying the Conceptual Space. / Andersen, Lotte Bøgh; Boesen, Andreas; Holm Pedersen, Lene. In: *Public Administration Review*, Vol. 76, No. 6, 2016, p. 852-862, which has been published in final form at <http://dx.doi.org/10.1111/puar.12578>.

This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#).

Uploaded to [Research@CBS](#): July 2017

Performance in Public Organizations: Clarifying the Conceptual Space

Lotte Bøgh Andersen

Aarhus University and KORA – Danish Institute for Regional and Local Government
Research

Bartholins Allé 7, 8000 Aarhus C., Denmark
(lotte@ps.au.dk)

Andreas Boesen

Aarhus University (aboesen@ps.au.dk)

Lene Holm Pedersen

Copenhagen Business School and KORA – Danish Institute for Regional and Local
Government Research (lhp.dbp@cbs.dk)

Abstract

Performance in public organizations is a key concept requiring clarification. Based on a conceptual review of articles in ten public administration journals, this article proposes six distinctions to describe the systematic differences between performance criteria: From which stakeholder's perspective is performance being assessed? Are the criteria formal or informal? Are the criteria subjective? Which process focus and product focus do they have (if any)? What is the unit of analysis? Based on these distinctions, we classify the performance criteria of existing studies used in an empirical review of management and performance. Our results illustrate how a systematization of the conceptual space of performance in public organizations can help researchers select what to study and what to leave out with greater accuracy while also bringing greater clarity to public debates about performance.

BIO

Lotte Bøgh Andersen (lotte@ps.au.dk) is professor at Aarhus University & Danish Institute for Local and Regional Government Research. Her research interests focus on leadership, administration and management in public organizations, especially motivation and performance of public employees, leadership strategies, professional norms and economic incentives. Right now, she is leading a field experiment (www.leap-project.dk) which investigates 500 public and private leaders to find out how transformational and transactional leadership affects employee motivation and organizational performance.

Andreas Boesen (aboesen@ps.au.dk) is a master student at Department of Political Science, Aarhus University, Denmark.

Lene Holm Pedersen (lhp.dbp@cbs.dk) is professor in public management at Copenhagen Business School (CBS) & Danish Institute for Local and Regional Government Research. She conducts research on steering, motivation and performance in the public sector. She has also published a number of articles on the political power of regulatory ideas, organizational change and reform, local government and democracy, and on how local governments are influenced by fiscal austerity.

Points for practitioners:

It is equally important to systematize that which is included in and left out of performance assessments. It is normally not possible or even desirable to include all possible aspects of performance in an assessment, but it is very useful to know explicitly which types of performance you measure—and which you do not.

Many performance criteria focus only on what is achieved in public organizations, but it can also be very relevant to include process aspects of performance such as equal access to services.

It is important to be aware of whose performance you are analyzing: Individuals, teams, organizations or other units of analyses. Performance can be conceptualized at many different levels in public organizations, and if teamwork is necessary for success, it might not be meaningful to look at individual performance.

In recent decades, both the quantity and quality of the empirical contributions to the performance literature have increased (Andrews, Boyne, and Walker 2006a, 14; 2006b, 14; Bommer 1995; Brewer 2006; Meier and O'Toole 2013; Meier et al. 2015; O'Toole and Meier 2013; Walker and Andrews, 2015). This is important to the field, as performance is an important—possibly *the* most important—concept in public administration. According to Rainey, “virtually all of management and organization theory concerns performance and effectiveness, at least implicitly” (1997, 125). Very broadly defined, performance is the actual achievement of a unit relative to its intended achievements, such as the attainment of goals and objectives (Jung 2011, 195).

However, the performance concept itself is not clear enough. In order to take the study of performance one step further and bring greater clarity to public debates about what performance is, the conceptual space of performance in public organizations must be clarified. Hirsh and Levin (1999: 208) prophesized 25 years ago that the performance concept would go through a life-cycle whereby it would fall by the wayside or narrow in scope once better applications were required. They anticipated that critical reviewers would question the validity of the concept at a later point in its life-cycle, and the strong demands for better research designs and applications indicate that they were right (e.g., Andrews, Boyne and Walker 2006a and Meier and O'Toole, 2013). However, there is tension between rigor and relevance (Hirsch and Levin, 1999). Especially in studies of public organization performance, focusing on one performance criterion without discussing what is excluded is problematic, as most public organizations have multiple, potentially conflicting goals. Public administration studies of performance that do not explicate their partial perspective will paint a biased picture of performance as a whole.

We argue that a valid conceptual space of performance can improve our ability to create research designs that fill important gaps in the literature and provide practitioners

with a tool to systematize that which is included in and left out of performance assessments. Thus, our research questions are: What are the central distinctions in the conceptual space of performance in public organizations? How can these distinctions be defined and understood? And how can they inform our understanding of existing studies of performance in public organizations?

The article contributes by building a valid conceptual space of performance consisting of six distinctions according to which performance criteria can be classified based on a conceptual review of the literature. A *performance criterion* is a theoretically defined standard based on achievements that can be evaluated. Effectiveness is an example of a specific performance criterion; in the case of a school, it could be the extent to which formal learning standards are reached. In contrast, a *conceptual space* consists of one or more *distinctions* representing various qualities or characteristics of criteria corresponding to how these criteria are judged similar or different. Which stakeholder has, for example, decided that effectiveness should be a performance criterion for the organization in question? And is the criterion formal or informal?

We test the conceptual space in an empirical re-classification of 110 performance criteria used in a recent high-quality review of the relationship between local government management and performance (Walker and Andrews 2015). This re-classification shows that the conceptualization is useful. The six distinctions provide oversight of the central differences and similarities between the performance criteria employed in the literature and facilitate the evaluation of the studies made. A conceptualization of performance based on distinctions and criteria facilitates a more explicit discussion of the relation between the performance concept and the specific performance criteria used in public administration studies. Analyzing different performance criteria in different studies is not in itself problematic but requires greater conceptual clarity as well as a framework that allows us to

compare different performance criteria. As such, the conceptual space will allow us to understand and use specific performance criteria better, and a stronger conceptual framework can make performance research less data-driven and sharpen the potential for theory building.

The article is structured in five sections. The next section discusses the relationship between the general performance concept and the specific concepts used in individual articles and books, arguing that a systematized conceptual space clarifies the relationship between these two conceptual levels. We then explain the methodology used to develop our proposed clarification of the conceptual space for performance and discuss the six identified distinctions. This is followed by a categorization of 110 performance criteria from 66 existing performance studies, which exemplifies how the distinctions can inform our understanding of performance in public organizations. In the final section, we discuss the benefits of a clear conceptual space for future research and practice.

Conceptualizing Performance

The literature has begun to develop the conceptual space for performance in public organizations and important contributions have been made, even if no final consensus exists. The 3E model focuses on economy (production costs for a given quality), efficiency (cost per unit of output), and effectiveness (achievement of formal objectives) (Boyne 2002; 2003a; 2003b). Similarly, the IOO model examines the sequence of inputs, outputs, and outcomes, where inputs are comparable with economy in the 3E model. Outputs concern the actions performed in the production process and include both quantity and quality. The ratio of outputs to inputs is defined as efficiency. Outcome concerns the changes in external units (e.g., individuals or organizations), which are the object or target of the relevant policy or service intervention (Behn 2014, 137; Lynn and Robichau 2013, 208). It can include

effectiveness as well as impact and equity in outcomes. The ratio of outcomes to inputs is defined as cost-effectiveness or “value for money” (Walker and Andrews., 2015).

Distinguishing between these different elements in the production process is useful and will also be part of the proposed conceptual space in this article, but many aspects of performance in public organizations are not included in the 3E and IOO models, which draw on the management of private firms. First, aspects related to the process are not included (e.g., accountability, probity, participation, due process), and these aspects can be very important in public organizations (Walker, Boyne, and Brewer 2010; Moynihan et al. 2011). Second, identifying the stakeholders is central in defining performance, because it is otherwise unclear who has the legitimate right to define what good performance is (Andrews, Boyne, and Walker 2006b: 29; Amirkhanyan, Kim, and Lambright 2014). Third, ambiguity about the relevant units of analysis can hinder comparisons of different performance studies. While the performance analysis of individual surgeons and hospital wards contribute to our understanding of public organization performance, it is necessary to be explicit about the level on which performance is conceptualized. Responding to the need for a coherent conceptualization of performance in the public sector, we clarify and refine the concept in the following according to the model proposed by Adcock and Collier (2001).

The relationship between concepts and observations can be illustrated as consisting of at least three levels, as shown in figure 1. The performance concept applied in a given study is based on the general performance concept in the relevant literature (also called the background concept). A specific concept can contain one or more criteria, depending on the complexity included in the relevant research. The same is the case for the indicators (also called measures).

[FIGURE 1 here]

Conceptualization is developed through the downward and upward movement in figure 1. On the left side, the specific concept used in a study should be based on the background concept. Here, a central choice relates to the number of criteria included. Background concepts often include a variety of meanings (this holds especially true for performance), meaning that it is normally necessary to prioritize when the specific concepts are formulated.

Adcock and Collier (2001) warn researchers against three common traps when establishing their specific concept, and we argue that a clear conceptual space can help avoid all three. First, the flexibility inherent in the choice of concept does not suggest that anything goes. In most literatures, scholars commonly associate a matrix of potential meanings with the background concept, limiting the range of plausible options, and the researcher who strays outside it risks being dismissed or misunderstood. A clear conceptual space makes this explicit. Although the performance of public organizations is inescapably contestable (Boyne et al. 2006: 6), clarifying the concept provides a vocabulary for discussing what performance is and what it is not.

Second, scholars should refrain from claiming that their choice of a specific concept is the only possible choice. A clear conceptual space facilitates cumulative research if scholars recognize that it is fruitful to emphasize different aspects of a background concept when developing specific concepts. Rather than make sweeping claims about what performance in public organizations “really” means, scholars should present specific arguments linked to the goals and context of their research that justify their choices.

The third problem occurs when scholars stop short of accounting in detail for their specific concept. Adcock and Collier argue that scholars should specify the meaning and entailments of their specific concept in relation to the background concept. If the translation from background to specific concept in performance studies included a more systematic

discussion of what is included and what is left out, it would be easier to make a comprehensive review of the findings in the performance literature.

Methodology for Reaching the Proposed Conceptual Clarification

Building on existing conceptual discussions in public administration books and journal articles, we perform a conceptual review of the existing understandings of the performance concept. Our aim is to identify the distinctions that are relevant to conceptualize and include in the overall concept (Petticrew and Roberts 2006, 39) rather than provide an overview of all of the empirical results produced in this field.

The identification of eligible studies was based on two principles. First, the review was delimited by articles published in top peer-reviewed journals in public administration. We included 10 journals in line with the selections in recent reviews of public administration concepts (e.g., Kuipers et al. 2014; Tummers et al. 2015; Voorberg, Bekkers, and Tummers 2014). Selecting the top journals should make it unlikely that we have omitted important trends in the existing discussions, and we did not find additional distinctions for the last many articles we conceptually reviewed. This indicates that a sufficient number of journals have been reviewed to identify the relevant distinctions. The second principle was the specific usage of the terms “performance” and “public,” but not synonyms such as “effectiveness” and “results.” As implied by the predictions made by Hirsh and Levin (1999) about the rise and fall of the performance concept, there are several related concepts, but analyzing the relationship between the performance concept and other relevant concepts would be another task entirely. Given that we aim to identify the state of the art in the conceptualization of performance, articles that focus empirically on the effect of specific explanatory variables were excluded unless they included a conceptualization of performance. As discussed below, studies were identified according to the steps depicted in

figure 2, and we reached a saturated conceptualization where no more dimensions were identified.

[Insert figure 2 here]

In the first step, a search was conducted in Scopus for articles published before April 15, 2014, with “performance” and “public” in the title, abstract, or keywords in 10 journals: *Public Administration Review* (PAR—190 results), *Journal of Public Administration Research and Theory* (JPART—130 results), *Public Administration* (PA—97 results), *American Review of Public Administration* (ARPA—50 results), *International Public Management Journal* (IPMJ—26 results), *Administration and Society* (A&S—35 results), *International Review of Administrative Science* (IRAS—103 results), *International Journal of Public Administration* (IIPA—64 results), *Public Management Review* (PMR—74 results), and *Review of Public Personnel Administration* (ROPPEA—25 results). In the second step, title, abstracts, and keywords in step 1 articles (n = 794) were screened, and 635 studies (step 3) were excluded, because it was clear from the title, abstract, and/or keywords that they did not conceptualize the performance in public organizations. The remaining 159 studies were screened on the specific eligibility principles via the title, abstract, and text (step 4), and 98 studies were then excluded in step 5, because they did not contribute with conceptual performance discussions. The remaining 61 articles were included in the conceptual review (n = 61), 28 of which are referred to in the final article. Additionally, snowball sampling based on the reading of the 159 studies identified in step 4 led to the identification of relevant studies, which were screened and included if relevant. Furthermore, references identified through peer discussions were included. This additional process (step 7) identified 27 references, nine of which are referred to in the final article. The process was similar for books (step 8).

This literature forms the basis of the six distinctions discussed below. Ideally, the distinctions should be comprehensive, mutually exclusive, and systematized in order to capture the most important aspects of how performance is conceptualized in public administration literature.

The Distinctions

Starting with the distinction consistently regarded as most important, namely who decides what performance is, this section discusses stakeholders, formality, subjectivity, process focus, product focus, and finally the unit of analysis for performance criteria in public administration research.

Stakeholders

The central question behind the first distinction is who actually decides what good performance is. It is broadly accepted that different stakeholders, inside and outside public organizations, hold diverse views on what it takes to perform well (Boyne 2003a, 368; Smith and Larimer 2004; Walker, Boyne, and Brewer 2010; Yang 2009). However, the distinction between different stakeholders' performance criteria is not always explicitly defined, and it is seldom discussed who ultimately has the authority to determine what good performance is.

Stakeholders can be defined as “any group or individual who can affect or is affected by the achievement of the organization's objectives” (Freeman [1984] 2010, 46). Thus, performance criteria and their relative weight vary across stakeholders, and performance criteria are not technical and universal, but rather politically constructed (Boyne 2003b, 224). The performance literature often points out that giving one stakeholder priority may have distributive consequences (Andrews et al. 2012; Boyne 2003a; Moynihan et al. 2011), but the identity of prioritized stakeholder is seldom explicitly mentioned in comparisons of

performance studies. Analyzing stakeholder identity as a distinction in the conceptual space of performance allows us to compare which stakeholders are taken into account and which are left out.

The authority to define good performance differs in public and private organizations. While private organizations often have close ties to their owners and customers, public organizations have stronger relations with political and government authorities and are more exposed to political influence (Mintzberg 1973). Boschken (1992) distinguishes between performance indexes according to the source of authority. For organizations resting on political authority, politicians formulate the goals and objectives, and the relevant performance question is whether the organization meets its public mandate, in which case it is performing well. Boschken (1992, 272) labels this “social effectiveness.” Here, performance is about attaining democratically stated goals. For the market, the source of authority is economic, and the central question for the organization is: Do our customers like us? (Boschken 1992). Even when the introduction of market-based principles has created “customers” in public service industries, the public mandate often remains crucial to the legitimate claim on authority, and public organizations rarely operate according to ideal-typical market principles as the financial resources depend on political authority.

A third unit of authority in addition to the market and state is the clan (Ouchi 1980). The clan is a group that may or may not be linked by kinship ties but is based on common, internalized goals and strong feelings of kinship (Scott 1995, 252). Clan systems are characterized by implicit, internalized control mechanisms, non-specialized roles and career paths, holistic rather than segmented concerns, and long-term employment (Ouchi 1980). The socialization process effectively eliminates goal incongruence between individuals—that is, among the members of the clan—while incompatibility can remain between the clan and outsiders or their managers (Behn 2014).

Professions are often seen as the prime examples of clans due to the existence of inter-occupational norms and specialized theoretical knowledge (Ouchi 1980). Accordingly, a profession is a unit of authority that has standards for defining good performance, and this authority rests on expert knowledge and to some extent a monopoly on the delivery of specialized services (Andersen 2005, 205; Collins 1990; Friedson 2001). In some types of service delivery, professions hold the authority to define good performance, but this may be challenged if other stakeholders find that professions fail to deliver high quality services. Professions can also have the power to influence the opinions of others based on recognized expertise and knowledge. In addition to public service professions, such as teachers and nurses, researchers from the public administration scholarly community can also have this type of clan power and become the arbiters of performance in public organizations by virtue of their expertise and the choices they make in their research. In contrast, the logic of the hierarchical model as implemented in representative democracies is that elected politicians have the power to make decisions and enforce obedience based on recognized democratic legitimacy.

In the literature on performance in public organizations, the stakeholder model is pluralist, favoring the worldview that there are multiple legitimate interests surrounding an organization. The performance literature also tends to emphasize a more polycentric relationship with multiple principals (Boschken 1992; Boyne 2003b; Moynihan et al. 2011; Talbot 2008, 1579) rather than a simplistic, vertical principal–agent relationship in the public sector. Stakeholders can, for example, be customers, clients, citizens, professional groups, organizations in the environment, governance networks, and international institutions (Stoney and Winstanley 2001). These various stakeholders possibly steer towards specific forms of performance in public organizations, including conformance to internal process standards

(Talbot 2008, 1580). In this complex institutional environment, it is relevant for analysts to ask who holds the authority to determine that a given criterion represents performance.

Formality

Research on organizational performance tends to see organizations as established to achieve specific objectives (Boyne 2003b, 214). The formality of these objectives may vary, however. Some public organizations do not have formal goals that are clearly expressed in legislation or other official documents, and formal goals (should they exist) can be broad mission statements rather than concrete objectives. Boyne (2003b) argues that a combination of the multiple stakeholder model as discussed above and the goal model is a positive basis for defining performance in public organizations. The key argument of the goal model is that the extent to which formal goals are attained should be used to assess performance. Although this steers us free of seeing performance as consisting only of the subjective impressions of powerful stakeholders (Boyne 2003b), it remains important to recognize that performance criteria can vary in formality.

The central question in the distinction concerning formality is the degree to which performance is formally or informally defined. Goals that are formalized in the sense that they are written, predefined, and/or made explicit are more formal than implicit expectations regarding the achievement of a unit. Compliance with professional norms, for example, can be central in public organizations, and these norms may be highly informal, even tacit. Politicians may also formally decide that satisfying a group's preferences is the goal, making user satisfaction a formal performance criterion. It would still, however, be an inherently subjective performance criterion, as discussed below.

Inherent Subjectivity

Subjectivity has been discussed at length with regard to the operationalization of performance measures, but performance criteria may conceptually also have varying degrees of inherent subjectivity. The central question behind this distinction is the degree to which a performance criterion concerns interior experiences and perceptions versus exterior, observable phenomena. Some criteria, such as trust in civil servants (Van Ryzin 2011), are inherently subjective. Boyne (2003b) mentions consumer satisfaction as either a criterion in itself or a proxy for another type of performance. As the former, it can be classified as inherently subjective, because satisfaction is an internal feeling. An example of a criterion related to exterior and observable phenomena is the physical health of hospital patients.

This is a conceptual distinction, but performance can also be measured more or less subjectively according to whether it concerns fixed criteria of performance, whether there is a process for verifying the accuracy of the measure, and whether there is external verification (Andrews, Boyne, and Walker 2006b). For instance, parent satisfaction with schools is an inherently subjective performance criterion and almost always measured subjectively compared to student stress, which is a school performance criterion with objective manifestations. Stress can be assessed in a relatively objective manner by testing cortisol levels using saliva, but it can also be measured subjectively by simply asking the students. Correspondingly, survey measures are not the only way to assess parent satisfaction; (objective) exit behavior from the school also serves as a measure.

Distinguishing between performance criteria with different subjectivity highlights the relevance of first discussing the distinction between degrees of subjectivity at the conceptual level and subsequently making decisions regarding measurement. Research that postulates that objective measures of performance are always preferable to subjective measures ignores conceptual differences between performance criteria, whereas our discussion of the performance concept implies that the choice of indicator depends on the conceptual

understanding of the relevant performance criterion. Similarly, it is difficult to make an *a priori* decision in favor of either inherently subjective or inherently objective performance criteria, because doing so depends on the understanding of performance that the relevant stakeholders find important.

Type of Process Focus

Performance in public organizations involves more than services and decisions; it also includes the manner in which the service is delivered. In the context of the private sector, Grönroos (1984, 39) argues that both product and process are relevant for customers. How a good or service is made functionally (the process) is often important to consumers and how they view what they receive (the product—which can also be a service, such as a haircut). In the public sector, this argument is even more valid because the specific user is rarely the sole benefactor of the service, and Van Ryzin (2011) emphasizes that some specific governmental processes matter a great deal to citizens. Moynihan et al. (2011, 143) also criticize that performance measurement systems displace attention from the democratic values linked to the working processes of the administration, such as due process, equity, integrity, and transparency, because these systems only focus on product criteria related to mission achievement and effectiveness and disregard process criteria related to traditional democratic values. Performance criteria can focus on process as well as product, but the distinction is not between means and ends (where means per definition is not performance), because democracy is a goal in itself in a democratic state; hence, administration and service delivery that respect democratic values is also a goal.

Type of Product Focus

Boyne (2002) establishes an influential classification of performance criteria by combining the 3E (economy–efficiency–effectiveness) and IOO (inputs–outputs–outcomes) models with two additional criteria (responsiveness and democratic outcomes). After omitting the redundant elements of the two models, Boyne (2002, 19) claims to have arrived at a comprehensive checklist. He distinguishes outputs, efficiency, service outcomes, responsiveness, and democratic outcomes and specifies 15 performance criteria. The problem with combining two different models and including responsiveness and democratic outcomes in an ad hoc manner is that it becomes unclear how the criteria differ.

We argue that the distinction in the original IOO model (i.e., production phase) can be useful to classify what type of product focus (if any) performance criteria have: input, output, outcome, output per input, or outcome per input. Outcome performance criteria such as effectiveness are normally closest to the achievement of defined objectives in public organizations, while input-related criteria are rarely relevant for performance in this understanding of the concept. Although outcome criteria are often seen as the gold standard, Lynn and Robichau (2013, 208) argue that they can be problematic in public organizations for two reasons. First, they may be imposed on organizations by external stakeholders and might therefore be incompletely embraced by organizational actors. Second, since public organizations are often unable to control all of the factors affecting outcomes (Ashworth, Boyne, and Entwistle 2010; Behn 2014; Talbot 2010), outcome criteria are poor reflections of the actual organizational effort.

Concerning the second reason, Boyne (2003b, 218) discusses whether the difficulty in attributing outcomes exclusively to the work of service providers may imply that the achievement of professional standards can be used as performance criteria. He argues that the evidence of following correct procedures or doing things correctly might be preferable to outcome information. This links back to our discussion of the relevant source of authority. If

the profession is seen as a relevant source, then Boyne clearly has a point and his argument in favor of the usefulness of the achievement of professional standards can also be valid and useful if it is impossible to assess the achievement of outcome-related goals (given that evidence indicates that compliance with professional standards leads to these goals).

Unit of Analysis

Cameron (1986, 542) argues that a number of questions must be answered prior to assessing effectiveness. The same goes for performance. The first question concerns the perspective from which effectiveness is being assessed, which corresponds to the stakeholder distinction discussed above. The next two questions concern the unit of analysis. What domain of activity is being assessed, and what is the level of analysis? These questions are about the identity of the unit whose achievements we analyze. Who performs: individuals (Andersen, Heinesen, and Pedersen 2014), groups (Van Ryzin 2011), organizations (Angle and Perry 1981; Caillier 2011; Kim 2005, 245), or programs (Collins and Gerber 2008; Jung 2014)? Even when we analyze individuals, it might be relevant to consider performance at a higher level, because the extra-role performance of individuals captures their contribution to broader organizational goals, whereas in-role performance is narrowly related to the individual's own work (Van Loon, Vandenabeele, and Leisink 2015; Williams and Anderson 1991). Although individuals are the agents behind extra-role performance, the organization might be the relevant level to analyze performance effects. For teachers, an individual in-role performance criterion could be their students' academic skills in their subject controlled for performance in other subjects. Extra-role performance is typically more difficult to capture. For teachers, the academic skills of students could be relevant, but this would also be a consequence of efforts from other school employees and factors outside of the school, such as student background.

Analyses of performance carried out at different analytical levels may speak to each other, but they are not directly cumulative. For instance, the organizational performance investigated at the school level (Andersen and Mortensen 2010) and individual performance investigated for each teacher (Andersen, Heinesen, and Pedersen 2014) represent fundamentally different performance criteria, although knowledge of the factors that increase individual performance is crucial in order to improve school performance.

The distinction between different analytical levels calls for a conscious choice of the unit of analysis in research designs. The consequences of individual factors, such as public service motivation, job satisfaction, and intrinsic motivation, may be better investigated in terms of the performance of individual employees, as much information is lost if individual scores are aggregated to the organizational level. In contrast, it may be less relevant to investigate individual-level performance in organizations such as hospitals, where production is the result of teamwork involving different professions and organizational units. Linking the survival of cancer patients to individual surgeons makes less sense than using students' math skills to assess the performance of their math teacher.

The fact that performance exists at different levels does not mean that all levels should necessarily be analyzed, but rather that we must prioritize between different units of analysis and take the explanatory variables into account when deciding the relevant analytical level. If explanatory variables with a causal impact on performance are conceptualized at a given analytical level, it speaks in favor of also conceptualizing performance at this level.

The unit of analysis is the last distinction. Table 1 summarizes our discussion; it lists the key question for each distinction and provides examples of how performance criteria can vary based on the distinctions.

[TABLE 1 HERE]

Compared to Boyne's conceptualizations of performance (Boyne 2002), our conceptual space explicates the distinctions between the criteria. As mentioned, Boyne's classification contains five main dimensions (outputs, efficiency, service outcomes, responsiveness, and democratic outcomes) and 15 sub dimensions. In our terminology, Boyne's dimensions can be seen as criteria, which can be classified in relation to one or more of our distinctions. For instance, output and service outcomes are different product focuses, while different types of democratic outcomes would be classified differently using our distinctions. While Boyne focuses on developing the criteria ("dimensions" in his terminology), we focus on the differences between the criteria. It is possible to map out upon which distinctions the criteria are placed, and the two conceptualizations thus supplement each other in building the overall conceptual space.

Using the Distinctions to Classify Specific Performance Criteria

Deriving and discussing the six distinctions is merely the first step. It is also central to test the usefulness of the distinctions on empirical studies of performance in public organizations. The main results from this empirical review are summarized in the following.

As shown in detail in the online appendix, we have classified the 110 performance criteria used in the Walker and Andrews (2015) review of local government management and performance. The advantage of choosing this material rather than the articles used in the conceptual review is that the six distinctions are not tested in the studies, which were used to derive them. That increases the external validity of our illustration. The criteria for being included in Walker and Andrews' (2015: 108) analysis were: (1) The articles appeared in an Anglophone public administration journal listed in the Web of Science Social Sciences Citation Index between 1970 to 2012; (2) the dependent variable was

performance, (3) the unit of analysis was an organization or part thereof, and (4) the statistical results were presented in a form that allowed them to be included in a meta-analysis. Except that we do not focus on only organizations as units of analysis, Walker and Andrew's selection of performance criteria is thus very well fitted to illustrate the usefulness of our framework. Our classification of the 110 performance criteria reveals that all six distinctions are relevant. The descriptions in the article of their performance criteria made it possible to classify almost all of them on all six distinctions. Some articles did not explicitly mention the information needed to make the classification, but it is often possible to infer from the context of the articles how the performance criteria should be classified.

The classification of the empirical studies (see table A2 in the online appendix) draws attention to two major groups of performance criteria. First, research connected to Cardiff University (e.g., Andrews et al. 2005) often combines many different criteria in indexes, thus covering the performance of the investigated organizations (typically British local authorities) broadly and with central government represented by the audit commission as a key stakeholder. Second, research connected to Meier and O'Toole's research program on public management typically focuses more narrowly on specific performance criteria, often related to student academic performance and often analyzed for school districts in Texas (e.g., Meier and O'Toole 2003).

The substantial number of strong contributions from these two traditions means that many of the criteria used are very similar. Of the 110 performance criteria, 29 are proportions of students with acceptable academic competences (typically operationalized using TAAS test scores), while additional 14 criteria concern equity in proportions of students with different ethnic backgrounds with acceptable academic competences, and 13 additional criteria concern proportions of students with high quality academic competences (e.g., above 1110 on the SAT college entrance examination). In terms of the comparability of

different results, it is definitely not a problem that performance criteria in existing studies are similar, but given that we also work towards a comprehensive picture of performance in all types of public organizations, it is advisable to also apply performance criteria that differ more on the selected distinctions. Specifically, our classification highlights how very few existing performance criteria (20 out of 110) focus on process. Existing examples include equal access to services, equity in the use of disciplinary actions, and user satisfaction with the process.

This highlights the usefulness of including distinctions concerning both process and product focus. The distinction concerning product focus ensures comparability to existing classifications—illustrated by the fact that our classification on this distinction is almost similar to the Walker and Andrews (2015) classification. However, this distinction, which is the basis of many existing classifications of performance, does not capture all of the relevant performance criteria. Examples of alternative criteria that can be systematized using our distinctions are participation, probity, and responsiveness in service delivery processes. Although Walker and Andrews (2015, 104) argue that “these additional dimensions of performance are critical for scholars to understand and assess the performance,” they do not present a systematic way to conceptualize them. We hope that our inclusion of the process–focus distinction will highlight how performance can relate to process as well as product aspects.

In line with this, our classification of existing performance criteria shows that it is relevant to differentiate between different types of equity and user satisfaction. For example, equity can concern either the product, as illustrated by the Meier and O’Toole (2003) analysis of whether outcomes are equally good for students with different backgrounds, or the process, as exemplified by the Andrews et al. (2010) analysis of how fairly services are distributed amongst citizens. Similarly, it can make a difference whether

performance is conceptualized as user satisfaction with the process or as user satisfaction with the product delivered by public organizations.

Discussions about the stakeholder identity are especially exemplary in the Cardiff tradition, but their combination of different criteria (e.g., user satisfaction and effectiveness) in the same index means that there is little information about each criterion. Conversely, the parsimonious criteria in the Meier–O’Toole tradition makes it easy to classify, for example, TAAS pass rates as effectiveness. Increased awareness of this trade-off between parsimony versus broad performance conceptualizations has at least two advantages: Researchers can make more explicit choices between accuracy and completeness based on systematic knowledge about other performance studies, and it can bring greater clarity to public debates about what performance is and which performance criteria are reasonable and legitimate.

Talking about legitimacy, our classification reveals that a large majority of the performance criteria are determined by external stakeholders. Democratically, this is reassuring, especially because many of these external stakeholders are democratically elected politicians. While only five of the 110 classified performance criteria are purely internal (and thus determined by the units whose performance are evaluated), 90 criteria have external stakeholders. Three performance criteria combine external and internal stakeholders, and seven are determined by the researchers alone. There are only five criteria in relation to which it is totally implicit who the stakeholders are. Still, for 23 of the criteria classified as having external stakeholders, this could not be seen directly from the text and was therefore inferred from context information in the article. This means that research could still benefit from greater clarity. This would also make the role of the researcher clearer; while only seven criteria are explicitly researcher determined, more explicitness would facilitate a systematic

evaluation of the extent to which researchers themselves are stakeholders or important arbiters, deciding how performance should be measured in public organizations.

Our last observation is directly relevant to the ongoing debate about subjective versus objective data (Andrews et al., 2006b; 2011; Bommer, 1995; Brewer, 2006). Although only a limited proportion of the classified performance criteria (approximately 20 percent as shown in table A2) are inherently subjective on the conceptual level, we argue that being aware that subjectivity can be present at both the conceptual and operational levels will qualify the discussion about the desirability of different types of performance data. While subjectivity in measures of performance criteria which concern observable outcomes is hardly desirable, it can become necessary to use subjective measures of interior experiences and perceptions, and criteria such as a user's subjective experiences with public organizations can be very important pieces in the total performance puzzle.

Conclusion

It is essential for public administration research to be aware of how we evaluate performance in public organizations, and this article aims to identify, define, and discuss the central distinctions in order to build a conceptual space.

This article defines and discusses six central distinctions needed to build a conceptual space of performance based on a conceptual review of the literature. First, a key insight is that public organizations often have multiple stakeholders and multiple goals and that the trade-offs between them are political choices (Radin 2006). The most important distinction therefore concerns the identity of the stakeholder behind a given performance criterion.

Second, such criteria can be more or less formally stated, and although vague goals are seen as an inherent feature of public organizations (Heffron 1989), a classification of existing performance studies shows that there is substantial variation in the level of formality.

Third, the degree of subjectivity clarifies whether a performance criterion conceptually concerns interior experiences, and perceptions contribute greatly to the discussion of subjective versus objective performance measures, because they highlight how subjectivity is both a question about conceptual definition and data sources, and the criticism of subjective measures sometimes fails to differentiate between these two levels. Existing studies (Meier and O'Toole 2013; Meier et al. 2015; O'Toole and Meier 2013) have convincingly shown that it can be problematic for measurement validity to use perceptual measures to measure objective performance criteria, but our third distinction implies that it is worthwhile to discuss how we measure performance criteria, which on the conceptual level concern interior experiences and perceptions. If our effort to increase measurement validity means that we do not include any inherently subjective performance criteria, it could be a problem for the conceptual validity of performance research in public administration, because an important aspect of the theoretical performance concept would be ignored; for example, it is difficult to measure user satisfaction without using perceptual measures.

The fourth and fifth distinctions concern the process and product focuses of performance criteria. Distinguishing between different types of product focuses corresponds to the IOO and 3E models in the existing literature, while the explicit inclusion of process answers several calls for attention to this type of performance in public organizations (Moynihan et al. 2011, 143; Van Ryzin 2011).

Finally, the last distinction relates to the level of analysis, which varies considerably. This choice depends on the explanatory variables and on the actor who is expected to produce

a specific type of performance. For example, if several health workers are involved in a surgical procedure, it becomes less meaningful to conceptualize performance as individual.

The successful application of the distinctions requires being able to describe the central differences among the relevant studies that use them. The fact that all of the distinctions can be used to describe key differences between existing performance criteria in the literature (illustrated by our classification of 110 performance criteria) indicates that we do not have too many distinctions. There is, however, a balance between accuracy and detail, on the one hand, and parsimony and applicability on the other. We find that the distinctions are comprehensive and provide oversight of the central differences and similarities between the performance criteria. As such, the distinctions can inform our understanding and comparison of performance in at least three ways. First, our classification of existing performance criteria illustrated how a clear conceptualization can facilitate the better evaluation of existing performance studies. One key insight from our analysis of the existing criteria was that the literature is dominated by criteria focusing only on product but that it is also possible to analyze process aspects of performance. Another insight is that many criteria are implicit and that articles should be more explicit about their performance criteria. A last example is that comparisons of different performance studies should be very aware of the units of analysis, because performance can be conceptualized at many different levels in the public sector, and these different conceptualizations are not necessarily comparable.

Second, one implication of this is that the conceptual distinctions can help guide central choices related to research questions, data, and methods in future performance studies. There is no logical method for having new ideas, but most researchers try to formulate research questions that contribute to an identifiable scholarly literature, and the continued discussion in the discipline of what is conceptualized as performance will therefore affect our future research questions. For example, the distinction between process and product criteria can

highlight how traditional democratic values can also be seen as performance (Moynihan et al. 2011, 143). The distinctions developed in the article can also draw attention to new types of research designs (e.g., comparisons of organizations with different types of stakeholders), and they might contribute to more informed choices of data sources (e.g., basing the data choice on the degree of subjectivity of the relevant performance criterion) and higher comparability between studies (ensuring that we do not compare apples and oranges). The conceptual discussion of performance is an example of how a background concept can be systematized. This might inspire future research to clarify the conceptual space of other key public administration concepts, such as public service motivation and red tape.

A systematic understanding of performance can also clarify public debates about what performance is and which performance criteria are reasonable and legitimate. The conceptual space developed in this article can potentially help public managers systematize their thoughts and decisions about which criteria to prioritize in their organizations. According to Behn (2003), public managers often measure performance because it helps them to achieve different managerial purposes (evaluate, control, budget, motivate, promote, celebrate, learn, and improve), and different purposes require performance criteria with different characteristics (Behn 2003, 593). We think that it is very important for managers in public organizations and for other participants in public debates about performance to understand the entire conceptual space of performance before they select the most important performance criteria, which are then measured. Hopefully, this article is helpful in that regard. Although completeness in measurement is not possible (or even desirable), it is very useful to know explicitly which types of performance you measure—and which you do not.

In sum, clarifying how performance in public organizations is assessed facilitates the better evaluation of existing performance studies, improves our ability to design new

performance studies which fill important gaps in the literature, and clarifies public debates about what performance is and which performance criteria are reasonable and legitimate.

References

- Adcock, Robert, and David Collier. 2001. Measurement Validity: A Shared Standard for Qualitative and Quantitative Research. *American Political Science Review* 95(3): 529–46.
doi:10.1017/S0003055401003100.
- Amirkhanyan, Anna A., Hyun J. Kim, and Kristina T. Lambright. 2014. The Performance Puzzle: Understanding the Factors Influencing Alternative Dimensions and Views of Performance. *Journal of Public Administration Research and Theory* 24(1): 1–31. doi:10.1093/jopart/mut021.
- Andersen, Lotte B. 2005. *Offentligt Ansattes Strategier: Aflønning, arbejdsbelastning og professionel status for dagplejere, folkeskolelærere og tandlæger* [Public Sector Employees' Strategies: Rewards, Workloads, and Professional Status for Day-Care Workers, Elementary School Teachers, and Dentists]. Aarhus: Politica.
- Andersen, Lotte B., Eskil Heinesen, and Lene H. Pedersen. 2014. How Does Public Service Motivation among Teachers Affect Student Performance in Schools? *Journal of Public Administration Research and Theory* 24(3): 651–71. doi:10.1093/jopart/mut082.
- Andersen, Simon C., and Peter B. Mortensen. 2010. Policy Stability and Organizational Performance: Is there a Relationship? *Journal of Public Administration Research and Theory – LA English* 20(1): 1–22. doi:10.1093/jopart/mup005.
- Andrews, Rhys, George A. Boyne, Jennifer Law, and Richard M. Walker. 2005. External Constraints on Local Service Standards: The Case of Comprehensive Performance Assessment in English Local Government. *Public Administration* 83(3): 639–56.
- Andrews, Rhys M., George A. Boyne, Kenneth J. Meier, Laurence J. O'Toole, and Richard M. Walker. 2012. Vertical Strategic Alignment and Public Service Performance. *Public Administration* 90(1): 77–98. doi:10.1111/j.1467-9299.2011.01938.x.

- Andrews, Rhys M., George A. Boyne, and Richard M. Walker. 2006a. Strategy Content and Organizational Performance: An Empirical Analysis. *Public Administration Review* 66(1): 52–63. doi:10.1111/j.1540-6210.2006.00555.x.
- . 2006b. Subjective and Objective Measures of Organizational Performance: An Empirical Exploration. In *Public Service Performance: Perspectives on Measurement and Management*, edited by George A. Boyne, Kenneth J. Meier, Laurence J. O'Toole, and Richard M. Walker, 314–34. Cambridge: Cambridge University Press.
- . 2011. The Impact of Management on Administrative and Survey Measures of Organizational Performance. *Public Management Review* 13(2): 227–55. doi:10.1080/14719037.2010.532968.
- Angle, Harold L., and James L. Perry. 1981. An Empirical Assessment of Organizational Commitment and Organizational Effectiveness. *Administrative Science Quarterly* 26(1): 1–14. doi:10.2307/2392596.
- Ashworth, Rachel, George A. Boyne, and Tom Entwistle, eds. 2010. *Public Service Improvement: Theories and Evidence*. Oxford: Oxford University Press.
- Behn, Robert D. 2003. Why Measure Performance? Different Purposes Require Different Measures. *Public Administration Review* 63(5): 586–606. doi:10.1111/1540-6210.00322.
- . 2014. *The Performance Stat Potential: A Leadership Strategy for Producing Results*. Washington, DC: Brookings Institution Press.
- Bommer, William H. 1995. On the Interchangeability of Objective and Subjective Measures of Employee Performance: A Meta-Analysis. *Personnel Psychology* 48(3): 3–27 doi:10.1111/j.1744-6570.1995.tb01772.x.
- Boschken, Herman L. 1992. Analyzing Performance Skewness in Public Agencies: The Case of Urban Mass Transit. *Journal of Public Administration Research and Theory* 2(3): 265–88.
- Boyne, George A. 2002. Concepts and Indicators of Local Authority Performance: An Evaluation of the Statutory Frameworks in England and Wales. *Public Money & Management* 22(2): 17–24. doi:10.1111/1467-9302.00303.

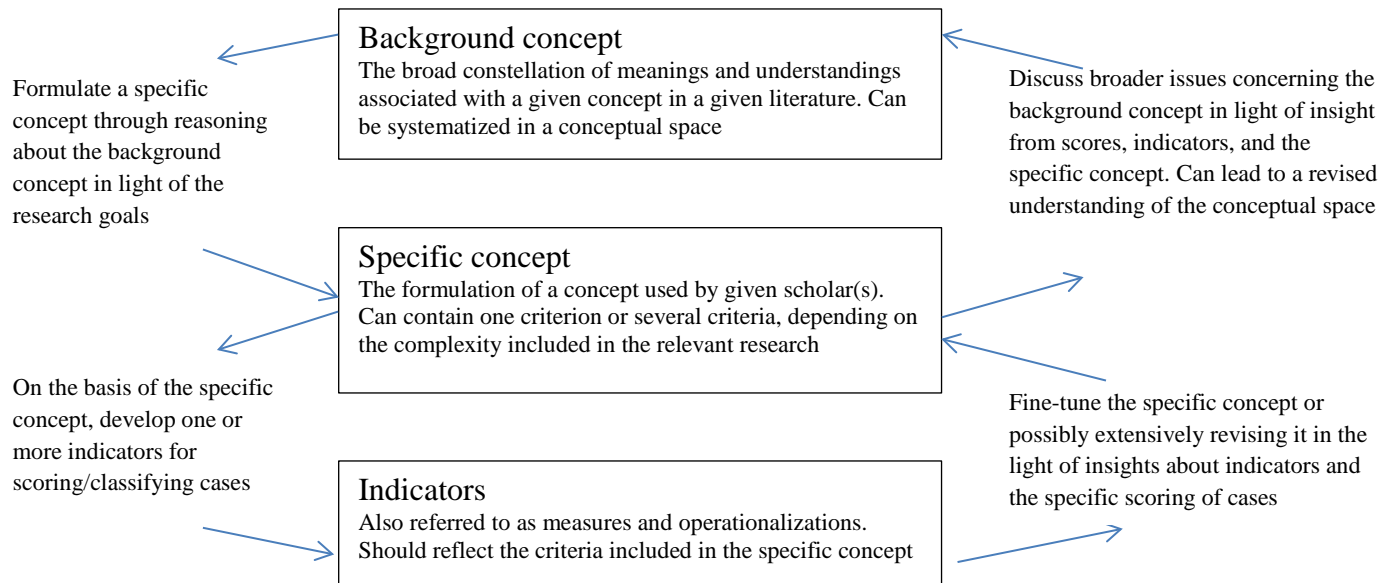
- . 2003a. Sources of Public Service Improvement: A Critical Review and Research Agenda. *Journal of Public Administration Research and Theory* 13(3): 367–94. doi:10.1093/jpart/mug027.
- . 2003b. What is Public Service Improvement? *Public Administration* 81(2): 211–27. doi:10.1111/1467-9299.00343.
- Boyne, George A., Kenneth J. Meier, Laurence J. O'Toole, and Richard M. Walker. 2006. Introduction. In *Public Service Performance: Perspectives on Measurement and Management*, edited by Boyne, George A., Kenneth J. Meier, Laurence J. O'Toole and Richard M. Walker, 1–13. Cambridge: Cambridge University Press.
- Brewer, Gene A. 2006. All Measures of Performance are Subjective: More Evidence on Federal Agencies. In *Public Service Performance: Perspectives on Measurement and Management*, edited by George A. Boyne, Kenneth J. Meier, Laurence J. O'Toole, and Richard M. Walker, 35–54. Cambridge: Cambridge University Press.
- Caillier, James G. 2011. Funding, Management, and Individual-Level Factors: What Factors Matter in Predicting Perceived Organizational Effectiveness? *International Journal of Public Administration* 34(7): 413–23. doi:10.1080/01900692.2011.569918.
- Cameron, Kim S. 1986. Effectiveness as Paradox: Consensus and Conflict in Conceptions of Organizational Effectiveness. *Management Science* 32(5): 539. doi:10.1287/mnsc.32.5.539.
- Collins, Brian K., and Brian J. Gerber. 2008. Taken for Granted? Managing for Social Equity in Grant Programs. *Public Administration Review* 68(6): 1128–1141. doi:10.1111/j.1540-6210.2008.00960.x.
- Collins, Randall. 1990. Market Closure and the Conflict Theory of the Professions. In *Professions in Theory and History: Rethinking the Study of Professions*, edited by Michael Burrage and Rolf Torstendahl, 24–43. London: Sage Publications.
- Freeman, R. Edward. [1984] 2010. *Strategic Management: A Stakeholder Approach*. Cambridge: Cambridge University Press.
- Friedson, Eliot. 2001. *Professionalism, the Third Logic: On the Practice of Knowledge*. Chicago: University of Chicago Press.

- Grönroos, Christian. 1984. A Service Quality Model and its Marketing Implications. *European Journal of Marketing* 18(4): 36–44. doi:10.1108/EUM00000000004784.
- Heffron, Florence A. 1989. *Organizations Theory and Public Organizations: The Political Connection*. Upper Saddle River, NJ: Prentice Hall.
- Hirsch, Paul M. and Daniel Z. Levin. 1999. Umbrella Advocates Versus Validity Police: A Life-Cycle Model. *Organization Science* 10 (2): 199-212.
- Jung, Chan Su. 2011. Organizational Goal Ambiguity and Performance. *International Public Management Journal* 14(2): 193–217. doi:10.1080/10967494.2011.589760.
- Jung, Chan Su. 2014. Extending the Theory of Goal Ambiguity to Programs: Examining the Relationship between Goal Ambiguity and Performance. *Public Administration Review* 74(2): 205–19. doi:10.1111/puar.12176.
- Kim, Sangmook. 2005. Individual-Level Factors and Organizational Performance in Government Organizations. *Journal of Public Administration Research and Theory* 15(2): 245–61. doi:10.1093/jopart/mui013
- Kuipers, Ben, Malcolm Higgs, Walter Kickert, Lars Tummers, Jolien Grandia, and Joris Van der Voet. 2014. The Management of Change in Public Organizations: A Literature Review. *Public Administration* 92(1): 1–20. doi:10.1111/padm.12040.
- Lynn, Laurence E., and Robbie W. Robichau. 2013. Governance and Organisational Effectiveness: Towards a Theory of Government Performance. *Journal of Public Policy* 33(2): 201–28. doi:10.1017/S0143814X13000056.
- Meier, Kenneth J., and Laurence J. O'Toole, Jr. 2003. Public Management and Educational Performance: The Impact of Managerial Networking. *Public Administration Review* 63(6): 689–99.
- . 2013. Subjective Organizational Performance and Measurement Error: Common Source Bias and Spurious Relationships. *Journal of Public Administration Research and Theory* 23(2): 429–56. doi:10.1093/jopart/mus057.

- Meier, Kenneth J., Søren C. Winter, Laurence J. O'Toole, Jr., Nathan Favero, and Simon C. Andersen. 2015. The Validity of Subjective Performance Measures: School Principals in Texas and Denmark. *Public Administration*, 93(4): 1084–101.
- Mintzberg, Henry. 1973. *The Nature of Managerial Work*. New York: Harper & Row.
- Moynihan, Donald P., Sergio Fernandez, Soonhee Kim, Kelly M. Leroux, Suzanne J. Piotrowski, Bradley E. Wright, and Kaifeng Yang. 2011. Performance Regimes amidst Governance Complexity. *Journal of Public Administration Research and Theory* 21(SUPPL. 1): i141-i155.
- O'Toole, Jr., Laurence J., and Kenneth J. Meier. 2013. I Think (I Am Doing Well), Therefore I Am: Assessing the Validity of Administrators' Self-Assessments of Performance. *International Public Management Journal* 16(1): 1–27. doi:10.1080/10967494.2013.796253.
- Ouchi, William G. 1980. Markets, Bureaucracies, and Clans. *Administrative Science Quarterly* 25(1): 129–41.
- Petticrew, Mark, and Helen Roberts. 2006. *Systematic Reviews in the Social Sciences: A Practical Guide*. Malden, MA: Blackwell.
- Radin, Beryl A. 2006. *Challenging the Performance Movement: Accountability, Complexity and Democratic Values*. Washington, DC: Georgetown University Press.
- Rainey, Hal G. 1997. *Understanding and Managing Public Organizations*. San Francisco: Jossey-Bass.
- Scott, W. Richard. 1995. *Institutions and Organizations: Ideas, Interests and Identities*. London: Sage Publications.
- Smith, Kevin B., and Christopher W. Larimer. 2004. A Mixed Relationship: Bureaucracy and School Performance. *Public Administration Review* 64(6): 728. doi:10.1111/j.1540-6210.2004.00419.x.
- Stoney, Christopher, and Diana Winstanley. 2001. Stakeholding: Confusion or Utopia? Mapping the Conceptual Terrain. *Journal of Management Studies* 38(5): 603–26.
<http://onlinelibrary.wiley.com/doi/10.1111/1467-6486.00251/pdf>.
- Talbot, Colin. 2008. Performance Regimes: The Institutional Context of Performance Policies. *International Journal of Public Administration* 31(14): 1569. doi:10.1080/01900690802199437.

- . 2010. *Theories of Performance: Organizational and Service Improvement in the Public Domain*. Oxford: Oxford University Press.
- Tummers, Lars, Victor Bekkers, Evelien Vink, and Michael Musheno. 2015. Coping During Public Service Delivery: A Conceptualization and Systematic Review of the Literature. *Journal of Public Administration Research and Theory*. Advance Access published January 12, 2015, doi:10.1093/jopart/muu056.
- Van Loon, Nina Mari, Wouter Vandenabeele, and Peter Leisink. 2015. Clarifying the Relationship between Public Service Motivation and In-Role and Extra-Role Behaviours: The Relative Contributions of Person–Job and Person–Organization Fit. *American Review of Public Administration*, DOI: 10.1177/0275074015617547, 1–21 (online before print).
- Van Ryzin, Gregg G. 2011. Outcomes, Process, and Trust of Civil Servants. *Journal of Public Administration Research and Theory* 21(4): 745–60. doi:10.1093/jopart/muq092.
- Voorberg, William, Viktor Bekkers, and Lars Tummers. 2014. A Systematic Review of Co-Creation and Co-Production: Embarking on the Social Innovation Journey. *Public Management Review*. doi:10.1080/14719037.2014.930505.
- Richard M. Walker, Richard M and Rhys Andrews. 2015. Local Government Management and Performance: A Review of Evidence *J Public Adm Res Theory* (2015) 25 (1): 101-133. doi:10.1093/jopart/mut038
- Walker, Richard M., George A. Boyne, and Gene A. Brewer, eds. 2010. *Public Management and Performance: Research Directions*. Cambridge: Cambridge University Press.
- Williams, Larry J., and Stella E. Anderson. 1991. Job Satisfaction and Organizational Commitment as Predictors of Organizational Citizenship and in-Role Behaviors. *Journal of Management* 17(3): 601–17. doi:10.1177/014920639101700305.
- Yang, Kaifeng. 2009. Examining Perceived Honest Performance Reporting by Public Organizations: Bureaucratic Politics and Organizational Practice. *Journal of Public Administration Research and Theory* 19(1): 81–105. doi:10.1093/jopart/mum042.

Figure 1: Illustration of the conceptualization process



Note: The figure is inspired by Adcock and Collier (2001, 531), who also include a fourth level (Scores for Cases) because they are also interested in specific measurement whereas we focus on conceptualization.

Figure 2: Flowchart for the process for identifying and retaining studies

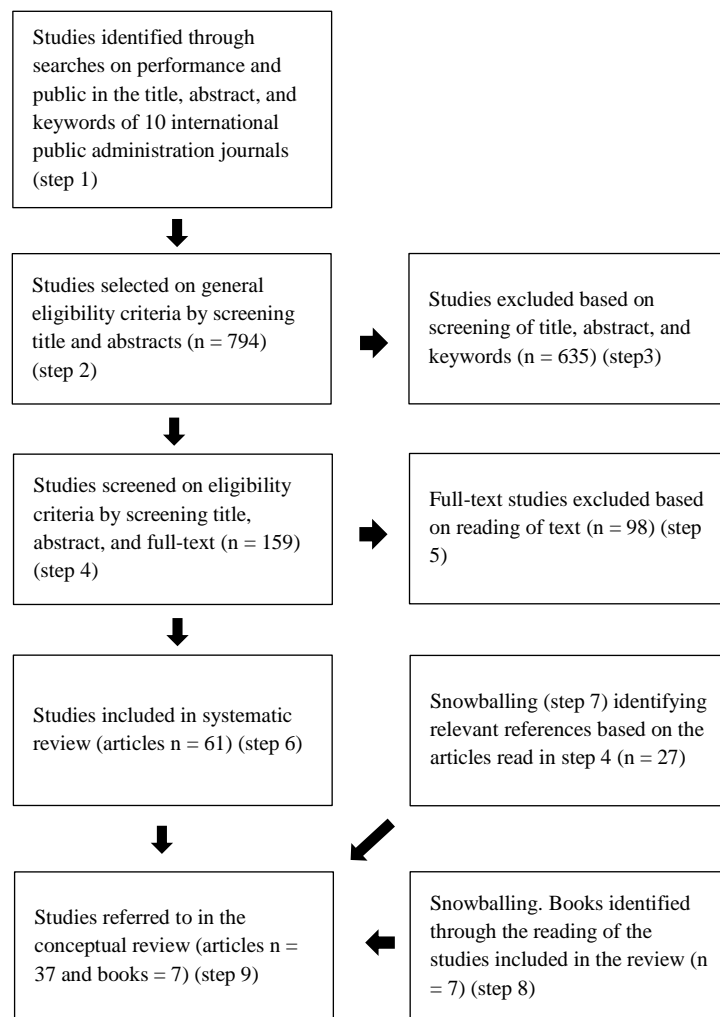


Table 1: A conceptual space of performance

Distinction	Question	Example
1. Stakeholders	Who decides what good performance is?	E.g., politicians, users, or professionals
2. Formality	To what degree is performance formally or informally defined?	E.g., written goals vs. implicit expectations
3. Subjectivity	To what degree does performance concern interior perceptions versus exterior phenomena?	E.g., consumer satisfaction vs. actual physical health as the performance criterion
4. Type of process focus	What type of process focus does the criterion have (if any)?	E.g., fair process or user participation in the decision-making process
5. Type of product focus	What type of process focus does the criterion have (if any)?	E.g., effectiveness or output quantity
6. Units of analysis	Who performs and on what level?	E.g., individual, organization, or program

Online Appendix: Categorization of Existing Studies

In this appendix, the studies included in tables 2–8 in Walker and Andrews (2015) are classified to illustrate how the six distinctions in the conceptual space of public organization performance can inform our understanding of performance criteria. The procedure for this classification was as follows:

1. Constructing categories for the classification

Based on the theoretical arguments behind the six distinctions combined with the purpose of this classification, we decided on categories for each distinction (Table A1). Here, it should be noted that the categories for other classifications could be more or less detailed. For example, this classification only distinguishes between external stakeholders, internal stakeholders, and researcher-determined criteria. Researchers obviously always decide what criteria they use, but this category is used when researchers explicitly use a criteria that is not used by other stakeholders. If the stakeholders are not mentioned and it is impossible to infer their identity, the category “implicit” is used. This category can also be used in parentheses behind another category if the stakeholder identity is inferred (rather than directly identified) from the text. This use of “implicit” in parentheses can be used for all six distinctions if necessary.

2. Identifying and sorting performance criteria to be classified

This classification is based on all of the performance criteria classified in Walker and Andrews (2015). Specifically, we copied tables 2–8 into one document and deleted doubles (criteria appearing more than once in tables 2–8). The criteria were then sorted alphabetically, and the lines between criteria from the same articles were dotted. Many performance studies have multiple performance criteria, and we ended up classifying 110 criteria from 66 studies. Some of the studies had criteria that were not included in Walker and Andrews (2015), because there had not been correlations with the independent variables investigated by them. These criteria are not included in this categorization either, as the idea is to see how our categorization expands an existing categorization.

3. Categorizing the performance criteria

Based on the original articles, we categorized all 110 performance criteria in relation to the six distinctions. Specifically, we first selected 10 articles that were categorized by both authors (to calibrate categorization practices and make sure that the procedures were inter-subjectively transferable). The rest of the criteria were then coded by one of the researchers.

4. Summarizing the results from the classification

The final classification shown in table 2 was transferred to SPSS to support the easy calculation of different summary statistics, such as the proportion of criteria with external stakeholders.

Table A1: List of categories for the classification of performance criteria from Walker and Andrews (2015)

Criterion	Categories
1: Stakeholder	<ul style="list-style-type: none"> - Internal - External - Both internal and external - Implicit (can also be used in parentheses if possible to deduce identity of stakeholder) - Researcher determined (used in parentheses)
2: Formality	<ul style="list-style-type: none"> - Informal (not written) - Medium formal (written/described, not legally binding) - Formal (legally binding) - Implicit (can also be added in parentheses if possible to infer formality)
3: Inherent subjectivity	<ul style="list-style-type: none"> - Inherently subjective - Inherently objective - Implicit (can also be added in parentheses if possible to infer subjectivity)
4: Type of process focus	<ul style="list-style-type: none"> - No process focus - Accountability - Equal assess = equity in process - Participation - Probity - User satisfaction with process - Responsiveness
5: Type of product focus	<ul style="list-style-type: none"> - No product focus - Input (e.g., economy, user satisfaction with input) - Output (e.g., quantity, quality, user satisfaction with output) - Outcome (effectiveness = achievement of formal service objectives, impact (informal outcome), equity in outcomes, user satisfaction with outcome) - Output/input (efficiency) - Outcome/input (cost-effectiveness = value for money)
6: Unit of analysis	<ul style="list-style-type: none"> - Individual citizens - Individual employees - Teams of employees - Organizations (also including local authorities) - Programs - Implicit (can also be added in parentheses)

Table A2: List of categories for the classification of performance criteria from Walker and Andrews (2015)

	Criterion	1 Stakeholder	2 Formality	3 Inherent subjectivity	4 Which process focus if any?	5 Which product focus if any?	6 Level of analysis
Andrews (2010)	Staff safety and wellbeing	External (central government)	Formal	Inherently objective	No process	Quality: output quality, outcome quality	Org./fire authority
Andrews (2010)	Promotion of equality and diversity	External (central government)	Formal	Inherently objective	Equity in process	No product focus in equity measure	Org./fire authority
Andrews and Boyne (2010)	Service performance index	External (central government)	Formal for effectiveness, medium for the rest of this index	Inherently subjective (for service satisfaction)	No process focus	Index 1: effectiveness, output quality, output quantity, efficiency	Org./local authorities
Andrews and Boyne (2010)	User product satisfaction	External (local citizens)	Medium	Inherently subjective	No process focus	Index 2: citizen satisfaction.	Org./local authorities
Andrews and Boyne (2011)	Service performance index	External (central government)	Formal for effectiveness, medium for the rest of this index	Inherently objective	No process focus	Index: effectiveness	Org./local authorities
Andrews and Boyne (2011)	Service performance per input	Researcher determined	Medium	Inherently objective	No process focus	Cost effectiveness	Org./local authorities
Andrews and Boyne (2011)	Promotion of equality	External (central government)	Medium	Inherently objective (but self-ass subj. measure of equity)	Equity in process (fair employment outcomes) and equal access to services	No product focus in equity measure	Org./local authorities
Andrews et al. (2005)	Service performance index	External (central government)	Formal for effectiveness, medium for the rest	Inherently subjective (for service	No process focus	Index: effectiveness, output quality, output quantity,	Org./local authorities

with O'Toole			of this index	satisfaction, no for rest) (implicit)		efficiency, service satisfaction. Outcome: value for money	
Andrews et al. (2005) with O'Toole	User product satisfaction	External (local citizens)	Medium	Inherently subjective	No process focus	User satisfaction	Org./local authorities
Andrews et al. (2005) with Law	Service performance index	External (central government)	Formal for effectiveness, medium for the rest of this index	Inherently subjective for service satisfaction	No process focus	Index: effectiveness, output quality, output quantity, efficiency, service satisfaction, outcomes	Org./local authorities
Andrews et al. (2008) with Boyne	Service performance index	External (national assembly of wales)	Medium	Inherently objective	Probity	Effectiveness: index with effectiveness, output quantity and outcome quality	Org./local authorities
Andrews et al. (2009)	Service performance index	External (national assembly of Wales)	Medium	Inherently objective	Probity, responsiveness	Index: effectiveness output quantity, outcome quality	Org./local authorities
Andrews et al. (2010)	Service performance index	Internal and external (corporate officers, politicians, service managers)	Medium	Inherently subjective for the two satisfaction measures	Equity in process (i.e., how fairly your services are distributed amongst citizens)	Effectiveness, quality, value for money, efficiency, consumer satisfaction, staff satisfaction, and promoting the well-being of local people	Org./local authorities

Andrews et al. (2010)	Service performance index	Central government agency	Medium	Inherently objective	No process focus	Index: effectiveness output quantity, outcome quality	Org./local authorities
Andrews et al. (2010)	User product and process satisfaction	External (citizens)	Medium	Inherently subjective	User satisfaction: process part of “how the authority runs things”	User satisfaction: output part of “how the authority runs things”	Org./local authorities
Andrews et al. (2012)	Service performance index	External (central government)	Formal for effectiveness, medium for the rest of this index	Inherently subjective (satisfaction with garbage collection)	Equity in process (equal access to public housing)	Index: effectiveness, output quality, output quantity, efficiency, service satisfaction. Outcome: value for money	Org./local authorities
Andrews, Boyne, and Walker (2006)	Service performance index	External (central government)	Formal for effectiveness, medium for the rest of this index	Inherently subjective	No process focus	Index: effectiveness, output quality, output quantity, efficiency, service satisfaction, outcomes	Org./local authorities
Aslam and Yilmaz (2011)	Magnitude of services	External politicians	Formal	Inherently objective	No process focus	Quantity	Org./villages (in Pakistan)
Bohte (2004)	Proportion of students with acceptable	External politicians	Formal	Inherently objective	No process focus	Effectiveness	Org./local authorities (Texas public)

	academic competences						school districts)
Boyne and Chen (2006)	Proportion of students with acceptable academic competences	External central government (implicit)	Formal	Inherently objective	No process focus	Effectiveness (exam results)	Org./local authorities
Boyne and Gould-Williams (2003)	Organizational cost effectiveness	External	Medium	Inherently objective	No process focus	Cost effectiveness	Org./local authorities (Welsh local governments)
Boyne and Gould-Williams (2003)	Organizational efficiency	External	Medium	Inherently objective	No process focus	Efficiency	Org./local authorities (Welsh local governments)
Boyne and Gould-Williams (2003)	Perceived organizational quality	External	Medium	Inherently subjective	No process focus	Quality	Org./local authorities (Welsh local governments)
Boyne and Gould-Williams (2003)	User product satisfaction	External	Medium	Inherently subjective	No process focus	User satisfaction Best Value has enhanced customer satisfaction	Org./local authorities (Welsh local governments)
Boyne et al. (2011)	Service performance index	External politicians (implicit)	Formal	Inherently subjective	Equity in process (equal access to public housing)	Index: effectiveness, output quality, output quantity, efficiency, service	Org./local authorities

						satisfaction. Outcomes: Value for money	
Carmeli (2006)	Financial performance	External (researcher)	Informal	Inherently objective	No process	Efficiency (not efficiency rather input and goal attainment)	Org./local authorities in Israel
Carmeli (2006)	Perceived organizational performance	Internal (local authority heads) for perceived criteria and external (researcher) for other criteria	Informal	Inherently subjective (called perceived perf.)	No process focus	Index: Several indexes (output and effectiveness)	Org./local authorities in Israel
Davies and Coles (1981)	Resources	External (implicit)	(Implicit)	Inherently objective (implicit)	No process focus	Efficiency, input/economy, efficiency	Org./local authorities
Davis, Barton, and McMillan (1971)	Resources	External (implicit)	(Implicit)	Inherently objective (implicit)	No process focus	Efficiency, input/economy: efficiency	Org./local authorities
DeSantis and Renner (1994)	Resources	External (implicit)	(implicit)	Inherently objective (implicit)	No process focus	Efficiency, input/economy	Org./local authorities
Fitzgerald and Durant (1980)	User product and process satisfaction	External (citizens)	Informal	Inherently subjective	User satisfaction (additional influence desire)	User satisfaction: satisfied with municipal services	Org./local authorities

Folz (2004)	Solid waste diversion rate	External (local officials)	Implicit	Inherently objective	Participation	Effectiveness: solid waste diversion rate	Org./city level
Folz and Hazlett (1991)	Solid waste diversion rate	External (local officials)	implicit	Inherently objective	Participation	Effectiveness: solid waste diversion rate	Other/city level
Goerdel (2006)	Proportion of students with acceptable academic competences	External (state politicians)	Formal	Inherently objective	No process focus	Effectiveness (percentage of overall students who pass TAAS test)	Org./school districts in Texas
Goerdel (2006)	Equity in proportion of students with acceptable academic competences	External (researchers)	Informal	Inherently objective	No process focus	Equity in outcome. Percentage of Latino, African American, and low-income students who pass the TAAS exam	Org./school districts in Texas
Goerdel (2006)	Proportion of students with high quality academic competences	External (researchers)	Medium formal	Inherently objective	No process focus	Quality (average ACT and SAT scores)	Org./school districts in Texas
Hansen and Kjellberg (1976)	Resources	(Implicit)	Implicit	Inherently objective	No process focus	Efficiency: input	Org./local authorities
Hill (2005)	Proportion of students with acceptable academic competences	State politicians (Texas State Board of Education at the behest of legislature)	Formal	Inherently objective	No process focus	Effectiveness (TAAS pass rate)	Org./school districts in Texas

Im and Lee (2012)	User product and process satisfaction	External (users)	Informal	Inherently subjective	User perceived process quality, office environment quality, and social quality	Efficiency: user-perceived outcome quality	Org./local authorities (25 district governments in Seoul City)
Johansen (2012)	Proportion of students with acceptable academic competences	External (implicit)	Medium	Inherently objective	No process focus	Effectiveness (SATS)	Org./school districts in the state of Texas
Johansen (2012)	Equity in proportion of students with acceptable academic competences	External (implicit)	Medium	Inherently objective	No process focus	Equity in results	Org./school districts in the state of Texas
May and Winter (2007)	Perceived organizational performance	External (politicians – but implicit)	Medium	Inherently objective	No process focus	Effectiveness: index of perceived outcome based on the responses of middle managers about the extent to which the municipality has succeeded in getting clients to search for jobs, to be available for work, and to enter ordinary employment	Org./local authorities (Danish municipalities)
Meier (1993)	Proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process	Effectiveness	Org./school districts Florida

Meier (1993)	Equity in process	External	Medium	Inherently objective	Equity in process disciplinary actions (e.g., corporal punishment)	No product focus in equity measure	Org./school districts Florida
Meier and Bothe (2001)	Equity in proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Equity in results	Org./school districts
Meier and Hicklin (2008)	Proportion of students with acceptable academic competences	External (state-level politicians)	Formal	Inherently objective	No process focus	Effectiveness	Org./school districts
Meier and Hicklin (2008)	Proportion of students with high quality academic competences	External (state-level politicians)	Formal	Inherently objective	No process focus	Quality (college bound)	Org./school districts
Meier and Nicholson-Crotty (2006)	Arrest rate	Implicit (researcher)	Medium	Inherently objective	No process focus	Effectiveness (rape arrest rate)	Org./police departments
Meier and O'Toole (2001)	Proportion of students with acceptable academic competences	External (state politicians)	Formal	Inherently subjective	No process focus	Effectiveness: percentage of students in each school district who pass state-required, standardized reading, writing, and mathematics tests each year	Org./Texas school districts

Meier and O'Toole (2002)	Proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Effectiveness (test-scores, average ACT and average SAT)	Org./school districts
Meier and O'Toole (2002)	Equity in proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Equity in outcome	Org./school districts
Meier and O'Toole (2002)	Proportion of students with high quality academic competences	External	Formal	Inherently objective	No process focus	Quality output	Org./school districts
Meier and O'Toole (2003)	Proportion of students with acceptable academic competences	External (state-level politicians)	Formal	Inherently objective	No process focus	Effectiveness: percentage of students in each school district who pass state-required, standardized reading, writing, and mathematics tests each year	Org./school districts (Texas)
Meier and O'Toole (2003)	Equity in proportion of students with acceptable academic competences	External (implicit)	Formal	Inherently objective	No process focus	Equity in outcome	Org./school districts (Texas)
Meier and O'Toole (2008)	Proportion of students with acceptable academic	External (implicit)	Formal	Inherently objective	No process focus	Effectiveness	Org./school districts (Texas)

	competences						
Meier and O'Toole (2008)	Equity in proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Equity in outcome	Org./school districts (Texas)
Meier and O'Toole (2008)	Proportion of students with high quality academic competences	External	Formal	Inherently objective	No process focus	Quality of outcome	Org./school districts (Texas)
Meier and O'Toole (2010)	Proportion of students with acceptable academic competences	External (implicit)	Formal	Inherently objective	No process focus	Effectiveness	Org./school districts (Texas)
Meier and O'Toole (2010)	Proportion of students with high quality academic competences	External	Formal	Inherently objective	No process focus	Quality of outcome	Org./school districts (Texas)
Meier and Stewart (1992)	Proportion of students with acceptable academic competences	External (implicit, but seem to be politicians)	Formal/medium formal ("are required to")	Inherently objective	No process focus	Effectiveness (student achievement tests)	Org./school districts in Florida
Meier and Stewart (1992)	Equity in product and process	External (researcher)	Informal	Inherently objective	Equity in process: equal punishments	Equity in outcome: test results, equal assignment to classes	School districts in Florida

Meier et al. (1999)	Proportion of students with acceptable academic competences	External (state politicians)	Formal	Inherently objective	No process focus	Effectiveness percentage of students who passed standardized competency tests	Org./school districts (Texas)
Meier et al. (1999)	Equity in proportion of students with acceptable academic competences	External (researchers)	Informal	Inherently objective	No process focus	Equity in outcome: percentage of minority (Afro-American or Latino) students who pass test divided by percentage of Anglo students who pass test	Org./school districts (Texas)
Meier et al. (2006)	Proportion of students with acceptable academic competences	External (implicit)	Formal	Inherently objective	No process focus	Effectiveness: effectiveness, output quantity	Org./school districts (Texas)
Meier et al. (2007)	Proportion of students with acceptable academic competences	External (implicit)	Formal	Inherently objective	No process focus	Effectiveness	Org./school districts (Texas)
Meier et al. (2007)	Equity in proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Equity in outcome	Org./school districts (Texas)
Meier et al. (2007)	Proportion of students with high quality academic competences	External	Formal	Inherently objective	No process focus	Quality output	Org./school districts (Texas)

Meier et al. (2010)	Proportion of students with acceptable academic competences	External (implicit)	Formal	Inherently objective	No process focus	Effectiveness	Org./school districts (Texas)
Meier et al. (2010)	Proportion of students with high quality academic competences	External	Formal	Inherently objective	No process focus	Quality output	Org./school districts (Texas) district
Meier, O'Toole, and Hicklin (2010)	Proportion of students with acceptable academic competences	External (state politicians)	Formal	Inherently objective	No process focus	Effectiveness TAKS pass rate	Org./school districts (Texas)
Melkers and Willoughby (2005)	Perceived importance for budget	Implicit	Medium	Inherently subjective (perceived effects)	No process focus	Cost-effectiveness? Rather input	Org./(see below)
Melkers and Willoughby (2005)	Perceived importance for budget	Implicit	Medium	Inherently subjective (perceived effects)	No process focus	Index: lasting effects: combined quality, effectiveness, responsiveness etc.	Org./"in the departments in your city/county"
Miranda and Lerner (1995)	Resources	Implicit	Implicit	Inherently objective	No process focus	Efficiency: input (expenditures)	Org./municipal police departments
Nicholson-Crotty and	Percentage of "index crimes" cleared by	External (FBI)	Medium	Inherently objective	No process focus	Effectiveness	Org./police departments)

O'Toole (2004)	arrest						
Nunn (2001)	Resources	Implicit	Medium	Inherently objective	No process focus	Efficiency: input	Org./municipal police agencies
O'Toole and Meier (2003)	Proportion of students with acceptable academic competences	External (implicit)	Formal	Inherently objective	No process focus	Effectiveness	Org./school districts (Texas)
O'Toole and Meier (2003)	Equity in proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Equity in outcome	Org./school districts (Texas)
O'Toole and Meier (2003)	Proportion of students with high quality academic competences	External	Formal	Inherently objective	No process focus	Quality output	Org./school districts (Texas) districts
O'Toole and Meier (2004b)	Proportion of students with acceptable academic competences	External (implicit)	Formal	Inherently objective	No process focus	Effectiveness	Org./school districts (Texas)
O'Toole and Meier (2004b)	Equity in proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Equity in outcome	Org./school districts (Texas)

O'Toole and Meier (2004b)	Proportion of students with high quality academic competences	External	Formal	Inherently objective	No process focus	Quality output	Org./school districts (Texas)
O'Toole and Meier (2006)	Proportion of students with acceptable academic competences	External (implicit)	Formal	Inherently objective	No process focus	Effectiveness	Org./school districts (Texas)
O'Toole and Meier (2006)	Equity in proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Equity in outcome	Org./school districts (Texas)
O'Toole and Meier (2006)	Proportion of students with high quality academic competences	External	Formal	Inherently objective	No process focus	Quality output	Org./school districts (Texas)
O'Toole and Meier (2009)	Proportion of students with acceptable academic competences	External (implicit)	Formal	Inherently objective	No process focus	Effectiveness	Org./school districts (Texas)
O'Toole and Meier (2009)	Equity in proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Equity in outcome	Org./school districts (Texas)
O'Toole and Meier (2009)	Proportion of students with high quality academic	External	Formal	Inherently objective	No process focus	Quality output	Org./school districts (Texas)

	competences						
Owens and Kukla-Acevedo (2012)	Proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Effectiveness state-required TAAS exam (all student pass rate)	Org./school districts (Texas)
Owens and Kukla-Acevedo (2012)	Equity in proportion of students with acceptable academic competences	External	Formal	Inherently objective	No process focus	Equity in outcome	Org./school districts (Texas)
Paletta (2012)	Proportion of students with acceptable academic competences	External (implicit)	Medium	Inherently objective	No process focus	Effectiveness	Org./schools
Pitts (2005)	Proportion of students with acceptable academic competences	External (state politicians)	Formal	Inherently objective	No process focus	Effectiveness: student pass rate on the Texas Assessment of Academic Skills (TAAS)	Org./school districts (Texas)
Pitts (2005)	Proportion of students with high quality academic competences	External (researcher)	Medium	Inherently objective	No process focus	Quality: percentage of students earning above 1110 on the SAT college entrance examination	Org./school districts (Texas)

Pitts (2007)	Proportion of students with acceptable academic competences	External (state politicians)	Formal	Inherently objective	No process focus	Effectiveness: SAT Pass rate	Org./school districts (Texas)
Pitts (2007)	Equity in proportion of students with acceptable academic competences	External (implicit)	Medium	Inherently objective	No process focus	Equity in results (TAAS pass rate for Whites, African-American, Latino)	Org./school districts (Texas)
Pitts (2007)	Proportion of students with high quality academic competences	External (implicit)	Medium	Inherently objective	No process focus	Quality: college-bound (i.e., 1110+) percentage of students earning > 1110 on SAT.	Org./school districts (Texas)
Pitts and Jarry (2009)	Proportion of students with acceptable academic competences	External (state politicians)	Formal	Inherently objective	No process focus	Effectiveness: percentage of students who pass TAAS examination	Org./school districts (Texas)
Roch and Pitts (2012)	Proportion of students with acceptable academic competences	External (state politicians – implicit)	Formal	Inherently objective	In-school suspension and out-of-school suspension	Effectiveness: student pass rates from standardized tests	Org./Georgia elementary schools
Ruggiero et al. (1995)	Proportion of students with acceptable academic competences	External (implicit)	Medium	Inherently objective	No process focus	Effectiveness: Efficiency	Org./school districts

Smith (2003)	Number of police killings of felons	External (implicit)	Medium	Inherently objective	No process focus	Quality	Other/cities over 100,000
Sorensen (2007)	Resources	External (implicit)	Medium	Inherently objective	No process focus	Efficiency: (costs/inhabitant for garbage collection)	Org./municipalities
Walker and Boyne (2006)	Perceived service performance index	Internal (implicit)	Medium	Inherently subjective	No process focus	Cost effectiveness	Org./local authorities
Walker and Boyne (2006)	Perceived service performance index	Internal (implicit)	Medium	Inherently subjective	No process focus	Efficiency	Org./local authorities
Walker and Boyne (2006)	Perceived service performance index	Internal (implicit)	Medium	Inherently subjective	No process focus	Effectiveness	Org./local authorities
Walker and Boyne (2006)	Perceived service performance index	Internal (implicit)	Medium	Inherently subjective	No process focus	User satisfaction	Org./local authorities
Walker and Boyne (2006)	Perceived service performance index	Internal (implicit)	Medium	Inherently subjective	No process focus	Index	Org./local authorities
Walker and Williams (1986)	Administrative product and process performance	External (implicit at that it is dep. of health and social security)	Medium (defined by researchers in items)	Inherently objective	Responsiveness	Quality: output	Org./local offices DHSS

Walker et al. (2010)	Service performance index	External (central government)	Formal	Inherently subjective (satisfaction with garbage collection)	Equity in process (equal access to public housing)	Index: effectiveness, output quality, output quantity, efficiency, service satisfaction. Outcomes: Value for money	Org./local authorities
Walker et al. (2011)	Service performance index	<i>External: citizens and the central government's agent, the Audit Commission</i>	Formal	Inherently subjective (e.g., satisfaction with garbage collection)	Equity in process (equal access to public housing)	Index: effectiveness, output quality, output quantity, efficiency, service satisfaction. Outcomes: Value for money	Org./local authorities
Walker et al. (2011)	User product and process satisfaction	External: citizens and the central government's agent, the Audit Commission	Formal	Inherently subjective (e.g., satisfaction with garbage collection)	User satisfaction with process	User satisfaction with product	Org./local authorities
Wilkins and Williams (2008)	No process discrimination	External	Medium	Inherently objective	Equity in process No discrimination/equal treatment	No product focus	Org./division of police (parts of San Diego Police Department)

Bibliography

Andrews, Rhys. 2010. The Impact of Modernisation on Fire Authority Performance: An Empirical Evaluation. *Policy & Politics* 38(4): 599–617.

Andrews, Rhys, and George A. Boyne. 2010. Capacity, Leadership, and Organizational Performance: Testing the Black Box Model of Public Management. *Public Administration Review* 70(3): 443–54.

———. 2011. Corporate Capacity and Public Service Performance. *Public Administration* 89(3): 894–908.

- Andrews, Rhys, George A. Boyne, Jennifer Law, and Richard M. Walker. 2005. External Constraints on Local Service Standards: The Case of Comprehensive Performance Assessment in English Local Government. *Public Administration* 83(3): 639–56.
- . 2008. Organizational Strategy, External Regulation and Public Service Performance. *Public Administration* 86(1): 185–203.
- . 2009. Strategy Formulation, Strategy Content and Performance: An Empirical Analysis. *Public Management Review* 11(1): 1–22.
- Andrews, Rhys, George A. Boyne, Kenneth J. Meier, Laurence J. O'Toole, Jr., and Richard M. Walker. 2005. Representative Bureaucracy, Organizational Strategy, and Public Service Performance: An Empirical Analysis of English Local Government. *Journal of Public Administration Research and Theory* 15(4): 489–504.
- . 2012. Vertical Strategic Alignment and Public Service Performance. *Public Administration* 90(1): 77–98.
- Andrews, Rhys, George A. Boyne, M. Jae Moon, and Richard M. Walker. 2010. Assessing Organizational Performance: Exploring Differences between Internal and External Measures. *International Public Management Journal* 13(2): 105–29.
- Andrews, Rhys, George A. Boyne, and Richard M. Walker. 2006. Strategy Content and Organizational Performance: An Empirical Analysis. *Public Administration Review* 66(1): 52–63.
- Aslam, Ghazia, and Serdar Yilmaz. 2011. Impact of Decentralization Reforms in Pakistan on Service Delivery: An Empirical Study. *Public Administration and Development* 31(3): 159–71.
- Bohte, John. 2004. Examining the Impact of Charter Schools on Performance in Traditional Public Schools. *Policy Studies Journal* 32(4): 501–20.
- Boyne, George A., and Alex A. Chen. 2007. Performance Targets and Public Service Improvement. *Journal of Public Administration Research and Theory* 17(3): 455–77.
- Boyne, George, and Julian Gould-Williams. 2003. Planning and Performance in Public Organizations an Empirical Analysis. *Public Management Review* 5(1): 115–32.
- Boyne, George A., Oliver James, Peter John, and Nicolai Petrovsky. 2011. Top Management Turnover and Organizational Performance: A Test of a Contingency Model. *Public Administration Review* 71(4): 572–81.
- Brewer, Gene A., and Richard M. Walker. 2010. The Impact of Red Tape on Governmental Performance: An Empirical Analysis. *Journal of Public Administration Research and Theory* 20(1): 233–57.

- Carmeli, Abraham. 2006. The Managerial Skills of the Top Management Team and the Performance of Municipal Organisations. *Local Government Studies* 32(2): 153–76.
- Davies, Bleddyn, Andrew Barton, and Ian McMillan. 1971. Variations in the Provision of Local Authority Health and Welfare Services for the Elderly: A Comparison between Counties and County Boroughs. *Social Policy & Administration* 5(2): 100–24.
- Davies, Bleddyn, and Oliver Coles. 1981. Towards a Territorial Cost Function for the Home Help Service. *Social Policy & Administration* 15(1): 32–42.
- DeSantis, Victor S., and Tari Renner. 1994. The Impact of Political Structures on Public Policies in American Counties. *Public Administration Review*: 291–95.
- Fitzgerald, Michael R., and Robert F. Durant. 1980. Citizen Evaluations and Urban Management: Service Delivery in an Era of Protest. *Public Administration Review*: 585–94.
- Folz, David H. 2004. Service Quality and Benchmarking the Performance of Municipal Services. *Public Administration Review* 64(2): 209–20.
- Folz, David H., and Joseph M. Hazlett. 1991. Public Participation and Recycling Performance: Explaining Program Success. *Public Administration Review*: 526–32.
- Goerdel, Holly T. 2006. Taking Initiative: Proactive Management and Organizational Performance in Networked Environments. *Journal of Public Administration Research and Theory* 16(3): 351–67.
- Hansen, Tore, and Francesco Kjellberg. 1976. Municipal Expenditures in Norway: Autonomy and Constraints in Local Government Activity. *Policy & Politics* 4(3): 25–50.
- Hill, Gregory C. 2005. The Effects of Managerial Succession on Organizational Performance. *Journal of Public Administration Research and Theory* 15(4): 585–97.
- Im, Tobin, and Seung Jong Lee. 2011. Does Management Performance Impact Citizen Satisfaction? *The American Review of Public Administration*: 0275074011408589.
- Johansen, Morgen S. 2012. The Direct and Interactive Effects of Middle and Upper Managerial Quality on Organizational Performance. *Administration & Society* 44(4): 383–411.
- May, Peter J., and Søren C. Winter. 2007. Collaborative Service Arrangements: Patterns, Bases, and Perceived Consequences. *Public Management Review* 9(4): 479–502.

- Meier, Kenneth J. 1993. Latinos and Representative Bureaucracy Testing the Thompson and Henderson Hypotheses. *Journal of Public Administration Research and Theory* 3(4): 393–414.
- Meier, Kenneth J., and John Bohte. 2001. Structure and Discretion: Missing Links in Representative Bureaucracy. *Journal of Public Administration Research and Theory* 11(4): 455–70.
- Meier, Kenneth J., George A. Boyne, Laurence J. O'Toole, Richard M. Walker, and Rhys Andrews. 2010. Alignment and results: Testing the Interaction Effects of Strategy, Structure, and Environment from Miles and Snow. *Administration and Society* 42: 160–92.
- Meier, Kenneth J., and Alisa Hicklin. 2008. Employee Turnover and Organizational Performance: Testing a Hypothesis from Classical Public Administration. *Journal of Public Administration Research and Theory* 18(4): 573–90.
- Meier, Kenneth J., Sharon H. Mastracci, and Kristin Wilson. 2006. Gender and Emotional Labor in Public Organizations: An Empirical Examination of the Link to Performance. *Public Administration Review* 66(6): 899–909.
- Meier, Kenneth J., and Jill Nicholson-Crotty. 2006. Gender, Representative Bureaucracy, and Law Enforcement: The Case of Sexual Assault. *Public Administration Review* 66(6): 850–60.
- Meier, Kenneth J., and Laurence J. O'Toole, Jr. 2001. Managerial Strategies and Behavior in Networks: A Model with Evidence from U.S. Public Education. *Journal of Public Administration Research and Theory* 11: 271–93.
- . 2002. Public Management and Organizational Performance: The Impact of Managerial Quality. *Journal of Policy Analysis and Management* 21(4): 629–43.
- . 2008. Management Theory and Occam's Razor: How Public Organizations Buffer the Environment. *Administration & Society* 39(8): 931–58.
- . 2010. Beware of Managers Not Bearing Gifts: How Management Capacity Augments the Impact of Managerial Networking. *Public Administration* 88(4): 1025–44.
- Meier, Kenneth J., Laurence J. O'Toole, Jr., George A. Boyne, and Richard M. Walker. 2007. Strategic Management and the Performance of Public Organizations: Testing Venerable Ideas Against Recent Theories. *Journal of Public Administration Research and Theory* 17(3): 357–77.
- Meier, Kenneth J., Laurence. O'Toole, Jr., and Alisa Hicklin. 2010. I've Seen Fire and I've Seen Rain: Public Management and Performance after a Natural Disaster. *Administration & Society* 41(8): 979–1003.
- . 2003. Public Management and Educational Performance: The Impact of Managerial Networking. *Public Administration Review* 63(6): 689–99.

- Meier, Kenneth J., and Joseph Stewart. 1992. The Impact of Representative Bureaucracies: Educational Systems and Public Policies. *The American Review of Public Administration* 22(3): 157–71.
- Meier, Kenneth J., Robert D. Wrinkle, and J.L. Polinard. 1999. Equity Versus Excellence in Organizations: A Substantively Weighted Least Squares Analysis. *The American Review of Public Administration* 29(1): 5–18.
- Melkers, Julia, and Katherine Willoughby. 2005. Models of Performance-Measurement use in Local Governments: Understanding Budgeting, Communication, and Lasting Effects. *Public Administration Review*: 180–90.
- Miranda, Rowan, and Allan Lerner. 1995. Bureaucracy, Organizational Redundancy, and the Privatization of Public Services. *Public Administration Review*: 193–200.
- Nicholson-Crotty, Sean, and Laurence J. O'Toole. 2004. Public Management and Organizational Performance: The Case of Law Enforcement Agencies. *Journal of Public Administration Research and Theory* 14(1): 1–18.
- Nunn, Samuel. 2001. Police Information Technology: Assessing the Effects of Computerization on Urban Police Functions. *Public Administration Review* 61(2): 221–34.
- O'Toole Jr, Laurence J., and Kenneth J. Meier. 2003. Plus Ça Change: Public Management, Personnel Stability, and Organizational Performance. *Journal of Public Administration Research and Theory: J-PART*: 43–64.
- . 2004. Desperately Seeking Selznick: Cooptation and the Dark Side of Public Management in Networks. *Public Administration Review* 64(6): 681–93.
- . 2006. Networking in the Penumbra: Public Management, Cooptative Links, and Distributional Consequences. *International Public Management Journal* 9(3): 271–94.
- . 2009. The Human Side of Public Organizations: Contributions to Organizational Performance. *The American Review of Public Administration*. 39 (5):499–518.
- Owens, Chris T., and Sharon Kukla-Acevedo. 2012. Network Diversity and the Ability of Public Managers to Influence Performance. *The American Review of Public Administration* 42(2): 226–45.
- Paletta, Angelo. 2012. Public Governance and School Performance: Improving Student Learning through Collaborative Public Management. *Public Management Review* 14(8): 1125–51.

- Pitts, David W. 2005. Diversity, Representation, and Performance: Evidence about Race and Ethnicity in Public Organizations. *Journal of Public Administration Research and Theory* 15(4): 615–31.
- . 2007. Representative Bureaucracy, Ethnicity, and Public Schools Examining the Link between Representation and Performance. *Administration & Society* 39(4): 497–526.
- Pitts, David W., and Elizabeth M. Jarry. 2009. Getting to Know You: Ethnic Diversity, Time and Performance in Public Organizations. *Public Administration* 87(3): 503–18.
- Roch, Christine H., and David W. Pitts. 2012. Differing Effects of Representative Bureaucracy in Charter Schools and Traditional Public Schools. *The American Review of Public Administration* 42(3): 282–302.
- Ruggiero, John, William Duncombe, and Jerry Miner. 1995. On the Measurement and Causes of Technical Inefficiency in Local Public Services: With an Application to Public Education. *Journal of Public Administration Research and Theory* 5(4): 403–28.
- Smith, Brad W. 2003. The Impact of Police Officer Diversity on Police-Caused Homicides. *Policy Studies Journal* 31(2): 147–62.
- Sørensen, Rune J. 2007. Does Dispersed Public Ownership Impair Efficiency? The Case of Refuse Collection in Norway. *Public Administration* 85(4): 1045–58.
- Walker, Richard M., Rhys Andrews, George A. Boyne, Kenneth J. Meier, and Laurence J. O'Toole, Jr. 2010. Wakeup Call: Strategic Management, Network Alarms, and Performance. *Public Administration Review* 70(5): 731–41.
- Walker, Richard M., and George A. Boyne. 2006. Public Management Reform and Organizational Performance: An Empirical Assessment of the UK Labour Government's Public Service Improvement Strategy. *Journal of Policy Analysis and Management* 25(2): 371–93.
- Walker, Richard M., Gene A. Brewer, George A. Boyne, and Claudia N. Avellaneda. 2011. Market Orientation and Public Service Performance: NPM Gone Mad? *Public Administration Review* 71: 707–17.
- Walker, Robert, and Julie Williams. 1986. Housing Benefit: Some Determinants of Administrative Performance. *Policy & Politics* 14(3): 309–34.
- Wilkins, Vicky M., and Brian N. Williams. 2008. Black or Blue: Racial Profiling and Representative Bureaucracy. *Public Administration Review* 68(4): 654–64.