Article

Review of Public Personnel Administration
2017, Vol. 37(2) 245–270

DOI: 10.1177/0734371X17698189
journals.sagepub.com/home/rop

SAGE

We Know the Yin—But Where Is the Yang? Toward a Balanced Approach on Common Source Bias in Public Administration Scholarship

Bert George[1] and Sanjay K. Pandey[2]

Abstract

Surveys have long been a dominant instrument for data collection in public administration. However, it has become widely accepted in the last decade that the usage of a self-reported instrument to measure both the independent and dependent variables results in common source bias (CSB). In turn, CSB is argued to inflate correlations between variables, resulting in biased findings. Subsequently, a narrow blinkered approach on the usage of surveys as single data source has emerged. In this article, we argue that this approach has resulted in an unbalanced perspective on CSB. We argue that claims on CSB are exaggerated, draw upon selective evidence, and project what should be tentative inferences as certainty over large domains of inquiry. We also discuss the perceptual nature of some variables and measurement validity concerns in using archival data. In conclusion, we present a flowchart that public administration scholars can use to analyze CSB concerns.

Keywords

common source bias, common method bias, common method variance, self-reported surveys, public administration

[1]Erasmus University Rotterdam, The Netherlands
[2]George Washington University, Washington, DC, USA

Corresponding Author:
Bert George, Erasmus University Rotterdam, Burgemeester Oudlaan 50, Mandeville Building T17-35, 3000 DR Rotterdam, The Netherlands.
Email: george@fsw.eur.nl

Introduction

Traditionally, public administration as a research field has used surveys extensively to measure core concepts (Lee, Benoit-Bryan, & Johnson, 2012; Pandey & Marlowe, 2015). Examples include survey items on public service motivation (PSM; for example, Bozeman & Su, 2015; Lee & Choi, 2016), bureaucratic red tape (e.g., Feeney & Bozeman, 2009; Pandey, Pandey, & Van Ryzin, 2016), public sector innovation (e.g., Audenaert, Decramer, George, Verschuere, & Van Waeyenberg, 2016; Verschuere, Beddeleem, & Verlet, 2014), and strategic planning (e.g., George, Desmidt, & De Moyer, 2016; Poister, Pasha, & Edwards, 2013). In doing so, public administration scholars do not differ from psychology and management scholars who often draw on surveys to measure perceptions, attitudes and/or intended behaviors (Podsakoff, MacKenzie, & Podsakoff, 2012). In public administration, surveys typically consist of a set of items that measure underlying variables, distributed to key informants such as public managers, politicians and/or employees. Such items can be targeted at the individual level, group level, and/or organizational level, where the latter two might require some form of aggregation of individual responses (Enticott, Boyne, & Walker, 2009). Such surveys offer several benefits, a key one being the efficiency and effectiveness in gathering data on a variety of variables simultaneously (Lee et al., 2012). However, despite the ubiquitous nature and benefits of surveys as an instrument for data collection in public administration, such surveys have not gone without criticism. Over the past years, one specific point of criticism has become a central focus of public administration journals: common source bias (CSB).

CSB, and interrelated terms such as common method bias, monomethod bias and common method variance (CMV), indicates potential issues when scholars use the same data source, typically a survey, to measure both independent and dependent variables simultaneously (Favero & Bullock, 2015; Jakobsen & Jensen, 2015; Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; Spector, 2006). Specifically, correlations between such variables are believed to be inflated due to the underlying CMV and the derived findings are thus strongly scrutinized and, often, criticized by reviewers (Pace, 2010; Spector, 2006). CMV is defined by Richardson, Simmering, and Sturman (2009, p. 763) as a "systematic error variance shared among variables measured with and introduced as a function of the same method and/or source." This variance can be considered "a confounding (or third) variable that influences both of the substantive variables in a systematic way," which might (but not necessarily will) result in inflated correlations between the variables derived from the same source (Jakobsen & Jensen, 2015, p. 5). When indeed CMV results in inflated correlations (or false positives), correlations are argued to suffer from CSB (Favero & Bullock, 2015).

In the fields of management and psychology, the debate on CSB includes a variety of perspectives (Podsakoff et al., 2012). While there are several scholars who argue the existence of and necessity to address CSB in surveys (e.g., Chang, van Witteloostuijn, & Eden, 2010; Harrison, McLaughlin, & Coalter, 1996; Kaiser, Schultz, & Scheuthle, 2007), there are others who argue that addressing CSB might require a more nuanced approach (e.g., Conway & Lance, 2010; Fuller, Simmering, Atinc, Atinc, & Babin,

2016; Kammeyer-Mueller, Steel, & Rubenstein, 2010; Spector, 2006). Similarly, editorial policies in management and psychology have ranged from, for instance, an editorial bias at the *Journal of Applied Psychology* toward any paper with potential CSB issues (Campbell, 1982) to a tolerance of these papers—as long as the necessary validity checks are conducted—at the *Journal of International Business Studies* (Chang et al., 2010) as well as the *Academy of Management Journal* (Colquitt & Ireland, 2009). In public administration, however, little has been written on CSB in general and studies that have discussed CSB typically center on CSB's impact on subjectively measured indicators of performance (e.g., Meier & O'Toole, 2013). Although these studies offer empirical evidence that self-reports of performance suffer from CSB and recommend avoiding self-report measures, counterarguments to the CSB problem for self-reports other than performance measures seem to be completely absent. As a result, an unbalanced approach on CSB has recently emerged in public administration, where papers that draw on a survey as single data source are greeted with a blinkered concern for potential CSB issues, reminiscent of Abe Kaplan's proverbial hammer (Kaplan, 1964).

Several illustrations indicate the existence of the proverbial CSB hammer. As part of the editorial policy of the *International Public Management Journal* (*IPMJ*), Kelman (2015) stipulates a "much-stricter policy for consideration of papers where dependent and independent variables are collected from the same survey, particularly when these data are subjective self-reports" and even "discourage[s] authors from submitting papers that may be affected by common-method bias" (pp. 1-2). Other top public administration journals also seem to be paying far more attention to CSB in recent years. For instance, when comparing the number of studies that explicitly mention CSB, common method bias, CMV or monomethod bias in 2010 and in 2015, we found that in 2010, the *Review of Public Personnel Administration* (*ROPPA*) published no such studies, *Public Administration Review* (*PAR*) published one and *Journal of Public Administration Research and Theory* (*JPART*) published six. Whereas, in 2015, those numbers increased to four studies for *ROPPA*, six studies for *PAR* and 10 studies for *JPART*. A recent tweet by Don Moynihan—the PMRA president at the time—at the Public Management Research Conference 2016 in Aarhus nicely summarizes public administration zeitgeist on CSB: "Bonfire at #pmrc2016: burning all papers with common source bias."

Our goal in this article is to move beyond "axiomatic and knee-jerk" responses and bring balance to consideration of CSB in public administration literature. We argue that the rapid rise of CSB in public administration literature represents an extreme response and thus there is a need to take pause and carefully scrutinize core claims about CSB and its impact. Our position is supported by four key arguments. First, we argue that the initial claims about CSB's influence might be exaggerated. Second, we argue that claims about CSB in public administration draw upon selective evidence, making "broad generalizing thrusts" suspect. Third, we argue that some variables are perceptual and can only be measured through surveys. Finally, we argue that archival data (collected from administrative sources) can be flawed and are not necessarily a better alternative to surveys. We conclude our article with a flowchart that public

administration scholars can use as a decision guide to appropriately and reasonably address CSB. As such, our study contributes to the methods debate within public administration by illustrating that the issues surrounding CSB need a more nuanced approach. There is need to move beyond reflexive and automatic invocation of CSB as scarlet letter and restore balance by using a more thoughtful and discriminating approach to papers using a survey as single data source.

In what follows, we first offer a brief review of the literature surrounding CSB both in public administration as well as management and psychology. Next, we strive for balance in the CSB debate by presenting the four arguments central to this article. In conclusion, we discuss the implications of our findings and present the indicated flowchart as a decision guide on how to reasonably address CSB concerns in public administration scholarship.

Literature Review

The best way to bring balance to considering CSB in public administration literature is to recognize that CSB—emphasized in public administration literature in the last decade—has a longer history in related social science disciplines. Because public administration research lags behind these sister social science disciplines in developing and incorporating methodological advances (Grimmelikhuijsen, Tummers, & Pandey, 2016), public administration scholarship is susceptible to rediscovering and selectively emphasizing methodological advancement. Therefore, to set the stage, we offer an overview of CSB in psychology and management literature before discussing relevant public administration studies.

CSB in Psychology and Management Literature

CSB, and related terms, have long been a focal point of methods studies in management and psychology (Podsakoff et al., 2003). Table 1 offers a brief overview of 13 of these studies based on a literature search through Google Scholar with a focus on highly cited articles in three top management and psychology journals (namely, *Journal of Applied Psychology*, *Journal of Management*, and *Organizational Research Methods*).

The selected studies cover a time range of nearly 25 years, illustrating that CSB seems to be an enduring topic in management and psychology. In the first article on the list, Podsakoff and Organ (1986) use the term *common method variance* to indicate issues when using the same survey for measuring the independent and dependent variables. They also offer almost prophetic insights into why CMV would become such an important topic in management and psychology: "It seems that organizational researchers do not like self-reports, but neither can they do without them" (Podsakoff & Organ, 1986, p. 531). They prognosticate that self-reports are the optimal means for gathering data on "individual's perceptions, beliefs, judgments, or feelings," and, as such, will remain central to the management and psychology scholarship (Podsakoff et al., 2012, p. 549).

Table 1. A Summary of Key Management and Psychology Articles on Common Source Bias and Related Terms.

Author(s)	Journal	Definition	Evidence	Key conclusion
Podsakoff and Organ (1986)	*Journal of Management*	"Because both measures come from the same source, any defect in that source contaminates both measures, presumably in the same fashion and in the same direction."	Own experience and articles.	"[W]e strongly recommend the use of **procedural or design remedies** for dealing with the common method variance problem as opposed to the use of statistical remedies or post-hoc patching up."
Spector (1987)	*Journal of Applied Psychology*	"[A]n artifact of measurement that biases results when relations are explored among constructs measured in the same way."	Articles with self-report measures of perceptions of jobs/work environments and affective reactions to jobs and data from Job Satisfaction Survey.	"The data and research results summarized here suggest that the problem [of CMV] may in fact be **mythical**."
Williams, Buckley, and Cote (1989)	*Journal of Applied Psychology*	*Not explicitly given*, but some explanation is provided, for example: "Values in the monomethod triangles are always inflated by shared method variance, because the correlation between the methods is 1.0 (the same method is used)."	Same articles as used by Paul E Spector (1987).	"In summary, this research indicates that the conclusions reached by Spector (1987) were an artifact of his method and that method variance is **real and present** in organizational behavior measures of affect and perceptions at work."
Bagozzi and Yi (1990)	*Journal of Applied Psychology*	"As an artifact of measurement, method variance can bias results when researchers investigate relations among constructs measured with the common method."	Same articles as used by Paul E Spector (1987) and Williams et al. (1989).	"Our reanalyses of the data analyzed by Spector (1987) suggest that the conclusions stated by Williams et al. (1989) could have been an artifact of their analytic procedure [. . .]. In the 10 studies [. . .], we found method variance to be sometimes significant, but **not as prevalent** as Williams et al. concluded."
Avolio, Yammarino, and Bass (1991)	*Journal of Management*	"[CMV] is defined as the overlap in variance between two variables attributable to the type of measurement instrument used rather than due to a relationship between the underlying constructs."	Survey data consisting of measures on leadership and outcomes (i.e., effectiveness of unit and satisfaction) gathered from immediate followers of a leader.	"The label "single-source effects" appears to be applied to the effects of an entire class of data collection that is rather wide-ranging. Stated more explicitly, single-source effects are **not necessarily an either-or issue** [. . .]."

(continued)

Table 1. (continued)

Author(s)	Journal	Definition	Evidence	Key conclusion
Doty and Glick (1998)	*Organizational Research Methods*	"[CMV] occurs when the measurement technique introduces systematic variance into the measure, [which] can cause observed relationships to differ from the true relationships among constructs."	Quantitative review of studies reporting multitrait-multimethod correlation matrices in six social science journals between 1980 and 1992.	"[CMV] is an **important concern** for organizational scientists. A significant amount of methods variance was found in all of the published data sets."
Keeping and Levy (2000)	*Journal of Applied Psychology*	"It is often argued that observed correlations are produced by the fact that the data originated from the same source rather than from relations among substantive constructs."	Survey data consisting of measures on appraisal and positive/negative affect gathered from employees.	"Although common method variance is a very prevalent critique of attitude and reactions research, additional support from other research areas seems very consistent with our findings and argues **against a substantial role** for common method variance [. . .]."
Lindell and Whitney (2001)	*Journal of Applied Psychology*	"[When] individuals' reports of their internal states are collected at the same time as their reports of their past behavior related to those internal states [. . .] the possibility arises that method variance (MV) has inflated the observed correlations."	Hypothetical correlations between leader characteristics, role characteristics, team characteristics, job characteristics, martial satisfaction and self-reported member participation.	"MV-marker-variable analysis should be conducted whenever researchers assess correlations that have been identified as being **most vulnerable** to CMV [. . .]. For correlations with low vulnerability [. . .], conventional correlation and regression analyses may provide satisfactory results."
Podsakoff, MacKenzie, Lee, and Podsakoff (2003)	*Journal of Applied Psychology*	"Most researchers agree that common method variance (i.e., variance that is attributable to the measurement method rather than to the constructs the measures represent) is a potential problem in behavioral research."	Literature review of previously published articles across behavioral research (e.g., management, marketing, psychology . . .).	"Although the strength of method biases may vary across research context, [. . .] CMV is **often a problem** and researchers need to do whatever they can to control for it. [. . .] this requires [. . .] implementing both procedural and statistical methods of control."

(continued)

Table 1. (continued)

Author(s)	Journal	Definition	Evidence	Key conclusion
Spector (2006)	*Organizational Research Methods*	"It has become widely accepted that correlations between variables measured with the same method, usually self-report surveys, are inflated due to the action of method variance."	Articles on turnover processes, social desirability, negative affectivity, acquiescence and comparison between multi and monomethod correlations.	"The time has come to **retire the term common method variance** and its derivatives and replace it with a consideration of specific biases and plausible alternative explanations for observed phenomena, regardless of whether they are from self-reports or other methods."
Pace (2010)	*Organizational Research Methods*	"86.5% [. . .] agreed or strongly agreed that CMV means that correlations among all variables assessed with the same method will likely be inflated to some extent due to the method itself."	Opinions of 225 editorial board members of *Journal of Applied Psychology*, *Journal of Organizational Behavior*, and *Journal of Management*.	"According to survey results, CMV is recognized as a frequent and potentially **serious issue** but one that requires much more research and understanding."
Siemsen, Roth, and Oliveira (2010)	*Organizational Research Methods*	"CMV refers to the shared variance among measured variables that arises when they are assessed using a common method."	Algebraic analysis, including extensive Monte-Carlo runs to test robustness.	"CMV can either **inflate or deflate** bivariate linear relationships [. . .]. With respect to multivariate linear relationships, [. . .] common method bias **generally decreases** when additional independent variables suffering from CMV are included [. . .]. [Q]uadratic and interaction effects **cannot** be artifacts of CMV."
Brannick, Chan, Conway, Lance, and Spector (2010)	*Organizational Research Methods*	"[M]ethod variance is an umbrella or generic term for invalidity of measurement. Systematic sources of variance that are not those of interest to the researcher are good candidates for the label 'method variance'."	Expert opinions from four scholars who have written about CMV: David Chan, James M. Conway, Charles E. Lance, and Paul E. Spector.	"Rather than considering method variance to be a plague, which, [. . .] leads inevitably to death (read: rejection of publication), method variance should be regarded in a **more refined way**. [. . .] [R]ather than considering method variance to be a general problem that afflicts a study, authors and reviewers should consider specific problems in **measurement** that affect the focal substantive inference."

Note. CMV = common method variance.

After the conceptualization by Podsakoff and Organ (1986), the term *common method variance* seemed to stick. Although the terms *CSB* and *common method bias* are sporadically used either as synonyms of CMV or to indicate the biases resulting from the presence of CMV. In addition, from a conceptual perspective, most studies seem to utilize similar definitions of CMV—indicating issues arising when variables are measured using the same method/source, typically a survey (e.g., Avolio, Yammarino, & Bass, 1991; Doty & Glick, 1998; Lindell & Whitney, 2001; Pace, 2010). Looking at the evidence used by these studies, we see a wide range of management and psychology topics (e.g., perceptions of jobs/work environments, appraisal, leadership outcomes), as well as wide range of research methods (e.g., expert opinions, algebraic analysis, literature review; for example, Avolio et al., 1991; Brannick, Chan, Conway, Lance, & Spector, 2010; Keeping & Levy, 2000; Podsakoff et al., 2003; Siemsen, Roth, & Oliveira, 2010; Spector, 1987). Interestingly enough, there is divergence in the core findings of these studies, which implies that the debate in management and psychology on whether or not CSB is a universal inflator of correlations in self-reported variables seems to be pointing toward a negative answer—it is not (Brannick et al., 2010; Spector, 2006).

Interesting examples of how this debate evolves over time are the studies of Spector (1987); Williams, Buckley, and Cote (1989); and Bagozzi and Yi (1990). The first author indicates that, based on previous articles with self-reported measures of jobs/work environment and affective reactions as well as own data on job satisfaction (JS), CMV is a nonissue (Spector, 1987). The second authors use the same data, but a different analysis, to illustrate that CMV is very much an issue (Williams et al., 1989). The third authors again use the same data, but a different analysis than number one and two, to illustrate that CMV is sometimes an issue—but not always (Bagozzi & Yi, 1990). We thus move from "no" to "yes" to "sometimes." Similar trends emerge in the other articles, with the consensus being that CSB is indeed an issue—but an issue that requires a deeper understanding of measurement problems and that cannot be considered a universal plague for all measures gathered from a common, self-reported source (e.g., Brannick et al., 2010; Pace, 2010; Podsakoff et al., 2003).

CSB in Public Administration Research

In contrast to the abundance of research studies as well as perspectives on CSB in management and psychology, public administration scholarship seems to lag behind. A search on the keywords "common method bias," "common source bias," or "common method variance" in some top public administration journals (*ROPPA*, *PAR*, *JPART*, *PAR*, *IPMJ*, and *The American Review of Public Administration*) indicates how little the field of public administration has written about the subject. Only four articles explicitly mention (some of) these keywords in their title (Andersen, Heinesen, & Pedersen, 2016; Favero & Bullock, 2015; Jakobsen & Jensen, 2015; Meier & O'Toole, 2013). Of those four articles, two focus on the presence of CSB when measuring performance through surveys (Andersen et al., 2016; Meier & O'Toole, 2013), whereas the other two center on the remedies that have been and/or could be applied

by public administration scholars to cope with CSB (Favero & Bullock, 2015; Jakobsen & Jensen, 2015). A paradox thus emerges where, on one hand, editorial policies in public administration journals are increasingly scrutinizing papers with potential CSB issues (i.e., see examples indicated in our "Introduction" section), while, on the other hand, very few public administration articles have actually studied the topic.

Meier and O'Toole (2013) use measurement theory to illustrate the theoretical underpinnings of CSB and test said theory by means of a data set from Texas schools that incorporates both self-reported measures of organizational performance as well as archival measures. The self-reported measures are drawn from a survey of school superintendents whereas the archival measures are drawn from a database, namely, the Academic Excellence Indicator System of the Texas Education Agency. Both measures center on the same concepts: Performance on the Texas Assessment of Knowledge and Skills (TAKS; that is, a student test) and performance on college-bound students (i.e., the number of college-ready students)—which are but a small subset of potential dependent variables, and specifically performance dimensions, prevalent in public administration scholarship. For the TAKS performance variables, the authors uncover 20 cases of false positives (i.e., a significant correlation is found with the perceptual measure but not with the archival) and three false negatives (i.e., a significant correlation is found with the archival measure but not with the perceptual) out of 84 tested regressions, which implies that 27% of the tested cases contain spurious results. For the college-bound performance variables, the authors uncover 31 false positives and 10 false negatives as well as one case where the relations were both significant—but opposite, which implies that 50% of these cases contain spurious results. Based on these findings, the authors conclude that CSB "is a serious problem when researchers rely on the responses of managers" for measuring organizational performance (Meier & O'Toole, 2013, p. 20). It is important to underscore that judgment of false positive/false negative is based on an uncritical acceptance of the archival measure with the implicit assumption about measurement validity of test scores.

Similarly, Andersen et al. (2016) combine survey data with archival data from Danish schools to test the associations between a set of independent variables (e.g., PSM and JS), teachers' self-report on their contribution to students' academic skills and archival sources of student performance. Again, the focus on student performance and self-reports on contribution to academic skills are but a subset of potential dependent variables, and specifically performance dimensions, investigated in public administration scholarship. The authors next compare several regression models between (a) PSM and performance, (b) intrinsic motivation (IM) and performance and (c) JS and performance. In (a), they find that PSM is significantly correlated with both the perceptual as well as three archival measures of performance. In (b), they find that IM is significantly correlated with the perceptual measure, but only with one out of three archival measures of performance—a pattern that repeats itself in (c). The authors argue that these "results illustrate that the highest associations in analyses using information from the same data source may not be the most robust associations when compared with results based on performance data from other sources" (Andersen et al., 2016, p. 74).

Jakobsen and Jensen (2015) offer further evidence on the existence of CSB by using an analysis of intrinsic work motivation and sickness absence drawn from a survey of Danish child care employees as well as archival data. In addition, they discuss a set of remedies to reduce CSB. First, in their analysis they construct two regression models—one with the perceptual measure of sickness absence and one with the archival measure. They find that intrinsic work motivation only has a significant effect on the perceptual measure and not on the archival, thus indicating issues of CSB. Second, they discuss how survey design, statistical tests and a panel data approach might reduce CSB. For the proposed remedies through survey design, they mostly summarize the advice of scholars from other domains on procedural remedies to cope with CSB (e.g., MacKenzie & Podsakoff, 2012; Podsakoff et al., 2012). For the statistical tests that can detect issues with CSB after data collection, they join authors from other fields (e.g., Chang et al., 2010; Richardson et al., 2009) in indicating that Harman's single factor test and the Confirmatory Factor Analysis marker technique are not enough. Finally, they describe how using panel data might minimize CSB—but simultaneously indicate that this method is not without its own flaws. The finding concerning the lack of statistical remedy to CSB was particularly startling to the chief editor of the journal in which the article was published (*IPMJ*) and resulted in an editorial policy where authors are discouraged to submit papers that suffer from CSB (Kelman, 2015).

In the final article, Favero and Bullock (2015) review the statistical remedies used by six articles published in the same journal (*JPART*) to identify the extent of CSB. They also use data from Texas schools as well as New York City schools to test the effectiveness of these remedies themselves. The six articles seem to be selected mainly from a convenience perspective as these fit the reviewed approaches to dealing with CSB. The dependent variables of the six articles cover a wide range of public administration outcomes—including Internet use, trust in government, citizen compliance, social loafing, JS, corruptibility, favorable workforce outcomes, board effectiveness, board-executive relations, and teacher turnover. Contrary to the wide scope of dependent variables present in the six articles, the dependent variables used by Favero and Bullock to test the effectiveness of the remedies engrained in these articles only focus on school violence and school performance (see Meier & O'Toole, 2013). Favero and Bullock find that when using the same statistical remedies to their data, none of the proposed statistical remedies satisfactorily address the underlying issue of CSB, and they conclude that the only reliable solution is to incorporate independent data sources in combination with surveys.

Conclusively, the literature review indicates that public administration, as a field, seems to lag behind management and psychology in its approach to CSB (see the contrast between Table 1 and the four reviewed public administration articles). Particularly potent are (a) the lack of counterarguments on the potential absence or limited impact of CSB in specific circumstances and (b) the narrow blinkered focus on specific dimensions of performance as dependent variables to prove the existence of CSB. Both of which seem to fuel editorial policies as well as review processes of major public administration journals on surveys as single data source. Similar to psychology

and management, we thus need some balance to the CSB debate within public administration and the next section aims to provide exactly that.

A Balanced Perspective on CSB

In this section, we present four arguments as to why CSB might have become somewhat of an urban legend within public administration that requires a nuanced and realistic perspective. Throughout these arguments, we draw on articles published in public administration journals as well as evidence derived from psychology and management studies.

Initial Claims of the CSB Threat Are Exaggerated

If all survey measures are influenced by CSB, evidence should be easy enough to find. Specifically, one would expect that measures drawn from the same survey source would be significantly correlated if these all suffer equally from CSB (Spector, 2006). We thus seek to refute the idea that CSB is "a universal inflator of correlations" by illustrating that the usage of "a self-report methodology is no guarantee of finding significant results, even with very large samples" (Spector, 2006, p. 224). To do so, we replicate Spector's (2006) test on public administration evidence by looking at correlation tables. We would like to point out that this test is aimed at identifying whether or not CMV always results in CSB (i.e., inflated correlations). Specifically, Spector argues that "[u]nless the strength of CMV is so small as to be inconsequential, [it] should produce significant correlations among all variables reported in such studies, given there is sufficient power to detect them" (p. 224).

We conducted a review of articles with "public service motivation" (PSM) in their title or keywords, published between 2007 and 2015 in the *Review of Public Personnel Administration* that present a correlation matrix between measures of specific variables (excluding demographic variables) and measures of PSM derived from the same survey. Six articles met our criteria. Characteristics of these articles and the subsequent correlations are presented in Table 2.

Were CSB indeed a universal inflator of correlations, we would expect to see evidence in the correlation matrices. Specifically, we would expect all correlations between PSM and other measures drawn from the same survey to be significant (Spector, 2006). Table 2 does not support this pattern. Out of the 36 reported correlations, 25 proved to be significant—indicating that more than 30% of the reported correlations were nonsignificant. Moreover, the average correlation per study was always <.30 and the average correlation across the six studies was .20, indicating a small to moderate correlation at best. Other examples show even more extreme patterns toward a lack of significant correlations. For instance, in their study on the relation between PSM and public values, Andersen, Jørgensen, Kjeldsen, Pedersen, and Vrangbæk (2013) use a cross-sectional survey with 501 individuals, including measures on PSM and public values. The sample size of 501 seems extensive enough to detect even limited amounts of CSB as correlations as small as .09 were statistically significant.

Table 2. Selected Articles on Public Service Motivation, Their Characteristics, and Correlations.

Article	Sample	Number of correlations tested[a]	Number of those correlations that were significant	Average correlation[b]
Quratulain and Khan (2015)	217 public servants	6	1	.07
van Loon (2015)	459 employees in people-changing organizations and 461 employees in people-processing organizations	8	6	.20
French and Emerson (2015)	272 employees of a local government	9	6	.22
Campbell and Im (2015)	480 ministry employees	6	5	.26
Ritz, Giauque, Varone, and Anderfuhren-Biget (2014)	569 public managers at the local level	3	3	.26
Vandenabeele (2014)	3.506 state civil servants	4	4	.20
	Sum	36	25	.20

[a]These are the correlations that were tested between public service motivation and other self-reported measures (excluding demographic characteristics) derived from the same survey.
[b]This is the average correlation (absolute value) in each study of all the identified correlations between public service motivation and other self-reported measures (excluding demographic characteristics) derived from the same survey.

However, only 11 out of 35 correlations proved to be significant (31%), and of those 11, three were .10 or less and none exceeded .20. Thus, both Table 2 and the findings of Andersen et al. (2013) hardly support the case of CSB as a particularly potent and universal inflator of correlations between variables derived from the same survey (Spector, 2006). This result is not that surprising when looking at previous management and psychology studies. Indeed, CMV has been shown to have no effect on correlations, an inflating effect and also a deflating effect—so concluding that correlations between variables from the same source are always inflated is not supported by empirical evidence (Conway & Lance, 2010).

Similar evidence and arguments—based on patterns of correlations—against CSB have been advanced in studies published in other public administration journals as well (e.g., Moynihan & Pandey, 2005; Pandey, Wright, & Moynihan, 2008). In a recent study, Fuller et al. (2016) reframe this question by asking a question about magnitude of CMV—At what level of CMV does CSB become a concern? Based on a simulation study, Fuller et al. conclude, "[f]or typical reliabilities, CMV would need to be on the

order of 70% or more before substantial concern about inflated relationships would arise. At lower reliabilities, CMV would need to be even higher to bias data" (p. 3197). The indicated threshold of 70% is based on the results of executing Harman's one-factor test and under the condition of typical scale reliabilities (i.e., Cronbach's α of .87-.90). If scale reliabilities are lower, CMV identified through Harman's one-factor test needs to be even higher to generate CSB. Hence, based on this argument Fuller et al. conclude that "today's reviewers may be asking more than is needed of authors in presenting evidence of a lack of CMB" (p. 3197). This observation is in line with evidence we have offered and may suggest a need to reframe the debate about CSB from a prima facie conclusion to a more circumspect and rigorous empirical basis for mounting such claims against studies using same data sources.

Initial Claims About CSB Threat Draw Upon Selective Evidence

In the literature review, we discussed four studies on CSB published in top public administration journals. These articles have, in many ways, delivered pioneering work about the impact of CSB on public administration scholarship. However, as the authors typically indicate themselves, these studies draw upon a particular set of evidence gathered in a particular context. Table 3 summarizes the characteristics of these four studies.

Table 3 illustrates that three out of four studies center on performance-related dependent variables. Two studies use U.S.-based samples, whereas the other two use Danish samples. Moreover, three out of four studies focus on an educational setting to test the impact of CSB. In addition, the incorporated measures of performance mostly center on one specific type of performance, namely, student test scores. The question that thus emerges is: Can the conclusions concerning CSB drawn from four studies centered on performance, education, student scores and the United States/Denmark be generalized toward public administration scholarship in general? Taking into account the insights from management and psychology studies (see Table 1), our answer is negative. Indeed, management and psychology studies have presented evidence that CSB might (a) deflate correlations (Siemsen et al., 2010), (b) depend on the type of variable that is measured (e.g., organizational performance measured through survey items or through archival data can be conceptually unrelated, see Wall et al., 2004), and (c) depend on the quality of the constructed survey and sample (e.g., complex questions and nonexpert respondents, see Podsakoff et al., 2003). So there is no "universal" truth when it comes to the impact of CSB on all variables gathered through the same survey.

Nevertheless, public administration scholars have used these four articles in their discussions on CSB. A citation analysis carried out on August 10, 2016, through Google Scholar indicates that, in total, these four studies have been cited by 172 articles in the past 3 years. The emerging question is "how" these studies were used: As reference to suggest the existence of CSB in all self-reported measures drawn from a single survey or as reference to justify use of specific performance outcome measures in specific contexts? Examples of both prevail. For instance, in their study

Table 3. Characteristics of Four Public Administration Articles Discussing Common Source Bias.

Article	Focus	Unit of analysis	Measures for the DV	Conclusion
Meier and O'Toole (2013)	CSB issues when measuring subjective organizational performance.	Texas schools	*Perceptual*: Perceived performance on student scores. *Archival*: Actual performance on student scores.	CSB problem in subjective measurement of organizational performance.
Andersen, Heinesen, and Pedersen (2016)	CSB issues when measuring subjective individual performance.	Danish lower secondary school teachers	*Perceptual*: Perceived contribution to students' academic skills. *Archival*: Actual performance on student scores.	CSB problem in subjective measurement of individual performance.
Jakobsen and Jensen (2015)	CSB issues in general self-reports—example sickness absence.	Danish child care employees	*Perceptual*: Perceptions of sickness absence. *Archival*: Archival numbers on sickness absence.	CSB problem in subjective measurement of sickness absence.
Favero and Bullock (2015)	Review of statistical remedies for CSB from 6 studies based on performance-related data.	New York City schools, Texas schools	*Perceptual*: Perceived performance on student scores and school violence. *Archival*: Actual performance on student scores and school violence.	Incorporate distinct source only reliable solution to CSB.

Note. DV = dependent variable; CSB = common source bias.

on the predictors of strategic-decision quality in Flemish student council centers, George and Desmidt (2016) use these articles to illustrate that CSB is particularly potent in self-reported measures of performance whereas their dependent variable, strategic-decision quality, might be less susceptible. Andersen, Kristensen, and Pedersen (2015) argue the necessity of using archival data for specifically measuring sickness absence in a Danish context to cope with CSB—which is in line with the evidence presented by Jakobsen and Jensen (2015). Jensen and Andersen (2015), in their study on prescription behavior of general practitioners (GP's), use these sources to argue that their choice of using register data to measure actual prescription behavior prevents CSB in their analysis—although the emergence of CSB when measuring prescription behavior by GP's is not the focus of the four studies. Finally, although Jakobsen and Jensen and Favero and Bullock (2015) argue against the usage of Harman's one-factor test to identify issues with CSB, studies referencing these articles still execute and report this test as a means of identifying CSB (e.g., George & Desmidt, 2016; Hsieh, 2016).

To conclude, when taking into account the importance of context in public administration (O'Toole & Meier, 2015) as well as the acknowledgment that not all self-reported variables might be equally prone to CSB (Fuller et al., 2016; Spector, 2006), we argue that the four studies are based on selective evidence from distinct populations and the case for broad generalizability of the CSB issue to different countries, public organizations, and variables is not merited.

Some Variables Are, by Their Very Nature, Perceptual

One of the reasons underlying the current performance-oriented evidence on CSB in public administration scholarship is the select group of scholars investigating this topic as well as their background. Specifically, authors such as Meier, O'Toole, and Andersen are well known for their studies of performance in public organizations—thus logically elucidating their interest for the effect of CSB on self-reported measures of performance (Andersen et al., 2016; Meier & O'Toole, 2013). As is also apparent from Table 1, a broader group of scholars with very diverse interests have investigated the topic of CSB within management and psychology—ranging from international business scholars (Chang et al., 2010) to industrial/organizational psychologists (Spector, 2006) to technology/operations management scholars (Siemsen et al., 2010). A far broader range of variables have thus been investigated within management and psychology journals on their predisposition toward CSB (Podsakoff et al., 2012)—and the emerging conclusion seems to be that when "both the predictor and criterion variables are capturing an individual's perceptions, beliefs, judgments, or feelings" surveys are appropriate measurement methods (Podsakoff et al., 2012, p. 549). Indeed, Meier and O'Toole (2013) also offer a list of variables less prone to CSB such as, "how managers spend their time, questions dealing with observable behavior, questions about environmental support, questions about a reactive strategy, and questions about managing in the network" (p. 447).

Human resource management (HRM) outcomes are also a potent example of variables in management and psychology that are, by their very nature, perceptual. Such outcomes include, for instance, JS and organizational commitment (Paauwe & Boselie, 2005). Investigating the determinants of such outcomes is no simple activity as these outcomes have been argued to unravel the black box "between HRM practices and policies on the one hand and the bottom-line performance of the firm on the other hand" (Paauwe, 2009, pp. 131-132). Indeed, a recent review of 36 empirical management and psychology studies published between 1995 and 2010 provides empirical evidence that human resource outcomes related to happiness and relationship can, at least partially, help explain the relation between HRM and organizational performance (Van De Voorde, Paauwe, & Van Veldhoven, 2012).

Within public administration, variables related to individuals' perceptions, beliefs, judgments or feelings have also received attention and are often argued to be antecedents of organizational performance-related outcomes (e.g., George & Desmidt, 2014; Poister, Pitts, & Edwards, 2010; Pollitt & Bouckaert, 2004). Examples include, for instance, studies on individual innovation in public organizations (e.g., Audenaert et al., 2016), perceived quality of strategic decisions (e.g., George & Desmidt, 2016) and HR outcomes such as JS, organizational commitment, and job involvement (e.g., Moynihan & Pandey, 2007). Taking into account the potential mediating role of such variables in explaining the public management-performance relation (Paauwe, 2009), editors, reviewers, and authors alike need to acknowledge their importance. This does not imply that CSB is not an issue in such studies, but it does imply that the self-reported nature of these variables cannot be the basis of prima facie conclusions that CSB makes such data unusable—especially if the authors followed a reasonable set of procedural remedies to minimize CSB (MacKenzie & Podsakoff, 2012).

Archival Data Can Be as Flawed as, or Even More Than, Self-Reported Data

Favero and Bullock (2015) indicate that using distinct data sources to measure dependent and independent variables in public administration scholarship might be the only way to avoid CSB. Such an approach is dubbed a multimethod approach (as opposed to a monomethod approach; Spector, 2006), but does not come without its own set of issues (Kammeyer-Mueller et al., 2010). When different data sources are used for the dependent and independent variables, a potential deficiency of measurement might emerge. Such a deficiency of measurement implies that designs with distinct data sources are susceptible to measurement error at the level of the database (Kammeyer-Mueller et al., 2010). For instance, imagine a dependent variable "student depression," which is rated by teachers and linked to a set of independent variables reported by students. Teachers are only experiencing a small part of the depression experienced by students, as the depression might by implicitly present, but not explicitly expressed in class (Courvoisier, Nussbeck, Eid, Geiser, & Cole, 2008). The student is simply the only person who can truly assess how he or she feels. Similarly, assessments of job

performance have shown low interrater agreement, which implies that raters (e.g., supervisor, colleague) might have a completely different assessment of job performance than the subject. Incorporating these other raters as distinct sources might solve issues of CSB, but results in a myriad of other problems (Lance, Hoffman, Gentry, & Baranik, 2008; Scullen, Mount, & Goff, 2000; Viswesvaran, Schmidt, & Ones, 2005). Concretely, distinct data sources result in downward biased correlations (i.e., the exact opposite of the assumed upward biased correlations in the presence of CSB; Kammeyer-Mueller et al., 2010).

An unfortunate side-effect of acceptance of CSB as a "mortal" threat to same source data is the relatively uncritical acceptance of data collected from distinct sources by researchers who do not necessarily do the due diligence to assess measurement validity of data from external sources. Whereas data collected from survey research using self-report leave measurement and construct validity in the researcher's hand, external sources of data, however, must be taken as such with no intentional effort guiding measurement validity and construct validity. Applied, for instance, to the distinct data source used by Meier and O'Toole (2013) as well as Favero and Bullock (2015; that is, student scores on the TAKS), what guarantee do we have that these data provide a valid and reliable measure of the different performance dimensions underlying Texas schools—as opposed to being one component of the bigger performance picture?

Not only do news reports often call into question integrity and validity of test scores (e.g., Anderson, 2016; Axtman, 2005; Wong, 2016), but education policy scholars also provide in-depth analysis that calls into question measurement validity of these scores as performance measures. Indeed, Booher-Jennings (2005) illustrates how teachers within the Texas school system committed "educational triage" to ensure an apparent improvement of test scores. Such triage implies, for instance, that students who might be "high-risk" in their potential to succeed for test scores are referred for special education thus removing them as a risk to the school test scores. Good student test scores might thus be a result of "selection" within schools—rather than true educational performance (Booher-Jennings, 2005). Similarly, Haney (2000) also substantiates this point of selection by teachers and schools to positively influence students test scores in Texas schools. He argues that the test scores resulted in more students from minorities being held back a grade as well as more students receiving the label of "special education" to avoid these students from taking the student test or for being counted in the schools accountability ratings (Haney, 2000). If such practices indeed underlie a database measuring "student performance," one might rightfully argue that such a measure of performance simply neither fits core public values nor offers a valid measure of school performance.

Furthermore, issues with selecting distinct sources are particularly relevant to public administration scholars due to potential transparency concerns with archival data sources (Tummers, 2016). Survey-based studies in public administration are expected to and typically do present a high standard of disclosure about the measurement procedures, instruments, and so on (Lee et al., 2012). Such disclosure allows other authors to replicate self-reported scales in different contexts to identify how biases might creep in (Jilke, Meuleman, & Van de Walle, 2015). The same standards might not

apply to data gathered from archival data sources because researchers do not have all insights into the intricacies underlying these databases. After all, a researcher has full control over the development, collection and analysis process of a survey whereas archival sources are, in their nature, already developed and collected by other parties. There thus seems to be a black box on the standards used by such archival data sources, which might be less scrutinized by reviewers because reviewers—primed to elevate CSB over other measurement validity concerns—settle for and readily accept the validity of an external data source. The lack of transparency on archival data sources is, however, not a trivial issue. Taking into account the previously indicated issues underlying a distinct source design (i.e., measurement error), data drawn from such sources might present very narrow interpretations of dependent variables or even interpretations strongly influenced by events and biases that indicate measurement error at the level of the database and not necessarily at the level of the unit of analysis (Kammeyer-Mueller et al., 2010).

Conclusion

In conclusion, we provide a decision guide on how to reasonably deal with CSB and offer final thoughts on our balance-seeking approach and its limitations. Our decision guide by no means seeks to diminish the role of CSB or the need to assess and control for it—but rather offers balance and reasonableness to authors taking into account the specific nature of public administration research. The guide is aimed at researchers who consider (or are constrained to) using a survey as single data source for their research. Figure 1 presents the subsequent flowchart underlying our proposed decision guide.

The first question to ask is whether or not a survey will be used as a single data source. If not, this does not exempt scholars from following the necessary standards. Indeed, as is apparent in our fourth argument, a distinct data source can result in a myriad of issues and authors should address measurement validity concerns. Among other things, authors need to report on the reliability of the distinct sources, the specific dimensions of a variable measured by the distinct source as well as the openness and transparency of the data and its origin (Kammeyer-Mueller et al., 2010; Tummers, 2016). Nevertheless, if a distinct data source is both available as well as relevant, we would still advise authors to incorporate it into their analysis, perhaps to complement measures gathered from a single survey (e.g., see Kroll, Neshkova, & Pandey, 2017). Moreover, authors need not limit themselves to finding distinct data sources for their dependent variable. Indeed, a distinct data source can also be used to measure the independent variable of a study. For instance, if one wants to study the impact of training on JS, the number of training hours registered in a database can be correlated to survey measures of JS.

If a distinct data source is not available or not relevant, the next question to ask is: Which type of dependent and independent variable is measured through the single survey? Do these variables pertain to perceptions or beliefs of individuals that are, by nature, only measurable to surveys (e.g., HR outcomes)—as reported in our third

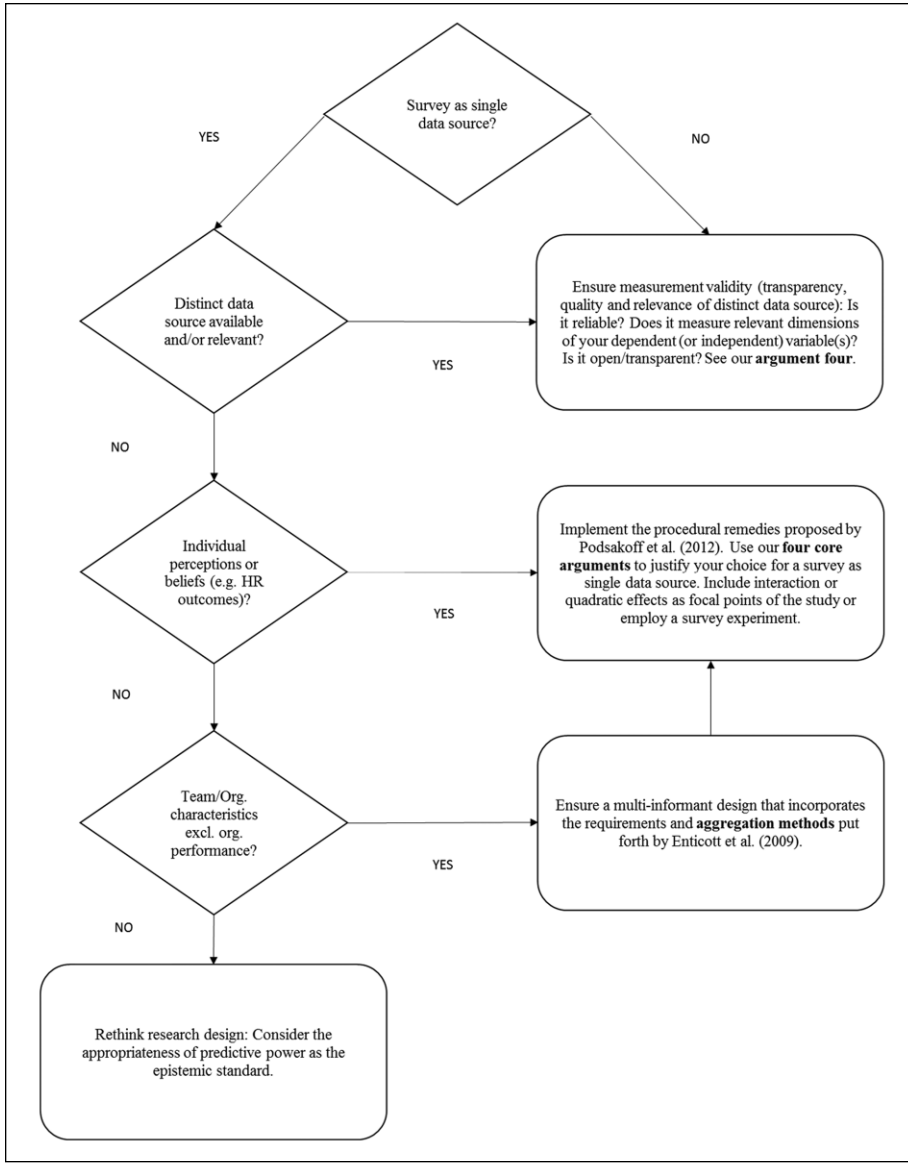

Figure 1. Flowchart underlying the usage of surveys in public administration.

argument (Van De Voorde et al., 2012)? If yes, authors should use the procedural remedies offered by Podsakoff et al. (2012) when devising and distributing their survey. In addition, they could consider four arguments discussed earlier to justify the inclusion of these self-reported variables: (a) Are the variables strongly correlated to all other

variables gathered from the same survey (identified through correlation tables) and is CMV high enough to generate CSB (identified through Harman's one-factor test; Fuller et al., 2016; Spector, 2006)? Although Harman's one-factor test has been widely criticized, recent empirical evidence has indicated that "the most commonly used post-hoc approach to managing CMV—Harman's one-factor test—can detect biasing levels of CMV under conditions commonly found in survey-based marketing research" (Fuller et al., 2016, p. 3197). The test can thus be an important (but not the only) argument of a balanced discussion on potential CSB issues in public administration papers. (b) Are the variables part of the set of CSB-susceptible variables included in the four studies of CSB within public administration or of other CSB studies within management and psychology? (c) Are the variables by nature perceptual such as typical HR outcomes that can only be measured through surveys? (d) Are other data sources nonexistent, irrelevant, or of poor quality, and if a survey is used, are multiple items used to measure variables and is the scale reliability (i.e., Cronbach's α) acceptable (Fuller et al., 2016)? We also encourage authors to include interaction or quadratic effects as core element of papers because these effects cannot be the product of CSB (Siemsen et al., 2010). Finally, the self-report nature of a dependent variable is less of a problem if a survey experiment is used. Such an experiment randomly assigns a specific independent variable to control and treatment groups—and thus allows an unbiased assessment of the impact of this independent variable on the self-reported dependent variable. Although the data (independent and dependent variables) come from the same source and remain subject to CSB, random assignment and experimental manipulation provide assurance about the most common weakness associated with CSB (endogeneity threat to causal inference).

If the variables measured through the same survey are not an individual's perception or belief of an individual-level variable—but rather focus on characteristics of teams (e.g., team performance, team dynamics) or organizations (e.g., enhanced managerial capacity, improved organizational communication) but excluding organizational performance—we encourage the authors to incorporate a multi-informant design that includes perspectives from different hierarchies within the organization and devise a relevant aggregation method (for advice, see Enticott et al., 2009). After having done that, the same set of remedies apply (i.e., implement procedural remedies, use our four arguments to justify design and include interaction or quadratic effects as focal study point).

Finally, we would not recommend authors to include measures of organizational performance together with specific independent variables in a single survey—the evidence points toward significant issues with these types of self-reports. Another research design might be considered in this stage. We believe that following the recommendations of our decision guide offers a reasonable and balanced approach to dealing with CSB in public administration scholarship—an approach which we encourage reviewers and editors to keep in mind when assessing the papers.

Neither do we advocate for single survey as the data source for analysis, nor do we believe that the CSB case against single-source data, built in recent public administration literature, is as robust as it seems at first blush. Before we elaborate more on our

goal of adding balance to the CSB discourse, we want to clearly note a key limitation of the use of a single data source—endogeneity as a result of unmeasured confounders. Cross-sectional survey-based research is of limited use in drawing causal conclusions, and CSB does pose endogeneity concerns. But it is also simplistic to believe that it is possible to use external data to "kill two birds with one stone" (endogeneity and CSB). If the imperative is to get the better of endogeneity and to make robust causal inferences using observational data, there is a need to seriously consider the "identification strategy" or elements of research design that help with making causal inferences (Angrist & Krueger, 1999). Despite the empirical leverage offered by the "identification revolution," there are some concerns about its negative effects on theoretical development (Huber, 2013).

We conclude with Yin and Yang, the two implacably and eternally opposing forces in Eastern Philosophy referenced in our title, and what the relationship between the two means for consideration of CSB in public administration literature. The injunction in Eastern Philosophy is about balancing these forces and often the guiding question is: Balance for what purposes? This is a question we must ask with CSB as well. What is it that we gain or lose by accounting for CSB and how does it relate to the state of science in public administration? Although economics, among social sciences, gets called out for a dubious focus on prediction, public administration scholarship is not far behind in embracing predictive power as an epistemic standard. There is a need to be more reflexive, and it is incumbent on us to ask whether the state of science in public administration is comparable with natural sciences like physics, with well-regarded explanatory and predictive theories that drive calls for correction of errors to obtain ever-precise predictions? Or is it that the state of science in public administration barely affords us the opportunity to identify relevant explanatory variables and it is hard to make defensible claims about precise effects? Are reasonableness and balance—as opposed to exactness—perhaps the keywords to look for when dealing with CSB?

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

References

Andersen, L. B., Heinesen, E., & Pedersen, L. H. (2016). Individual performance: From common source bias to institutionalized assessment. *Journal of Public Administration Research and Theory*, *26*, 63-78.

Andersen, L. B., Jørgensen, T. B., Kjeldsen, A. M., Pedersen, L. H., & Vrangbæk, K. (2013). Public values and public service motivation: Conceptual and empirical relationships. *The American Review of Public Administration*, *43*, 292-311.

Andersen, L. B., Kristensen, N., & Pedersen, L. H. (2015). Documentation requirements, intrinsic motivation, and worker absence. *International Public Management Journal*, *18*, 483-513.

Anderson, L. (2016, July). Who's who in the EPISD cheating scheme. *El Paso Times*. Retrieved from http://www.elpasotimes.com/story/news/education/episd/2016/07/30/whos-who-episd-cheating-scheme/87618624

Angrist, J. D., & Krueger, A. B. (1999). Empirical strategies in labor economics. In O. C. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 3). (pp. 1277-1366). North Holland, The Netherlands.

Audenaert, M., Decramer, A., George, B., Verschuere, B., & Van Waeyenberg, T. (2016). When employee performance management affects individual innovation in public organizations: The role of consistency and LMX. *The International Journal of Human Resource Management*, DOI: 10.1080/09585192.2016.1239220

Avolio, B. J., Yammarino, F. J., & Bass, B. M. (1991). Identifying common methods variance with data collected from a single source: An unresolved sticky issue. *Journal of Management*, *17*, 571-587.

Axtman, C. (2005, January). When tests' cheaters are the teachers: Probe of Texas scores on high-stakes tests is the latest case in series of cheating incidents. *Christian Science Monitor*. Retrieved from http://www.csmonitor.com/2005/0111/p01s03-ussc.html

Bagozzi, R. P., & Yi, Y. (1990). Assessing method variance in multitrait-multimethod matrices: The case of self-reported affect and perceptions at work. *Journal of Applied Psychology*, *75*, 547-560.

Booher-Jennings, J. (2005). Below the bubble: "Educational triage" and the Texas Accountability System. *American Educational Research Journal*, *42*, 231-268.

Bozeman, B., & Su, X. (2015). Public service motivation concepts and theory: A critique. *Public Administration Review*, *75*, 700-710.

Brannick, M. T., Chan, D., Conway, J. M., Lance, C. E., & Spector, P. E. (2010). What is method variance and how can we cope with it? A panel discussion. *Organizational Research Methods*, *13*, 407-420.

Campbell, J. P. (1982). Editorial: Some remarks from the outgoing editor. *Journal of Applied Psychology*, *67*, 691-700.

Campbell, J. W., & Im, T. (2015). PSM and turnover intention in public organizations: Does change-oriented organizational citizenship behavior play a role? *Review of Public Personnel Administration*, *36*, 323-346.

Chang, S.-J., van Witteloostuijn, A., & Eden, L. (2010). From the editors: Common method variance in international business research. *Journal of International Business Studies*, *41*, 178-184.

Colquitt, J. A., & Ireland, R. D. (2009). From the editors: taking the mystery out of Amj's reviewer evaluation form. *Academy of Management Journal*, *52*, 224-228.

Conway, J. M., & Lance, C. E. (2010). What reviewers should expect from authors regarding common method bias in organizational research. *Journal of Business and Psychology*, *25*, 325-334.

Courvoisier, D. S., Nussbeck, F. W., Eid, M., Geiser, C., & Cole, D. A. (2008). Analyzing the convergent and discriminant validity of states and traits: Development and applications of multimethod latent state-trait models. *Psychological Assessment*, *20*, 270-280.

Doty, D. H., & Glick, W. H. (1998). Common methods bias: Does common methods variance really bias results? *Organizational Research Methods*, *1*, 374-406.

Enticott, G., Boyne, G. A., & Walker, R. M. (2009). The use of multiple informants in public administration research: Data aggregation using organizational echelons. *Journal of Public Administration Research and Theory*, *19*, 229-253.

Favero, N., & Bullock, J. B. (2015). How (not) to solve the problem: An evaluation of scholarly responses to common source bias. *Journal of Public Administration Research and Theory*, 25, 285-308.

Feeney, M. K., & Bozeman, B. (2009). Stakeholder red tape: Comparing perceptions of public managers and their private consultants. *Public Administration Review*, *69*, 710-726.

French, P. E., & Emerson, M. C. (2015). One size does not fit all: matching the reward to the employee's motivational needs. *Review of Public Personnel Administration*, *35*, 82-94.

Fuller, C. M., Simmering, M. J., Atinc, G., Atinc, Y., & Babin, B. J. (2016). Common methods variance detection in business research. *Journal of Business Research*, *69*, 3192-3198.

George, B., & Desmidt, S. (2014). A state of research on strategic management in the public sector: An analysis of the empirical evidence. In P. Joyce & A. Drumaux (Eds.), *Strategic management in public organizations: European practices and perspectives* (pp. 151-172). New York, NY: Routledge.

George, B., & Desmidt, S. (2016). Strategic-decision quality in public organizations: An information processing perspective. *Administration & Society*. DOI: 10.1177/0095399716647153

George, B., Desmidt, S., & De Moyer, J. (2016). Strategic decision quality in Flemish municipalities. *Public Money & Management*, *36*, 317-324.

Grimmelikhuijsen, S., Tummers, L., & Pandey, S. K. (2016). Promoting state-of-the-art methods in public management research. *International Public Management Journal*, *20*(1), 7-13.

Haney, W. (2000). The myth of the Texas miracle in education. *Education Policy Analysis Archives*, *8*, 41.

Harrison, D. A., McLaughlin, M. E., & Coalter, T. M. (1996). Context, cognition, and common method variance: Psychometric and verbal protocol evidence. *Organizational Behavior and Human Decision Processes*, *68*, 246-261.

Hsieh, J. Y. (2016). Spurious or true? An exploration of antecedents and simultaneity of job performance and job satisfaction across the sectors. *Public Personnel Management*, *45*, 90-118.

Huber, J. (2013). Is theory getting lost in the "identification revolution."http://themonkeycage.org/2013/06/is-theory-getting-lost-in-the-identification-revolution/.

Jakobsen, M., & Jensen, R. (2015). Common method bias in public management studies. *International Public Management Journal*, *18*, 3-30.

Jensen, U. T., & Andersen, L. B. (2015). Public service motivation, user orientation, and prescription behaviour: Doing good for society or for the individual user? *Public Administration*, *93*, 753-768.

Jilke, S., Meuleman, B., & Van de Walle, S. (2015). We need to compare, but how? Measurement equivalence in comparative public administration. *Public Administration Review*, *75*, 36-48.

Kaiser, F. G., Schultz, P. W., & Scheuthle, H. (2007). The theory of planned behavior without compatibility? Beyond method bias and past trivial associations. *Journal of Applied Social Psychology*, *37*, 1522-1544.

Kammeyer-Mueller, J., Steel, P. D. G., & Rubenstein, A. (2010). The other side of method bias: The perils of distinct source research designs. *Multivariate Behavioral Research*, *45*, 294-321.

Kaplan, A. (1964). *The conduct of inquiry: Methodology for behavioral science*. San Francisco, CA: Chandler.

Keeping, L. M., & Levy, P. E. (2000). Performance appraisal reactions: Measurement, modeling, and method bias. *Journal of Applied Psychology*, *85*, 708-723.

Kelman, S. (2015). Letter from the editor. *International Public Management Journal*, *18*, 1-2.

Kroll, A., Neshkova, M. I., & Pandey, S. K. (2017). Spillover effects from customer to citizen orientation: How performance management reforms can foster public participation. *Administration & Society*. Advance online publicaiton.

Lance, C. E., Hoffman, B. J., Gentry, W. A., & Baranik, L. E. (2008). Rater source factors represent important subcomponents of the criterion construct space, not rater bias. *Human Resource Management Review*, *18*, 223-232.

Lee, G., Benoit-Bryan, J., & Johnson, T. P. (2012). Survey research in public administration: Assessing mainstream journals with a total survey error framework. *Public Administration Review*, *72*, 87-97.

Lee, G., & Choi, D. L. (2016). Does public service motivation influence the college students' intention to work in the public sector? Evidence from Korea. *Review of Public Personnel Administration*, *36*, 145-163.

Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology*, *86*, 114-121.

MacKenzie, S. B., & Podsakoff, P. M. (2012). Common method bias in marketing: Causes, mechanisms, and procedural remedies. *Journal of Retailing*, *88*, 542-555.

Meier, K. J., & O'Toole, L. J. (2013). Subjective organizational performance and measurement error: Common source bias and spurious relationships. *Journal of Public Administration Research and Theory*, *23*, 429-456.

Moynihan, D. P., & Pandey, S. K. (2005). Testing how management matters in an era of government by performance management. *Journal of Public Administration Research and Theory*, *15*, 421-439.

Moynihan, D. P., & Pandey, S. K. (2007). Finding workable levers over work motivation: Comparing job satisfaction, job involvement, and organizational commitment. *Administration & Society*, *39*, 803-832.

O'Toole, L. J., & Meier, K. J. (2015). Public management, context, and performance: In quest of a more general theory. *Journal of Public Administration Research and Theory*, *25*, 237-256.

Paauwe, J. (2009). HRM and performance: Achievements, methodological issues and prospects. *Journal of Management Studies*, *46*, 129-142.

Paauwe, J., & Boselie, P. (2005). HRM and performance: What next? *Human Resource Management Journal*, *15*, 68-83.

Pace, V. L. (2010). Method variance from the perspectives of reviewers: Poorly understood problem or overemphasized complaint? *Organizational Research Methods*, *13*, 421-434.

Pandey, S. K., & Marlowe, J. (2015). Assessing survey-based measurement of personnel red tape with anchoring vignettes. *Review of Public Personnel Administration*, *35*, 215-237.

Pandey, S. K., Pandey, S., & Van Ryzin, G. G. (2016). Prospects for experimental approaches to research on bureaucratic red tape. In J. Oliver, S. Jilke, & G. G. Van Ryzin (Eds.), *Experiments in public management research: Challenges and contributions*. Cambridge, UK: Cambridge University Press. DOI: 10.5465/AMBPP.2016.11116

Pandey, S. K., Wright, B. E., & Moynihan, D. P. (2008). Public service motivation and interpersonal citizenship behavior in public organizations: Testing a preliminary model. *International Public Management Journal*, *11*, 89-108.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*, 879-903.

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. *Annual Review of Psychology*, *63*, 539-569.

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management*, *12*, 531-544.

Poister, T. H., Pasha, O. Q., & Edwards, L. H. (2013). Does performance management lead to better outcomes? Evidence from the U.S. Public Transit Industry. *Public Administration Review*, *73*, 625-636.

Poister, T. H., Pitts, D. W., & Edwards, L. H. (2010). Strategic management research in the public sector: A review, synthesis, and future directions. *The American Review of Public Administration*, *40*, 522-545.

Pollitt, C., & Bouckaert, G. (2004). *Public management reform: A comparative analysis*. Oxford, UK: Oxford University Press.

Quratulain, S., & Khan, A. K. (2015). Red tape, resigned satisfaction, public service motivation, and negative employee attitudes and behaviors: Testing a model of moderated mediation. *Review of Public Personnel Administration*, *35*, 307-332.

Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A tale of three perspectives: Examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods*, *12*, 762-800.

Ritz, A., Giauque, D., Varone, F., & Anderfuhren-Biget, S. (2014). From leadership to citizenship behavior in public organizations: When values matter. *Review of Public Personnel Administration*, *34*, 128-152.

Scullen, S. E., Mount, M. K., & Goff, M. (2000). Understanding the latent structure of job performance ratings. *Journal of Applied Psychology*, *85*, 956-970.

Siemsen, E., Roth, A., & Oliveira, P. (2010). Common method bias in regression models with linear, quadratic, and interaction effects. *Organizational Research Methods*, *13*, 456-476.

Spector, P. E. (1987). Method variance as an artifact in self-reported affect and perceptions at work: Myth or significant problem? *Journal of Applied Psychology*, *72*, 438-443.

Spector, P. E. (2006). Method variance in organizational research: Truth or urban legend? *Organizational Research Methods*, *9*, 221-232.

Tummers, L. G. (2016). Moving towards an open research culture in public administration: *Public Administration Review*. https://publicadministrationreview.org/speak-your-mind-article/

Vandenabeele, W. (2014). Explaining public service motivation: The role of leadership and basic needs satisfaction. *Review of Public Personnel Administration*, *34*, 153-173.

Van De Voorde, K., Paauwe, J., & Van Veldhoven, M. (2012). Employee well-being and the HRM-organizational performance relationship: A review of quantitative studies. *International Journal of Management Reviews*, *14*, 391-407.

van Loon, N. M. (2015). Does context matter for the type of performance-related behavior of public service motivated employees? *Review of Public Personnel Administration*. DOI: 10.1177/0734371X15591036

Verschuere, B., Beddeleem, E., & Verlet, D. (2014). Determinants of innovative behaviour in Flemish nonprofit organizations: An empirical research. *Public Management Review*, *16*, 173-198.

Viswesvaran, C., Schmidt, F. L., & Ones, D. S. (2005). Is there a general factor in ratings of job performance? A meta-analytic framework for disentangling substantive and error influences. *Journal of Applied Psychology*, *90*, 108-131.

Wall, T. D., Michie, J., Patterson, M., Wood, S. J., Sheehan, M., Clegg, C. W., & West, M. (2004). On the validity of subjective measures of company performance. *Personnel Psychology*, *57*, 95-118.

Williams, L. J., Buckley, R. M., & Cote, J. A. (1989). Lack of method variance in self-reported affect and perceptions at work: Reality or artifact? *Journal of Applied Psychology*, *74*, 462-468.

Wong, A. (2016, April). Why would a teacher cheat? *The Atlantic*. Retrieved from https://www.theatlantic.com/education/archive/2016/04/why-teachers-cheat/480039/

Author Biographies

Bert George, PhD, is assistant professor in public management at Erasmus University Rotterdam, the Netherlands. His research focuses on the decision-making impact of strategic planning and performance measurement in public organizations, both from an observational and experimental perspective.

Sanjay K. Pandey, PhD, is professor and Shapiro Chair of public policy and public administration at George Washington University, USA. His research focuses on public management and deals with questions central to leading and managing public organizations.