PAR PUBLIC ADMINISTRATION REVIEW

Geon Lee
Seoul National University
Jennifer Benoit-Bryan
Timothy P. Johnson
University of Illinois at Chicago

Survey Research in Public Administration: Assessing Mainstream Journals with a Total Survey Error Framework

Survey research is a common tool for assessing public opinions, perceptions, attitudes, and behaviors for analyses in many social science disciplines. Yet there is little knowledge regarding how specific elements of survey research methodology are applied in practice in public administration. This article examines five mainstream public administration journals over an eight-year period regarding current methodological practice, organized around the total survey error framework. The findings show that survey research in the field of public administration features mainly small-scale studies, heavy reliance on a single data collection mode, questionable sample selection procedures, and suspect sample frame quality. Survey data largely are analyzed without careful consideration of assumptions or potential sources of error. An informed evaluation of the quality of survey data is made more difficult by the fact that many journal articles do not detail data collection procedures. This study concludes with suggestions for improving the quality and reporting of survey research in the field.

Since McCurdy and Cleary (1984) first discussed quality issues in doctoral dissertations in public administration, a number of other scholars have explored the quality of research methodologies used in the field (e.g., Box 1992; Cleary 1992, 2000; Forrester and Watson 1994; Houston and Delevan 1990; Perry and Kraemer 1986; Stallings 1986; Stallings and Ferris 1988; White 1986; White, Adams, and Forrester

Editor's Note: This article was accepted by Richard Stillman during his tenure as editor-in-chief. I am pleased that Richard had the foresight to accept it. I concur with the authors' findings, and Michael McGuire and I will use this article when making decisions about manuscripts using survey research. We will look carefully at the quality of survey research in manuscripts that are submitted to *PAR* both before we decide to review them and before we accept manuscripts employing survey methods. I encourage potential contributors to *PAR* to take seriously the authors' suggestions for improvements in survey research.

JLP

1996). Most of these reviews have examined quantitative research, and among such methods currently available, survey research is one of the most widely recognized and applied in public administration (Enticott, Boyne, and Walker 2009).

Survey research is a systematic data collection methodology in which samples are drawn, respondents are interviewed, and data are analyzed in order to extrapolate to a population of interest. The survey instrument allows researchers to assess, with a small sample, population attitudes, perceptions, and opinions about particular social issues, as well as factual knowledge (Swidorski 1980). For this efficiency, survey research has gained popularity in many academic disciplines (Folz 1996). In addition to academic domains, practitioners at all levels of government increasingly have turned to these techniques to measure citizen needs and feedback (Daneke and Klobus-Edwards 1979).

With its growing popularity, the demand for survey research has been increasing in professional fields as well as in academia. No longer considered merely a method, survey research now is recognized as an independent academic discipline in the United States. Largely in acknowledgment of the importance of continually developing the rigor of this methodology, most university programs in survey research methodology emphasize reducing survey errors in order to maximize the congruency between sample estimates and population parameters.

Although survey research is an important tool, there are no reviews available that examine the conduct and reporting of survey research in the empirical literature in public administration. The purpose of this article is to review the survey research methods now employed by researchers in public administration. Specifically, we aim to document the current practice of survey research within our field and to evaluate the degree to which potential sources of survey error are considered. In doing so, we will assess data collection procedures,

Geon Lee is a BK21 researcher in the Graduate School of Public Administration at Seoul National University, South Korea. His research interests are survey research methods, human resource management, public sector organizational behavior, and public service motivation and ethics.
E-mail: glee153@gmail.com

Jennifer Benoit-Bryan is a doctoral candidate in public administration and a research associate in the Survey Research Laboratory at the University of Illinois at Chicago. Her research interests are survey research methods, digital inclusion, transparency in government, and local government management.
E-mail: jbenoi2@uic.edu

Timothy P. Johnson is a professor of public administration and director of the Survey Research Laboratory at the University of Illinois at Chicago. His primary area of research is survey error.
E-mail: timj@uic.edu

Public Administration Review,
Vol. 72, Iss. 1, pp. 87–97. © 2011 by The American Society for Public Administration.
DOI: 10.111/j.1540-6210.2011.02482.x.

analytic issues, and reporting practices and offer thoughts regarding standards for professional practice.

> The purpose of this article is to review the survey research methods now employed by researchers in public administration.

Review of Public Administration Research

Research reviews in public administration first began with investigations of the quality of doctoral dissertations and subsequently were extended to debates over the quality of scholarly journal articles. Through these review assessments of public administration research over the past 25 years, public administration scholars have diagnosed the state and rigor of contemporary research and suggested constructive ideas for improving research quality.

McCurdy and Cleary (1984) first evaluated the quality of 142 public administration doctoral dissertations written in 1981 using six criteria: research purpose, validity, testing of theory, causal relationships, importance of the topic, and whether the topic was considered cutting-edge. They found that the majority of dissertations failed to meet these basic criteria and claimed that the quality of public administration research lagged that of mainstream social science research. These findings were confirmed by White (1986), who evaluated 305 dissertations using similar criteria. In addition, White found that a large amount of dissertation research never was published in journals, thereby limiting the contribution of dissertations to knowledge development. Both studies expressed serious concerns about the lack of rigorous research methodology in the field.

Issues regarding the research quality of doctoral dissertations subsequently directed scholarly attention to the quality of research published in professional scholarly journals. Perry and Kraemer (1986) and Stallings and Ferris (1988) investigated research articles published in *Public Administration Review* to examine the quality of research methodology for professional scholarly works. Both studies identified a need for more advanced statistical techniques and improvement of methodology in the work published in that journal.

Subsequently, Houston and Delevan (1990) examined a broader variety of public administration journals. These authors argued that research in public administration was underfunded, nonempirical, and not oriented toward theory testing. They went on to argue that the research designs employed in published journal articles were monotonous and lacked rigor. They recommended increased utilization of more rigorous research designs, including quasi-experiments and experiments. The authors concluded that low-quality research in public administration might be attributable in part to poor research training in public administration programs (see also Adams and White 1995; Cozzetto 1994; DeLorenzo 2000; McCurdy and Cleary 1984; Stallings 1986; White 1986).

Cozzetto (1994) narrowed his focus to address the quality of quantitative research articles in order to assess their level of statistical sophistication. He found that 40 percent of sampled articles incorrectly used statistical techniques and 83 percent showed a lack of methodological sophistication. In another assessment, Wright and his colleagues (2004) also emphasized quantitative research, raising issues of potential biases in the process of data collection and inadequate information on measurement in journal articles.

While invaluable, these prior reviews necessarily have been very broad in scope and thus unable to examine specific details of the methodologies examined. Researchers in other fields have undertaken more focused analyses specific to the use of survey research—for example, the review by Pinsonneault and Kraemer (1993) of survey research in the field of management information systems—yet this type of study has not yet been undertaken in the field of public administration. This study will attempt to fill that gap by focusing on survey research methodology in order to provide a more detailed assessment of its applications within the public administration literature. We begin by presenting a brief overview of a commonly accepted framework for understanding survey-related error.

Sources of Error in Survey Research

The total survey error (TSE) framework addresses all possible sources of error that can bias survey findings, including sampling error, coverage error, nonresponse error, measurement error, and processing error (Groves 1987, 1989). The last four are classified as nonsampling errors. Sampling error arises when the entire target population, or universe, is not selected and decreases as the sample size increases. Thus, the best way to minimize sampling error is to increase the size of random samples (while also controlling for other sources of survey error). Coverage error takes place when a sample frame does not fully represent the total population sampled, which results in selection bias, leading to generalizability issues. Nonresponse error occurs when there are systematic differences in responses between respondents and total sampled persons; enhancing response rates is generally considered a good strategy to minimize this type of error (Fowler 2002).

The sources of measurement error are many, including interviewer, respondent, questionnaire, and question wording. In face-to-face and telephone surveys, an interviewer can inadvertently influence respondents to answer in a certain direction, generating interviewer variance in which average responses to a particular variable varies across interviewers (Tourangeau, Rips, and Rasinski 2000). Typical examples of measurement error attributable to respondents might include inaccurate answers to retrospective questions because of recall or retrieval problems, difficulty interpreting question meaning, or difficulty mapping responses onto the available answer options (Golden 1992; Tourangeou, Rips, and Rasinski 2000). Likewise, social desirability demands may systematically bias responses to some survey questions (Fisher 1993; Johnson and Van de Vijver 2003). Poorly designed questionnaires that are ambiguous or overly complicated also make it difficult for respondents to comprehend and answer adequately (Holbrook, Cho, and Johnson 2006). Further, some question the assumption that survey questions can correctly measure opinions and beliefs (Zaller and Feldman 1992). Processing error tends to arise in postdata collection procedures such as data coding, editing, weighting construction, and estimation procedures (Groves et al. 2004).

Quality survey data can be obtained when TSE is minimized. More importantly, increased bias and variance attributable to survey errors may negatively affect analytical results—particularly, overestimation (or underestimation) of descriptive statistics and biased estimation or type II errors in regression analyses (Biemer and Lyberg 2003). For instance, survey measurement error attributable to social desirability may establish spurious relationships between variables in regression analysis (Fisher 1993; Moorman and Podsakoff 1992; Zerbe and Paulhus 1987). In addition, selection bias may influence analytical results in survey research. Enticott, Boyne, and Walker (2009), for example, show that a sample that contains multiple informants provides more accurate measures of organizational properties than can be obtained with only a single informant. While important as a research methodology, it is also clear that survey research can generate unreliable data that, in turn, may generate biased analytical results unless the variety of errors that commonly surround the research process are controlled properly. Next, we apply the TSE perspective to evaluate current survey practices in the public administration literature.

While important as a research methodology, it is also clear that survey research can generate unreliable data that, in turn, may generate biased analytical results unless the variety of errors that commonly surround the research process are controlled properly.

Data and Methods

We examined five peer-reviewed journals listed in the Social Science Citation Index: the *American Review of Public Administration, Administration & Society,* the *Journal of Public Administration Research and Theory, Public Administration,* and *Public Administration Review.* These five journals were chosen because they are often employed to represent the mainstream public administration research literature (see Brower, Abolafia, and Carr 2000; Forrester and Watson 1994; Lan and Anders 2000; Wright, Manigault, and Black 2004).[1]

We reviewed these journals for the eight-year period from 2000 to 2007, selecting 264 articles reporting either primary or secondary survey data.[2] One coder systematically reviewed all articles, coding each study characteristic of interest. We then selected 27 articles—about 10 percent of the total—through systematic random sampling; a second coder reviewed and recoded these. Both coders are coauthors of this paper. To judge coding reliability, we used Cohen's kappa to examine the consistency between both coders' work for these 27 articles. Findings of this analysis, reported in the appendix, indicate that coefficients for each study design feature examined were 0.8 or greater; attaining a kappa value of 0.6 or higher is considered a substantial level of mutual agreement (Landis and Koch 1977). Having examined kappa with only a small sample of articles, we calculated confidence intervals to predict the upper and lower levels of kappa coefficients for the total population of articles. Based on this analysis, we are 95 percent confident that the kappa coefficient from the full sample of articles lies between 0.38 and 1.00 for all coded contents. Accordingly, we concluded that our coding of these articles was substantially reliable.

Findings

Our coding scheme differentiated between primary and secondary research, though there were some unidentified articles for which the distinction was difficult to judge. Primary survey research includes both data collection and analysis, while secondary survey research focuses on the application of analytical techniques to survey data collected by others (Kiecolt and Nathan 1985). We identified 153 articles (58 percent of the total) reporting primary survey research in the five journals during this period, while 93 articles (35 percent) reported secondary survey analysis. Eighteen articles did not specify whether the data were primary or secondary survey data. Also, 90 percent of the survey data sets were cross-sectional; only 9 percent were longitudinal. In all, 95 percent of the primary studies reviewed were collected using a cross-sectional survey design; 5 percent collected longitudinal data at multiple time points.

Coverage Error

A target population is the group to which researchers are interested in generalizing their findings. Table 1 indicates that 57 percent of the studies examined focused on the public sector, including federal, state, and local government employees in each survey category. Excluding the "other" category, local government employees most commonly served as study subjects (35 percent), followed by those in state governments (13 percent). Employees of nonprofit organizations (11 percent) and citizens (12 percent) also were common target populations. This suggests that issues of local government, the behaviors of public managers in local government, and local governance are central themes in empirical studies in our field. Research on public perceptions of administration (or policy), civil participation, and public attitudes mainly used secondary data: primarily public polls and national survey data.

Coverage errors are a form of nonsampling error occurring when the target population and the sample frame(s) are mismatched in coverage. In other words, if the sample frame list does not cover the complete population to be studied, it suffers from undercoverage bias (Groves et al. 2004). An example of this problem would be if a target of public managers in a certain county were being sampled through a sample frame of public officials in several large cities within the county, while several small townships were not included. In practice, few studies specify how researchers construct their sample frames and how well they overlap with the intended target population. As shown in table 2, only about 27 percent (primary,

Table 1 Number (Percent) of Journal Articles, by Target Population and Survey Type, 2000–2007

Target Population	Primary	Secondary	Did Not Specify	Total
Federal employees	12 (8)	11 (12)	1 (5)	24 (9)
State employees	12 (8)	18 (20)	3 (17)	33 (13)
Local employees	62 (40)	24 (26)	7 (39)	93 (35)
Nonprofit employees	23 (15)	6 (6)	–	39 (11)
Citizens	13 (9)	19 (20)	–	32 (12)
Other	31 (20)	15 (16)	7 (39)	53 (20)
Total	153 (58)	93 (35)	18 (7)	264 (100)

Notes: The unit of the target population is employees; percentage in parentheses; "other" refers to any combinations of sectors, private employees, public (or private) schools or specialists (e.g., police, firefighter).

Table 2 Number (Percent) of Journal Articles Reporting How Sample Frame Was Constructed, by Survey Type, 2000–2007

Construction Of Sample Frame	Primary	Secondary	Did Not Specify	Total
Yes	45 (30)	24 (26)	2 (11)	71 (27)
No	2 (1)	3 (3)	–	5 (2)
Did not specify	106 (69)	66 (71)	16 (89)	188 (71)
Total	153 (100)	93 (100)	18 (100)	264 (100)

Table 3 Number (Percent) of Journal Articles Employing Survey Data, by General Sampling Method and Survey Type, 2000–2007

Sampling Type	Primary	Secondary	Did Not Specify	Total
Probability sampling	41 (26)	31 (33)	2 (11)	74 (28)
Nonprobability sampling	13 (9)	2 (2)	–	15 (6)
Other*	4 (3)	–	1 (6)	5 (2)
Did not specify	95 (62)	60 (65)	15 (83)	170 (65)
Total	153 (100)	93 (100)	18 (100)	264 (100)

* Includes surveys based on a census of the population of interest.

Table 4 Number (Percent) of Journal Articles Employing Probability Sampling Methods, by Type of Probability Sample Design and Survey Type, 2000–2007

Probability Sampling Type	Primary	Secondary	Did Not Specify	Total
Simple random sampling	13 (33)	6 (19)	–	19 (26)
Stratified sampling	10 (25)	7 (23)	–	17 (23)
Cluster sampling	1 (2)	–	–	1 (1)
Other complex design	3 (7)	7 (23)	–	10 (14)
Did not specify	14 (33)	11 (35)	2 (100)	27 (36)
Total	41 (100)	31 (100)	2 (100)	74 (100)

30 percent; secondary, 26 percent) of the articles reported how the sample frame was constructed for sample selection. Undoubtedly, many more in fact did construct excellent sample frames. Failure to report this information, of course, makes it impossible for readers to reach this or any other conclusion.

Sampling Error

There are two types of sampling techniques used in social surveys: probability and nonprobability sampling. Probability sampling techniques involve the selection of samples from a defined target population using a random mechanism such that every sample unit in the target population has a known probability of selection. In contrast, nonprobability sampling, also referred to as convenience sampling, does not rely on random selection. Instead, samples are collected based on nonrandom mechanisms that render it impossible to know the probability of selection for each sample unit (Folz 1996; Fowler 2002; Henry 1990).

Table 3 presents the general sampling techniques reported in the articles reviewed. About 28 percent of all articles explicitly reported that surveys were undertaken with probability sampling methods. Approximately 6 percent indicated that data were collected using nonprobability methods, and about 65 percent did not specify how sampled persons were selected, making it impossible for readers to judge the quality of the sampling plan.

Of the probability sample designs reported, simple random sampling was most common: 26 percent overall, 33 percent of primary survey studies, and 19 percent of secondary survey studies (see table 4). In total, 23 percent of the probability samples employed stratification, 14 percent used multistage sampling, and 1 percent used cluster sampling. About 36 percent of these surveys used probability sampling without specifying the specific type of sample design.

Table 5 Number (Percent) of Journal Articles Reporting Sample Size, by Survey Type, 2000–2007

Reporting Sample Size	Primary	Secondary	Did Not Specify	Total
Yes	143 (93)	70 (75)	14 (78)	227 (86)
No	10 (7)	23 (25)	4 (22)	37 (14)
Total	153 (100)	93 (100)	18 (100)	264 (100)

Table 6 Percentage of Journal Articles, by Initial and Completed Sample Size and Survey Type, 2000–2007

	Initial Sample Size (percent)		Completed Sample Size (Percent)	
Sample Size	Primary	Secondary	Primary	Secondary
Less than 250	29	–	42	3
250–499	25	3	29	25
500–999	17	31	11	28
1,000–1,999	18	28	11	19
More than 2,000	11	38	8	25
Mean	1,430	4,195	822	2,619
Standard deviation	2,967	9,080	1,580	5,464

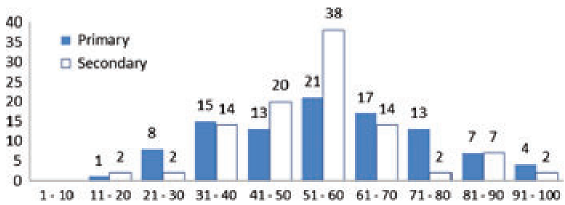
Under random sampling, sample size is negatively related to sampling errors: the smaller the sample size, the greater the potential sampling errors.[3] Smaller sample sizes also lead to larger confidence intervals, which, in turn, produce less accurate predictions. A large sample size has more precision and usually has better statistical properties in analytical models. However, it is important to note that large sample sizes do not reduce nonsampling errors, which in many cases are more influential in determining the quality of survey data. In other words, a large sample size does not guarantee quality survey data unless nonsampling errors also are controlled.

Reporting sample size is a fundamental requirement when presenting survey research findings (Johnson and Owens 2003). Table 5 indicates that of the 264 journal articles reviewed, 86 percent (N = 227 studies) reported survey sample sizes.[4] Reporting this information was more common for studies using primary data (93 percent) than for those using secondary data (75 percent). Few researchers, of course, are successful in collecting data from all units sampled for a survey. Commonly referred to as unit nonresponse, this occurs for many reasons, most typically because some portion of sampled respondents is unwilling or unable to participate.

Table 6 presents the distribution of initial sample size by survey type. About 30 percent of primary survey studies initially selected samples of more than 1,000, whereas nearly 66 percent of secondary data did so. On the other hand, 54 percent of primary surveys reported an initial sample size of less than 500, compared to 3 percent of the secondary studies reviewed. The mean initial sample size for primary surveys was 1,430 (standard deviation, 2,967); that of secondary data was 4,195 (standard deviation, 9,080). The table illustrates how distributions of completed sample size differ between the two survey types. Studies with more than 500 respondents were more common among those examining secondary survey data, compared to primary survey studies: 72 percent versus 30 percent, respectively. About 71 percent of the articles that used primary survey data analyzed samples of fewer than 500 cases.

Table 7 Number (Percent) of Journal Articles Reporting Response Rate, by Survey Type, 2000–2007

Reporting of Response Rate	Primary	Secondary	Did Not Specify	Total
Yes	124 (81)	55 (58)	11 (61)	190 (72)
No	29 (19)	38 (42)	7 (39)	74 (28)
Total	153 (100)	93 (100)	18 (100)	264 (100)



Primary: mean = 56.7, standard deviation = 19.2, median = 56.7.
Secondary: mean = 53.3, standard deviation = 15.0, median = 53.0.

Figure 1 Percentage of Journal Articles with Primary and Secondary Survey Date, By Response Rates: 2000–2007

Table 8 Number of Journal Articles Indicating Follow-up Contact for Nonresponse, by Survey Type, 2000–2007

Follow-up Contact	Primary	Secondary	Did Not Specify	Total
Yes	30 (19)	10 (11)	3 (17)	43 (16)
Did not indicate	123 (81)	83 (89)	15 (83)	221 (84)
Total	153 (100)	93 (100)	18 (100)	264 (100)

Nonresponse Error

Nonresponse error is adversely associated with response rate.[5] Table 7 presents information on response rate reporting. In total, 72 percent of the studies reported response rates: 81 percent for primary and 58 percent for secondary survey studies. As shown in figure 1, the distribution of response rates is approximately normal for both primary and secondary surveys. The values of the mean and median response rates are nearly identical: 56.7 versus 56.7 for primary and 53.3 versus 53.0 for secondary. About half of all primary surveys lie between 40 percent and 60 percent in response rates, whereas about 70 percent of secondary surveys fall into the same range. The mean response rate for primary survey research was 56.7 percent (standard deviation, 19.2), which was remarkably close to the mean (53.3 percent; standard deviation, 15.0) for secondary research.

Enhancing response rates is considered one of the best ways to minimize nonresponse error in survey estimates (Fowler 2002). As shown in table 8, about 16 percent (primary, 19 percent; secondary, 11 percent) of the articles reported using a follow-up contact to increase response rates, whereas a majority of the studies did not specify whether they employed follow-up procedures. We have little doubt that many of these surveys in fact did employ such procedures. This information, however, was not documented in these papers. As response rate is considered an important indicator for assessing nonresponse error (Fowler 2002), it should be calculated in a systematic manner that facilitates comparisons across studies. The American Association for Public Opinion Research (AAPOR 2010) provides six formulas for calculating response rates for surveys with probability sampling but does not offer any methods for surveys with nonprobability sampling.[6] The AAPOR standards are used to calculate response rates for telephone, in-person, mail, and Web surveys. It is important to note that none of the studies reviewed specifically reported employing an AAPOR standard formula to compute response rates.

Table 9 Number (Percent) of Journal Articles Employing Survey Data, by Data Collection Mode and Survey Type, 2000–2007

Data Collection Mode		Primary	Secondary	Did Not Specify	Total
Self-administered	Mail	104 (69)	35 (37)	7 (39)	146 (55)
	Web	7 (5)	3 (3)	–	10 (4)
Interviewer-administered	In-person	6 (4)	8 (9)	1 (5)	15 (6)
	Telephone	10 (6)	8 (9)	3 (17)	21 (8)
Mixed mode		12 (7)	2 (2)	2 (11)	16 (6)
Did not specify		14 (9)	37 (40)	5 (28)	56 (22)
Total		153 (100)	93 (100)	18 (100)	264 (100)

Measurement Error

Although many factors may affect measurement errors in surveys, appropriate mode selection and the quality of the questionnaire are recognized as integral components for reducing errors. Table 9 examines the survey modes employed in public administration papers reporting survey research, including mail, Web, in-person, telephone, and mixed-mode methodologies. Though computer-assisted instruments have been utilized widely in large-scale surveys such as the General Social Survey and the American National Election Survey, in practice, we found that computer-assisted technologies rarely were reported in primary survey research in public administration.[7]

Surveys are either self- or interviewer-administered. Self-administered modes include mail and Web surveys, while interviewer-administered surveys generally are completed by phone or in person. Of 153 articles reporting primary surveys, about 70 percent involved mail surveys, followed by mixed-mode surveys (7 percent), telephone interviews (6 percent), Web surveys (5 percent), and in-person interviews (4 percent). In secondary survey research studies, about 37 percent used mail surveys; followed by in-person interviews (9 percent), telephone interviews (9 percent), and the Web (3 percent). A somewhat surprising 40 percent of secondary studies did not specify the method of survey administration, which is much higher than primary surveys (9 percent). Overall, the most used survey mode is mail (55 percent of all studies reported). In terms of nonsampling error, one potential problem for mail surveys is nonresponse bias (Bridge 1974; Ellis, Endo, and Armer 1970; Filion 1975; Fowler 2002; Kanuk and Berenson 1975; Wright, Manigault, and Black 2004). In contrast, measurement error stemming from social desirability bias is believed to be reduced in self-administered modes such as mail surveys (Fowler 2002). Social desirability bias arises when survey questions are viewed by respondents as being sensitive. Sensitive topics tend to be more intrusive and include the threat of disclosure such as criminal activities, sexual behavior, and voting behavior (Tourangeau and Yan 2007).

Given that the predominant topic of public administration research employing survey research methods is organizational or administrative behavior, in which respondents' motivations, attitudes, and perceptions of work and organizations are examined

Table 10 Number (Percent) of Journal Articles Employing Questionnaire Evaluation Methods, by Survey Type, 2000–2007

Questionnaire Evaluation Method	Primary	Secondary	Did Not Specify	Total
Focus group	3 (2)	–	–	3 (1)
Expert review	2 (1)	–	–	2 (1)
Pretesting	10 (6)	1 (1)	2 (11)	13 (5)
Other	4 (2)	–	–	4 (2)
Did not specify	134 (89)	92 (99)	16 (89)	242 (92)
Total	153 (100)	93 (100)	18 (100)	264 (100)

Note: "Other" includes cognitive interview and any combination of two or more methods.

(rather than more personal and/or private matters), social desirability bias would seem to be less of a concern in public administration research compared to other behavioral sciences. Accordingly, loss of data quality due to nonresponse bias may outweigh the gain due to reducing socially desirable bias via mail surveys in public administration research.

We also investigated documentation of questionnaire evaluation, as shown in table 10. Although most of the studies did not specify if any strategies were employed to develop and/or refine the survey instrument (92 percent), pretesting, focus groups, and expert reviews were each reported by a small number of studies. These activities were more likely to be reported in primary survey research studies.

Processing Error

Processing errors can emerge when data management operations, including editing, coding, and weight construction, are activated after data collection (Groves et al. 2004). Here, we extend this error category to the appropriateness of statistical techniques because survey research methodology includes both data collection *and* analysis. Table 11 illustrates the main statistical techniques applied in the studies reviewed.[8] Nearly 44 percent were designed to test formal hypotheses to verify theories or previous findings, and about 56 percent did not use hypothesis testing, being more descriptive in nature. Survey research within this context thus appears to be more focused on exploring administrative phenomenon than on theory verification.

The statistical techniques used in 35 percent of the primary survey studies were basic univariate or bivariate methods such as descriptive statistics (i.e., means, standard deviations, z-scores), *t*-tests, χ^2-tests, and Pearson correlations, while only 20 percent of the secondary survey research studies used these methods. Primary survey studies were more likely to rely on simple statistical techniques than studies using secondary data. Linear and nonlinear regression methods such as ordinary least squares (OLS) and multinomial regression were employed in 59 percent of the primary survey studies and 69 percent of the secondary survey studies.

Linear regression with the OLS method (35 percent) was the dominant statistical approach used. Although widely used, it has many assumptions—e.g., individual independence, multicollinearity, heteroscedasticity—to be satisfied. Approximately 17 percent of the studies using linear regression reported whether models met the assumptions. Of the 85 studies, 10 percent reported transforming variables to satisfy model assumptions; 36 percent reported using robust standard errors. About 47 percent did not specifically discuss whether assessments of regression assumptions were investigated. Furthermore, we observed that some studies utilized linear regression without carefully considering the scale of the dependent variables (i.e., ordinal scale of dependent variable). Analytic techniques must match variable measurement structures.

In addition, we investigated statistical appropriateness in complex sample designs. In typical statistical models, standard errors and *p*-values of estimates are assumed to be drawn with simple random sampling. If a complex sample design is used, other procedures such as sample weighting and/or standard error adjustments are needed to develop correct estimates and their variances (Campbell and Berbaum 2010; Lee, Forthofer, and Lorimor 1989; Lohr 1999). We investigated 28 articles that reported using a probability sampling method other than simple random sampling. Only about 21 percent also indicated that their analyses took into account the complex nature of their survey sample. In summary, about 65 percent of all survey studies reviewed reported the use of simple statistical methods such as linear regression analysis, descriptive statistics, and simple bivariate tests. This finding appears consistent with previous reviews claiming that public administration empirical research relies on less advanced statistical techniques relative to other social sciences (see Cozzetto 1994; DeLorenzo 2000; Gill and Meier 2000; Perry and Kraemer 1986). More importantly, we found that linear regression was a common statistical technique, although a majority of the studies employing it did so without reporting their treatment of underlying assumptions.

Table 11 Number (Percent) of Journal Articles Employing Survey Data Analysis Techniques, by Survey Type and Hypothesis Testing, 2000–2007

Statistical Technique		Primary			Secondary			Total
		Hypothesis Testing			Hypothesis Testing			
		Yes	No	Total	Yes	No	Total	
Descriptive statistics		4 (6)	28 (33)	32 (21)	–	7 (13)	7 (7)	39 (16)
t-test/chi square/ANOVA		6 (9)	11 (13)	17 (11)	4 (10)	6 (12)	10 (11)	27 (11)
Pearson correlation		2 (3)	2 (2)	4 (3)	1 (2)	1 (2)	2 (2)	6 (2)
Linear regression	OLS	25 (38)	28 (33)	53 (35)	18 (44)	14 (27)	32 (34)	85 (35)
	GLS/WLS/ 2SLS	2 (3)	2 (2)	4 (3)	2 (5)	1 (2)	3 (3)	7 (3)
	SEM/Path	9 (13)	1 (1)	10 (7)	2 (5)	3 (6)	5 (5)	15 (6)
	HLM	1 (2)	1 (1)	2 (1)	–	–	–	2 (1)
Logistic/multinomial		14 (21)	6 (7)	20 (13)	10 (24)	15 (29)	25 (27)	45 (18)
Factor analysis		1 (2)	3 (4)	4 (3)	–	2 (4)	2 (2)	6 (2)
Other		3 (4)	3 (4)	6 (4)	4 (10)	3 (6)	7 (4)	13 (5)
Total		67 (44)	86 (56)	153 (100)	41 (44)	52 (56)	93 (100)	246 (100)

Note: Unspecified cases (n = 18) are excluded.

Discussion and Suggestions

Prior research evaluations have focused largely on measurement issues and analytical techniques and often assume that data quality is perfect. In contrast, this study examined survey data collection methodologies and evaluated them within the total survey error framework in order to assess the quality of survey data being reported in public administration journals. Our findings show that survey research in the public administration literature generally features small-scale surveys, heavy reliance on a single specific mode (mostly mail surveys), unspecified sample frame quality, unspecified sample selection procedures, and wide use of linear regression analysis models but without specifying the extent to which model assumptions were evaluated.

Our findings show that survey research in the public administration literature generally features small-scale surveys, heavy reliance on a single specific mode (mostly mail surveys), unspecified sample frame quality, unspecified sample selection procedures, and wide use of linear regression analysis models but without specifying the extent to which model assumptions were evaluated.

Quality of Survey Data

Survey data quality can be maximized when TSE is minimized. Although TSE is not perfectly observed here because so many articles did not report information about the data collection procedures used, we nonetheless are able to make several conclusions. Our findings show that survey research in the public administration literature relies heavily on self-administered instruments. One important concern with this particular mode is a potentially high degree of nonresponse bias because of low response rates (Bridge 1974; Ellis, Endo, and Armer 1970; Filion 1975; Kanuk and Berenson 1975; Wright, Manigault, and Black 2004). There are also several important advantages of self-administered surveys: they are less costly than other modes, they are not affected by interviewer-related bias, and they are less susceptible to socially desirable responses (Bridge 1974; Fowler 2002; Kanuk and Berenson 1975). The main reason for the preference for mail surveys in public administration survey research is likely funding constraints; mail surveys are less costly to carry out.

Web surveys are also comparatively inexpensive. Although they represented only 5 percent of the surveys reported in the public administration literature between 2000 and 2007, we have no doubt that reliance on this mode will increase in subsequent years. Web surveys present new opportunities and challenges that will merit future attention. Currently, there are a few sources that address specific aspects of Web survey research methodology (Couper 2008; Sue and Ritter 2007).

Indeed, lack of financial resources may be an impediment to the application of more sophisticated survey data collection methods among public administration researchers (Gill and Meier 2000; Perry and Kraemer 1986). In this regard, it is worth noting that there is a trade-off between costs and survey data quality (Groves 1989).

We make no claim that the mail survey is a problematic data collection strategy. Rather, we regard it as an efficient and practical tool that can be undertaken within the constraints of research funding, as long as nonresponse bias and measurement errors are adequately controlled. The use of pre- and postnotification (Assael and Keon 1982; Kanuk and Berenson 1975), monetary incentives (Armstrong 1975; Brown and Coney 1977; Goodstadt et al. 1977; Kanuk and Berenson 1975), mixed modes, or special questionnaire design protocols such as the tailored design method (Dillman, Smyth, and Christian 2009) are some recommended approaches to improve data quality in mail surveys.[9]

Further, efforts to reduce measurement errors stemming from complex questionnaires or poor question wording in survey items must be considered. To reduce these sources of error, focus groups, expert reviews, pretesting, and other questionnaire design strategies are necessary for the refinement of questionnaires (e.g., De Leeuw, Borgers, and Smits 2004), yet we observed only a few studies that reported such activities.

The common goal of sample surveys is to accurately estimate a population of interest with a small sample. To do so, the sample must be representative of the study population. Sample representativeness, known as external validity, is a cornerstone of the sample survey. In this respect, the quality of a sample frame and the sampling methodology exert a considerable influence on a survey's representativeness. As shown in our findings, nearly 70 percent of all studies reviewed did not detail whether a frame was used when samples were drawn. An imperfect or nonexistent sampling frame can give rise to coverage problems, throwing into doubt the representativeness of the sample.

Another important element for assessing representativeness is whether probability sampling methods are employed. Approximately 65 percent of the surveys reviewed did not specify how respondents were selected. As with all other social science methodologies, those employing surveys are obligated to inform readers about the degree to which their data may be biased. Papers in psychology, for instance, generally are expected to report whether subjects are assigned at random to experimental and control groups when an experimental design is utilized in a study (APA 2008). This reflects the recognition that random assignment is an integral element of experimental research as a criterion of research quality. Similarly, whether probability sampling is used or not is a critical issue in judging the quality of survey data. Of course, there are multiple types of probability sampling commonly employed in survey research. Survey results with nonprobability sampling, however, are difficult to generalize because it is unknown to what extent the sample represents the population of interest. Moreover, there is no way of estimating the degree of precision of estimates attained from convenience sampling (Lavrakas 1987). At present, AAPOR recommends against the reporting of margin of errors for survey data from nonprobability samples. In addition, a fundamental assumption of many statistical models such as linear regression is that samples are drawn at random (Wooldridge 2002). This serves as an important reminder that random selection from a well-defined sample frame

As with all other social science methodologies, those employing surveys are obligated to inform readers about the degree to which their data may be biased.

is an integral assumption of survey data analysis. However, as we noted previously, few studies specified how their sample frames were constructed. Frame quality determines the degree of coverage error in survey research. If the frame list does not fully represent a target population, generalizability issues again emerge.

Survey Data Analytical Techniques

In addition to the accurate collection of survey data, properly analyzing these data is critical to making valid inferences. Our findings indicate that the linear regression model estimated with OLS is currently the predominant technique in public administration for modeling survey data. In reviewing the literature, we found that many researchers either did not assess the underlying assumptions of regression when using this technique or failed to report doing so. There are some critical assumptions that must hold for linear regression to be used, including independent observations, a lack of multicollinearity, and homoscedasicity of residuals, that must be satisfied to obtain valid results using this technique. The assumption that every observation is independent, however, may be easily violated when observations are nested within upper contexts such as agencies, schools, neighborhoods, and work units because observations within the same context are known to be more homogeneous (or heterogeneous between contexts) in respondent behaviors and attitudes. In this particular case, rather than the generic linear regression model, another method such as hierarchical linear modeling (HLM) would be more appropriate.

Further, the weighted least squares (WLS) model generally should be used when an analytic model suffers from heteroscedasticity, in which error variance is not constant, or the heteroscedasticity-robust standard error (or robust standard error) is required to avoid increased type I error. The two-stage least squares (2SLS) model is recommended when there is an endogeneity problem—violation of recursivity assumption—in the generic linear regression model. The model needs to be transformed to the generalized least square (GLS) when both (either) heteroscedasticity and (or) autocorrelation occur(s). When error variance is not constant, linear regression's OLS estimators are unbiased but no longer efficient. The GLS makes estimators more efficient (Gujarati 2003). Few studies detailed a rationale for choosing a statistical model or reported a consideration of their underlying assumptions before analyzing their data.

Reporting Issues

We also would like to acknowledge the importance of careful reporting of survey data collection procedures as well as analytical findings. We constructed a reporting index to observe the detail with which authors reported nine specific aspects of their survey data collection procedures.[10] The average reporting score was 4.2 (primary, 4.6; secondary, 3.7). This indicates that authors in general reported less than half of these criteria; papers reporting secondary survey data analyses were more likely to omit this basic information. While it certainly would not be possible for researchers to provide all of the details of their survey data collection methodologies, we nonetheless argue that the basic information discussed in this paper should be described when reporting scientific survey results. That such key elements of survey methodology generally are not fully documented in public administration mainstream journals suggests that authors and editors need to reconsider the importance of full disclosure of the scientific methods used in professional papers.

While it certainly would not be possible for researchers to provide all of the details of their survey data collection methodologies, we nonetheless argue that the basic information discussed in this paper should be described when reporting scientific survey results.

One also might think that there is no need to specify the description of data collection procedures in each research paper, particularly when multiple papers are published using the same data. However, we believe that each paper should be independent in describing its methodology.[11] Merely providing references to the availability of methodological details that can be found elsewhere is usually inadequate, as researchers, for example, should seriously consider how various sources of TSE may directly influence the quality of the measures being examined in each unique study. For example, nonresponse error may be expected to have differential implications for the quality of various survey measures that are dependent on the degree to which the underlying nonresponse mechanisms believed to be associated with a given survey might also be correlated with specific independent or dependent variables of interest. These effects can vary considerably across the measures included in any given survey, requiring renewed consideration each time the data file is reexamined to address new hypotheses. Wright, Manigault, and Black (2004) considered the issue of the scarcity of information provided regarding methodological procedures and measures in the published literature. They suggested two potential reasons for this scarcity: (1) many professional journals lack explicit guidelines or policies regarding the full disclosure of data collection procedures (see also Johnson and Owens 2004); (2) authors omit this information because reviewers of journals demand condensed papers because of space limitations (see also Luton 2007; Lynn, Heinrich, and Hill 2001).

Conclusion

The total survey error framework is a useful tool for categorizing the types of survey error possible when conducting research and for understanding the impacts of those errors on data quality. Although survey research is being used widely in the public administration field, there have been no prior reviews of the use or reporting of indicators of survey research quality. We have applied the TSE framework to the public administration literature using survey research methodology and found significant gaps in reporting on possible sources of error. Journal editors and reviewers should consider not only how to best *analyze* but also how to best *collect* survey data while minimizing a variety of error sources. This will help ensure that readers have sufficient information with which to judge the quality of the survey data underlying analytical results. Thus, to the extent that they do not currently exist, we encourage public administration journals to consider establishing standards for methodological transparency in the disclosure of survey data collection procedures.

Acknowledgments

The earlier version of the manuscript was presented at the 2009 Conference of National Public Management Research in Columbus, OH. We are grateful to Young Ik Cho, Allyson Holbrook, Mary Feeney, and three anonymous reviewers for their useful suggestions.

Notes

1. The average impact factors of these journals in the past five years are between 0.99 and 2.49 on a standard of ISI journal citation reports: *Administration & Society* (0.99), *American Review of Public Administration* (1.01), *Journal of Public Administration Research and Theory* (2.49), *Public Administration* (1.65), and *Public Administration Review* (1.92).
2. We considered only quantitative survey methods, excluding in-depth interviews and focus groups as qualitative research. In coding, when authors published several articles using the same primary survey data, we counted them multiple times as primary research. Studies using multiple data sources—a combination of survey data and other types of data such as administrative records—were classified as survey research.
3. The statistical formula for sampling error in simple random sampling is as follows: $V(\bar{x}) = (1 - f)\frac{s^2}{n}$, where $(1 - f)$ = the finite population correction (fpc); n = sample size; and $s^2 = (x_i - \bar{x})/(n - 1)$. Sampling error of the sample mean is the square root of $V(\bar{x})$. This formula clearly shows that sampling error is adversely associated with sample size. It is important to note that there is no theoretical basis for using this formula in surveys that do not have probability sample designs.
4. We coded initial and completed sample sizes. The former refers to a sample initially selected; the latter is one in which unit nonresponse cases are excluded. Some articles indicated both sample sizes; others reported either initial or completed size only. For a case in which either initial or complete sample size was reported, we calculated the sample size that was not provided using reported response rates and reported sample size (initial or completed sample) and coded the study as reporting sample size. When multiple data sources were used, we coded the sample size with the arithmetic average of each for both initial and completed sample sizes. For coding of response rate, we used arithmetic average of response rates when multiple surveys were used in the same study.
5. $\bar{y}_r - \bar{y}_s = m_s/n_s(\bar{y}_r - \bar{y}_m)$, where $\bar{y}_r$ = mean of the respondents; $\bar{y}_s$ = mean of the initial sample members; m_s = total number of nonrespondents; n_s = total number of sample members; and $\bar{y}_m$ = mean of the nonrespondents (Groves et al. 2004). Nonresponse error is determined by the two terms (m_s/n_s) and $(\bar{y}_r - \bar{y}_m)$: the first term is nonresponse rate and second term is difference in mean between respondents and nonrespondents in the sample. The second term, however, is unknown because there is no way of knowing the mean of nonrespondents. For this reason, reducing the nonresponse rate (or increasing response rate) is considered the best strategy for reducing nonresponse error.
6. AAPOR provides six formulas for calculating response rates: RR1 through RR6. To apply these formulas, a survey must use probability sampling methods and have a clearly defined sample frame.

$$\text{RR1} = \frac{I}{(I+P)+(R+NC+O)+(UH+UO)}$$

$$\text{RR2} = \frac{(I+P)}{(I+P)+(R+NC+O)+(UH+UO)}$$

$$\text{RR3} = \frac{I}{(I+P)+(R+NC+O)+e(UH+UO)}$$

$$\text{RR4} = \frac{(I+P)}{(I+P)+(R+NC+O)+e(UH+UO)}$$

$$\text{RR5} = \frac{I}{(I+P)+(R+NC+O)}$$

$$\text{RR6} = \frac{(I+P)}{(I+P)+(R+NC+O)}$$

where I = complete interview; P = partial interview; NC = noncontact; O = other; UH = unknown eligibility if household; UO = unknown eligibility (other); and e = estimated proportion of cases of unknown eligibility that are eligible (AAPOR 2010). Most articles that focused on mail surveys indicated that response rates were calculated as a ratio of the number of returned questionnaires to the total number of mailed questionnaires. Strictly speaking, this is a complete rate or mail return rate, not a response rate. For most of the mail surveys we investigated, researchers did not consider partial interview (P), noncontact (NC), and unknown eligibility (UO), making use of only complete interview (I) based on returned mails to compute a response rate.
7. Computer-assisted interviewing is a data collection mode in which interviewers use computers rather than paper-and-pencil questionnaires. Large-scale national surveys commonly employ this data collection mode to enhance response rates.
8. For studies using multiple statistical analyses, we chose only the one technique most relevant to the research questions or hypotheses.
9. The tailored design method is a questionnaire design technique in which words, pictures, and coloring are mobilized to produce user-friendly self-administered questionnaires. It is designed to improve the quality of survey data in self-administered (e.g., mail, Web) surveys. Evidence suggests that this method can improve respondent understanding of survey questions as well as survey cooperation (Dillman, Smyth, and Christian 2009).
10. In constructing the index, we coded a value of 1 for reporting and 0 for nonreporting of each of the following elements: survey design type, sampling frame construction, sampling method, sample size, response rate, mode, questionnaire development, addressing nonresponse, and target population. We constructed the index by summing these values. The hypothetical range of the index was 0–9; the actual range was 0–7.
11. To assess the consistency with which basic survey procedures were reported by authors, we examined multiple papers by the same authors that employed the same survey data sets within the literature reviewed for this paper. Five key indicators were examined (survey mode, sample size, response rate, sample frame, and sample design). We found that 28 authors published at least two papers (mean = 2.36 papers; n = 66 total papers) using the same data. The concordance with which they did or did not report each of these indicators was as follows: mode (0.75), sample size (0.82), response rate (0.79), sample frame (0.68), and sample design (0.71). Although there was reasonable concordance in the reporting by these authors, there nonetheless was considerable variability in the likelihood that each of these indicators were reported: mode (reported in 81.6 percent of these papers), sample size (87.5 percent), response rate (80.4 percent), sample frame (27.4 percent), and sample design (32.1 percent).

Appendix. Interrater Reliability Tests for Data Abstracted from Sampled Articles (n = 27)

Contents	Cohen's kappa (SE)	Confidence Interval
Survey type	0.85 (0.10)	0.65–1.00
Survey design	0.87 (0.13)	0.61–1.00
Questionnaire development	0.79 (0.22)	0.38–1.00
Data collection mode	0.82 (0.10)	0.55–1.00
Sampling method	0.84 (0.09)	0.66–1.00
Sample size	1.00 (0)	1.00
Response rate	0.91 (0.09)	0.73–1.00
Follow-up contact	1.00 (0)	1.00
Sample frame	0.85 (0.10)	0.65–1.00
Target population	0.81 (0.09)	0.63–0.98
Use of hypothesis testing	0.89 (0.11)	0.67–1.00
Survey data analysis technique	0.95 (0.05)	0.84–1.00

Note: All coefficients of Cohen's kappa are significant at the level of .001; SE = standard error.

References

Adams, Guy B., and Jay D. White. 1995. The Public Administration Doctorate: A Degree in Search of a Purpose. *Journal of Public Administration Education* 1(1): 67–76.

American Association for Public Opinion Research (AAPOR). 2010. Standard Definitions: Final Disposition of Case Codes and Outcome Rates for Surveys. http://www.aapor.org/Standard_Definitions2.htm [accessed October 7, 2011].

American Psychological Association (APA). 2008. Reporting Standards for Research in Psychology: Why Do We Need Them? What Might They Be? *American Psychologist* 63(9): 839–51.

Armstrong, J. Scott. 1975. Monetary Incentives in Mail Surveys. *Public Opinion Quarterly* 39(1): 111–16.

Assael, Henry, and John Keon. 1982. Nonsampling vs. Sampling Errors in Survey Research. *Journal of Marketing Research* 46(2): 114–23.

Biemer, Paul P., and Lars E. Lyberg. 2003. *Introduction to Survey Quality.* Hoboken, NJ: Wiley.

Box, Richard C. 1992. An Examination of the Debate over Research in Public Administration. *Public Administration Review* 52(1): 62–70.

Bridge, R. Gary. 1974. *Nonresponse Bias in Mail Surveys: The Case of the Department of Defense Post-Service Survey.* Santa Monica, CA: RAND Corporation.

Brower, Ralph S., Mitchel Y. Abolafia, and Jered B. Carr. 2000. On Improving Qualitative Methods in Public Administration Research. *Administration & Society* 32(4): 363–97.

Brown, Stephen W., and Kenneth A. Coney. 1977. Comments on "Mail Survey Premiums and Response Bias." *Journal of Marketing Research* 14(3): 385–87.

Campbell, Richard T., and Michael L. Berbaum. 2010. Analysis of Data from Complex Surveys. In *Handbook of Survey Research,* edited by Peter V. Marsden and James D. Wright, 221–62. Bingley, UK: Emerald.

Cleary, Robert E. 1992. Revisiting the Doctoral Dissertation in Public Administration: An Examination of the Dissertations of 1990. *Public Administration Review* 52(1): 55–61.

———. 2000. The Public Administration Doctoral Dissertation Reexamined: An Evaluation of the Dissertations of 1998. *Public Administration Review* 60(5): 446–55.

Couper, Mick P. 2008. *Designing Effective Web Surveys.* New York: Cambridge University Press.

Cozzetto, Don A. 1994. Quantitative Research in Public Administration: A Need to Address Some Serious Methodological Problems. *Administration & Society* 26(3): 337–43.

Daneke, Gregory A., and Patricia Klobus-Edwards. 1979. Survey Research for Public Administrators. *Public Administration Review* 39(5): 421–26.

de Leeuw, Edith, Natacha Borgers, and Astrid Smits. 2004. Pretesting Questionnaires for Children and Adolescents. In *Methods for Testing and Evaluating Survey Questionnaires,* edited by Stanley Presser et al., 409–29. Hoboken, NJ: Wiley.

DeLorenzo, Lisa. 2000. Stars Aren't Stupid, but Our Methodological Training Is: A Commentary on Jeff Gill and Ken Meire's Article "Public Administration Research and Practice: A Methodological Manifesto." *Journal of Public Administration Research and Theory* 11(1): 139–45.

Dillman, Don A., Jolene D. Smyth, and Leah Melani Christian. 2009. *Internet, Mail, and Mixed-Mode Surveys: The Tailored Design Method.* 3rd ed. Hoboken, NJ: Wiley.

Ellis, Robert A., Calvin M. Endo, and J. Michael Armer. 1970. The Use of Potential Nonrespondents for Studying Nonresponse Bias. *Pacific Sociological Review* 13(2): 103–9.

Enticott, Gareth, George A. Boyne, and Richard M. Walker. 2009. The Use of Multiple Informants in Public Administration Research: Data Aggregation Using Organizational Echelons. *Journal of Public Administration Research and Theory* 19(2): 229–53.

Filion, F. L. 1975. Estimating Bias Due to Nonresponse in Mail Surveys. *Public Opinion Quarterly* 39: 482–92.

Fisher, Robert J. 1993. Social Desirability Bias and the Validity of Indirect Questioning. *Journal of Consumer Research* 20(2): 303–15.

Folz, David H. 1996. *Survey Research for Public Administration.* Thousand Oaks, CA: Sage Publications.

Forrester, John P., and Sheilah S. Watson. 1994. An Assessment of Public Administration Journals: The Perspective of Editors and Editorial Board Members. *Public Administration Review* 54(5): 474–85.

Fowler, Floyd J., Jr. 2002. *Survey Research Methods.* 3rd ed. Thousand Oaks, CA: Sage Publications.

Gill, Jeff, and Kenneth J. Meier. 2000. Public Administration Research and Practice: A Methodological Manifesto. *Journal of Public Administration Research and Theory* 10(1): 157–99.

Golden, Brian R. 1992. The Past Is the Past—or Is It? The Use of Retrospective Accounts as Indicators of Past Strategy. *Academy of Management Journal* 35(4): 848–60.

Goodstadt, Michael S., Linda Chung, Reena Kronitz, and Gaynoll Cook. 1977. Mail Survey Response Rates: Their Manipulation and Impact. *Journal of Marketing Research* 14(3): 391–95.

Groves, Robert M. 1987. Research on Survey Data Quality. Special issue, *Public Opinion Quarterly* 51: 156–72.

———. 1989. *Survey Errors and Survey Costs.* New York: Wiley.

Groves, Robert M., et al. 2004. *Survey Methodology.* Hoboken, NJ: Wiley Interscience.

Gujarati, Damodar N. 2003. *Basic Econometrics.* 4th ed. New York: McGraw-Hill.

Henry, Gary T. 1990. *Practical Sampling.* Newbury Park, CA: Sage Publications.

Holbrook, Allyson, Young Ik Cho, and Timothy Johnson. 2006. The Impact of Question and Respondent Characteristics on Comprehension and Mapping Difficulties. *Public Opinion Quarterly* 70(4): 565–95.

Houston, David J., and Sybil M. Delevan. 1990. Public Administration Research: An Assessment of Journal Publications. *Public Administration Review* 50(6): 674–81.

Johnson, Timothy P., and Linda Owens. 2004. Survey Response Rate Reporting in the Professional Literature. In *2003 Proceedings of the Section on Survey Research Methods,* 127–33. Alexandria, VA: American Statistical Association.

Johnson, Timothy P., and Fons J. R. van de Vijver. 2003. Social Desirability in Cross-Cultural Research. In *Cross-Cultural Survey Methods,* edited by Janet A. Harkness, Fons J. R. van de Vijver, and Peter Ph. Mohler, 193–202. Hoboken, NJ: Wiley.

Kanuk, Lelie, and Conrad Berenson. 1975. Mail Surveys and Response Rates: A Literature Review. *Journal of Marketing Research* 12(4): 440–53.

Kiecolt, K. Jill, and Laura E. Nathan. 1985. *Secondary Analysis of Survey Data.* Newbury Park, CA: Sage Publications.

Lan, Zhiyong, and Kathleen K. Anders. 2000. A Paradigmatic View of Contemporary Public Administration Research: An Empirical Test. *Administration & Society* 32(2): 138–65.

Landis, J. Richard, and Gary G. Koch. 1977. The Measurement of Observer Agreement for Categorical Data. *Biometrics* 33(1): 159–74.

Lavrakas, Paul J. 1987. *Telephone Survey Methods: Sampling, Selection, and Supervision.* Newbury Park, CA: Sage Publications.

Lee, Eun Sul, Ronald N. Forthofer, and Ronald J. Lorimor. 1989. *Analyzing Complex Survey Data.* Newbury Park, CA: Sage Publications.

Lohr, Sharon L. 1999. *Sampling: Design and Analysis.* Pacific Grove, CA: Brooks/Cole.

Luton, Larry S. 2007. Deconstructing Public Administration Empiricism. *Administration & Society* 39(4): 527–44.

Lynn, Jr., Laurence E., Carolyn J. Heinrich, and Carolyn J. Hill. 2001. *Improving Governance: A New Logic for Empirical Research.* Washington, DC: Georgetown University Press.

McCurdy, Howard E., and Robert E. Cleary. 1984. Why Can't We Resolve the Research Issue in Public Administration? *Public Administration Review* 44(1): 49–55.

Moorman, Robert H., and Philip M. Podsakoff. 1992. A Meta-Analytic Review and Empirical Test of the Potential Confounding Effects of Social Desirability Response Sets in Organizational Behavior Research. *Journal of Occupational and Organizational Psychology* 65(2): 131–49.

Perry, James L., and Kenneth L. Kraemer. 1986. Research Methodology in the *Public Administration Review,* 1975–1984. *Public Administration Review* 46(3): 215–26.

Pinsonneault, Alain, and Kenneth L. Kraemer. 1993. Survey Methodology in Management Information Systems: An Assessment. *Journal of Management Information Systems* 10(2): 75–105.

Stallings, Robert A. 1986. Doctoral Programs in Public Administration: An Outsider's Perspective. *Public Administration Review* 46(3): 235–40.

Stallings, Robert A., and James M. Ferris. 1988. Public Administration Research: Work in *PAR,* 1940–1984. *Public Administration Review* 48(1): 580–87.

Sue, Valerie M., and Lois A. Ritter. 2007. *Conducting Online Surveys.* Thousand Oaks, CA: Sage Publications.

Swidorski, Carl. 1980. Sample Surveys: Help for the "Out-of-House" Evaluator. *Public Administration Review* 40(1): 67–71.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response.* New York: Cambridge University Press.

Tourangeau, Roger, and Ting Yan. 2007. Sensitive Questions in Surveys. *Psychological Bulletin* 133(5): 859–83.

White, Jay D. 1986. Dissertations and Publications in Public Administration. *Public Administration Review* 46(3): 227–34.

White, Jay D., Guy B. Adams, and John P. Forrester. 1996. Knowledge and Theory Development in Public Administration: The Role of Doctoral Education and Research. *Public Administration Review* 56(5): 441–52.

Wooldridge, Jeffrey M. 2003. *Introductory Econometrics: A Modern Approach.* 2nd ed. Cincinnati, OH: South-Western College.

Wright, Bradley E., Lepora J. Manigault, and Tamika R. Black. 2004. Quantitative Research Measurement in Public Administration: An Assessment of Journal Publications. *Administration & Society* 35(6): 747–64.

Zaller, John, and Stanley Feldman. 1992. A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences. *American Journal of Political Science* 36(3): 579–616.

Zerbe, Wilfred J., and Delroy L. Paulhus. 1987. Social Desirable Responding in Organizational Behavior: A Reconception. *Academy of Management Review* 12(2): 250–64.