October 2007

Method Effects in IS Survey Research: An Assessment and Recommendations

William R. King
University of Pittsburgh, billking@katz.pitt.edu

Charles Z. Liu
University of Pittsburgh

Mark H. Haney
University of Pittsburgh

Jun He
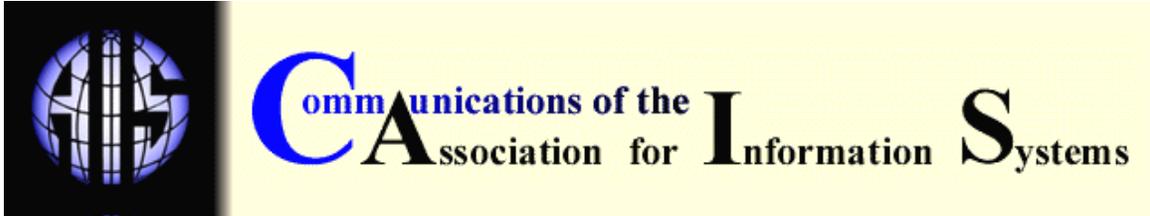*University of Pittsburgh*, JunHe@katz.pitt.edu

Follow this and additional works at: https://aisel.aisnet.org/cais

METHOD EFFECTS IN IS SURVEY RESEARCH: AN ASSESSMENT AND RECOMMENDATIONS

William R. King
Charles Z. Liu
Mark H. Haney
University of Pittsburgh,
Katz Graduate School of Business
billking@katz.pitt.edu

Jun He
University of Michigan Dearborn
School of Management Studies

ABSTRACT

Issues involving method effects are routinely taught to PhD students in Information Systems (IS). Unfortunately, the results of an assessment of a population of 128 survey-based studies published in three top IS journals over a seven-year period (1999-2005) reveal that relatively little attention is being paid to method bias and that the threat of serious method bias is great in many of the published studies. For instance, even the best-understood variety of method bias— common source bias—is found to have gone unnoted in over one-third of the papers that used a single respondent for all construct measures for reasons other than necessity. This study was motivated by studies in other areas of the social sciences which have resulted in calls for areas such as IS to conduct empirical assessments of the frequency of the various forms of method bias. Here, the myriad sources of method bias are reviewed and methods for minimizing or eliminating method effects, both in the design of a study and in the subsequent analysis of the data, are discussed. Data on the frequency of appearance of a wide variety of potential sources of method bias are provided and conclusions are drawn. A series of recommendations is made concerning creating greater awareness of method bias on the part of the IS research community and greater use of method bias evaluation criteria in the screening reviews of papers that are done by IS journals.

I. INTRODUCTION

There is considerable evidence that method bias ("method error," "method variance," or "method effect") constitutes a serious problem for the validity and credibility of social science research [eg., Cote and Buckley 1987; Podsakoff et al. 2003]. Despite this evidence, some have argued that strong method bias may not be universal across fields and have suggested the need for assessments in specific research domains [Crampton and Wagner 1994].

We set out to make such an assessment in the domain of IS survey research. We chose this domain because of our personal interest in it, because it represents more than one-half of all IS research published in the "top three" IS journals [King and He 2005] and because the pervasive use of perceptual measures in this domain makes it particularly susceptible to method effects.

We wished to develop answers to a number of questions:

1. How prevalent is potential method bias in IS survey research?

2. Is the potential for method bias recognized and acknowledged in IS survey research studies?

3. What techniques are used to reduce the impact of method bias?

4. What can be done by IS researchers and journal editors/reviewers to reduce method bias and/or focus greater attention on it?

First, we review the myriad means of inadvertently introducing method effects into a survey study as well as the methods that may be employed to eradicate such effects or to minimize their magnitude through the design of a study. We also discuss statistical methods that can be used to estimate the seriousness of such effects subsequent to data collection.

We assessed all of the survey research studies published in the three top-ranked IS journals over a six-year period by coding each study in terms of a profile of all previously recognized sources of method effects in survey research. We assessed whether the manner in which the study was designed and executed was likely to have introduced such effects, whether the likelihood of these effects was recognized by the researchers and whether they took steps in the design and analysis phases of each study to minimize or ameliorate potential method effects.

Our assessments indicate that there is the potential for significant levels of method bias in IS survey research and that there appears to be too little awareness of it, or interest in it. We propose a concerted campaign, conducted through the review and editing processes of IS journals, to create greater awareness of this issue and to reduce the potential for method bias in papers published in IS journals.

II. METHOD BIAS IN IS SURVEY RESEARCH

Method bias is the systematic variability that can be introduced into the data that are gathered in a study by the method that is used to gather the data ("artifactual covariation"). Method bias is particularly important for data that are perceptual in nature and/or self-reported by respondents— both common attributes of IS survey research studies.

Method bias may be minor, as may be the case when demographic data are self-reported rather than being objectively obtained from an organization's records, or its effects may threaten the validity and credibility of the entire study, as may be the case when common source bias is introduced into the data by having perceptual assessments of both the independent and dependent variables made by the same respondent.

Method bias is not unique to survey research using perceptual measures, but is also important in experimental and other organizational and IS research methodologies.

In any of these types of studies, the variance in the data can be attributed to the sum of the "true" variance in the trait being measured and to measurement error. Measurement error, in turn, can be thought of as having random and systematic components. Possible systematic error is extremely important since it may provide an explanation for the observed relationships between study variables other than the hypothesized explanation.

One of the major sources of systematic measurement error is "method effect," that is, the portion of measurement error that is caused by the method used. "Method" can refer to specific characteristics of the form of measurement, such as item wording, scale type, response format, and general context [Fiske 1982]. It can also refer to response biases such as halo effects, social desirability, acquiescence, leniency effects, or yea- and nay-saying [Bagozzi and Yi 1991].

A major and common variety of method effect—common source bias—can exist when two variables that have a hypothesized relationship are measured using the perceptions of the same individual. This is particularly troublesome when the variables are the primary independent and dependent variables in a study, but it can also be problematic with mediator and moderator variables.

In such instances, common source bias may be introduced due to a respondent's innate desire to be consistent or it may reflect an explicit or implicit theory that the rater holds about how the world works. For instance, a believer in the importance and efficacy of IS strategic planning may subconsciously respond to items related to how planning is done (independent variables) and how effective it is (dependent variables) in a way that reflects his/her strongly held beliefs. Even if respondents do this only subconsciously, it can have a major impact on the results of research studies involving these variables.

III. THE IMPORTANCE OF METHOD EFFECT

Method effect can bias results by making the relationship between constructs seem greater or less than it would have been without the method effect [Bagozzi et al. 1991]. There is a significant body of empirical evidence in the social sciences that this is a serious problem.

Cote and Buckley [1987] used a confirmatory factor analysis procedure to analyze 70 datasets from 64 published studies in a variety of social science disciplines. They found significant method variance in all but three of the studies. On average, they found that the measures used in the studies reflected 41.7 percent trait variance (variance explained by relationships between constructs in the research model), 26.3 percent method variance (variance attributable to the method used), and 32.0 percent random error variance (variance due to measurement errors or other random factors).

Another body of research in social science has examined the extent to which method variance affects the observed relationships between measures. Podsakoff et al [2003] summarized the findings of several such studies, concluding that when method variance was present, the amount of variance explained by the relationship between two variables was 35 percent, whereas it was only 11 percent when method variance was controlled for.

Thus, method variance is not only present in measures of constructs in the social sciences, it also has a significant effect on the observed relationships between those constructs. This threatens the validity of the conclusions reached when observed relationships between constructs are used to test theory.

Podsakoff et al. [2003] also estimated the average amounts of trait variance, method variance, and random error present in typical social science measures, as well as the average method intercorrelations from various studies, and used them to estimate the impact of method variance on the observed correlation between measures of different constructs. The results of their analysis suggest that if two traits are perfectly correlated, and if the typical amount of method variance is present, it will halve the observed correlation between the measures and reduce the variance explained by 70 percent. Conversely, if the true correlation between two traits is zero, and if the typical amount of method variance is present, it will cause the observed correlation between the measures to be significantly greater than zero.

Although the prime focus of attention given to method effects is their impact on the validity of research results, these potential errors also affect the credibility associated with research studies, as has been addressed by Luong and Rogelberg [1998] in a different, but similar, context. They argue that despite the adequacy of the sample and the statistical methodology that is employed in the analysis, studies that involve the possibility of unrecognized or unmeasured error are given low credibility by the research community.

Although this evidence suggests that method variance can be a significant problem in social science research, some researchers argue that the effect of method variance may not be as general or significant as these studies suggest, and that domain-specific research should be conducted to identify those domains of research where such bias may be especially prevalent [Crampton and Wagner 1994]. Even Cote and Buckley [1987], who estimated method variance at an overall average of about 26 percent, found that there was significant variation across different fields.

This study is, in part, reflective of that admonition to assess method bias in specific disciplines; in this case, information systems [IS].

IV. THE SOURCES OF METHOD EFFECTS

Podsakoff et al [2003] classified the sources of method bias into four types: common rater effects, item characteristic effects, item context effects, and measurement context effects. Tables 1-3 list the specific ways in which these effects may be inadvertently introduced into a research study [adapted from Podsakoff et al., 2003 and the other sources cited in the tables].

"Common rater effects" (common source bias) refers to bias produced as a result of having a single respondent provide the measures of multiple variables. Several sources of common rater effects are listed in Table 1.

Table 1. Sources of Common Rater Effects

Source	Definition	Reference
consistency motif	the tendency of respondents to maintain response consistency	[Johns 1994]
explicit or implicit theories	the respondents' explicit or implicit beliefs about the relationships that exist between study variables	[Smither, Buda, & Collins 1989]
social desirability	the tendency to give responses that are socially acceptable rather than responding in a totally honest fashion	[Ganster, Hennessey, & Luthans 1983]
leniency biases	the tendency of respondents to attribute more positive traits to things or people they like than to those they don't like	[Guilford 1954]
acquiescence biases	the general tendency of some individuals to agree or disagree independent of content	[Winkler, Kanouse, & Ware Jr. 1982]
mood state	the tendency of respondents to view the world in a generally positive or negative way	[Burke, Brief, & George 1993; Watson & Clark 1984]
transient mood state	the tendency of respondents to react to recent mood-inducing events	

"Item characteristic effects" refers to any bias that can be attributed to characteristics of the measurement items. Table 2 lists several characteristics of measurement items that can contribute to such effects.

"Item context effects" refers to biases that may occur because of the positioning of items in the survey, or the positioning of the items in relationship to each other [Wainer and Kiely 1987]. Table 3 lists possible causes of these effects.

Table 2. Sources of Item Characteristic Effects

Source	Definition	Reference
item social desirability	items may be constructed to reflect socially desirable perceptions	[Nederhof 1985]
item demand characteristics	items may give clues as to how to respond	
item complexity/ambiguity	items may allow multiple interpretations	[Peterson 2000]
scale formats	covariance resulting from using the same scale format [e.g., Likert scales] throughout a questionnaire	[Tourangeau, Rips, & Rasinski 2000]
scale anchors	covariance resulting from using the same scale anchors [e.g. "extremely"] throughout a questionnaire	[Tourangeau, Rips, & Rasinski 2000]
positive/negative wording	the fact that positive/negative wording may produce artifactual relationships	[Hinkin 1995]

"Measurement context effects" refers to any artifactual covariation produced from the context in which a survey instrument is completed by a respondent, such as the time, location, and medium of measurement [Bouchard 1976; Podsakoff et al. 2003; Richman, Weisband, Kiesler, and Drasgow 1999].

Table 3. Sources of Item Context Effects

Source	Definition	Reference
item priming effects	the position of the item can make the item more salient to the respondent, or may hint at the item or construct's suspected relationship to other items or constructs	[Salancik 1984]
item embeddedness	neutral items surrounded by positively or negatively worded items may take on the nature of those items	[Harrison & McLaughlin 1993]
context-induced mood	when earlier items on a questionnaire induce a mood that could potentially bias the answers to subsequent questions	[Peterson 2000]
scale length	if scales have fewer items, they are more likely to be accessible in short-term memory	[Harrison, Mclaughlin, & Coalter 1996]
grouping of items or constructs	grouping together items that measure the same construct may influence correlations	[Kline, Sulsky, & Rever-Moriyama 2000]

V. ADDRESSING METHOD EFFECTS IN SURVEY RESEARCH STUDIES

The design and procedures of a survey study can be adjusted to minimize the method bias that is introduced. Alternatively, after data have been collected, various statistical procedures may be performed to test for the presence of method variance or to account for it.

ADDRESSING METHOD EFFECTS IN STUDY DESIGN

Considering common source bias as a potential major impediment to the validity and acceptance of IS research, we first consider what may be done to alleviate this problem.

To control common source variance in the design of the study, the simplest solution is obtaining different respondents for each of the constructs. This prevents the mindset of a single informant from biasing the relationship between the variables. When it is not feasible to obtain different respondents for each construct, the second-best solution is to make sure that the measurements of the most important independent variables are taken from different respondents than those of the most important dependent variables.

Obtaining multiple respondents is not always possible, however. For example, the research may examine relationships among the attitudes or perceptions of individual subjects, in which case the measurements must come from the same person. In IS, this would be the case with studies based on the Technology Acceptance Model (TAM) [Davis 1989], for instance.

Another problem is that measurements from separate sources may be linked together through an identifying variable. If that identifying variable compromises anonymity, biases could be introduced.

Moreover, this approach may be costly for the researcher as the requirement of having multiple respondents may negatively influence the response rate.

When data cannot be obtained from separate sources, procedures that separate the measurements of the independent and dependent variables temporally, proximally, psychologically, or methodologically may be used. The potential disadvantages of such procedures are that they may allow contaminating factors to intervene between measurement of the independent and dependent variables, they may cause respondent attrition, they are costly, and they may confound the true underlying relationship if the separation has an influence on the respondents' answers.

Item characteristic effects may be best addressed if the researcher is familiar with the sources of these effects, as described in Table 2, and through the reuse of previously developed and validated scales.

McLaughlin [1999] has summarized ways to address item context effects due to the fact that individuals use information that is available in their memory at the time at which they give a judgment (a perception, attitude or the perception of a behavior) [Schwarz and Bless 1992; Strack and Martin 1987]. Respondents utilize both "chronically available" information—that which is accessible every time that an issue is brought to mind—and "temporarily available" information— that which is only accessible at certain times [Tourangeau 1992]. Avoiding item context effects requires that the researcher be aware of each and that he/she "test" preliminary items using these criteria. Sudman et al. [1996] and McLaughlin [1999] provide a good summary of such approaches. Table 4 presents a summary of the actions that they suggest.

Effective pretesting of the instrument is a basic method for assessing the impact of all of these approaches to preventing or minimizing method effects. Pretests may either be done in a "think aloud" mode or through focus groups [Sudman et al. 1996].

ADDRESSING METHOD EFFECTS THROUGH ANALYTIC ASSESSMENTS

When it is impractical to fully address method effects by altering the study design in these ways, various statistical techniques may be used to assess method variance. Several such techniques are reviewed by Podsakoff et al. [2003], who provide recommendations about which statistical techniques to use in different circumstances. Two statistical techniques which Podsakoff et al. [2003] do not recommend are Harman's single-factor test and various partial correlation procedures. In Harman's single-factor test all study variables are loaded into an exploratory factor analysis. If the unrotated factor solution indicates that a majority of the covariance among study variables is accounted for by a single factor, this is taken as evidence of common method bias. Conversely, if a single factor does not emerge, researchers conclude that their study does not suffer from common method bias. This is not a valid conclusion. In effect, this test can only detect

method bias if it accounts for most or all of the covariance among the measures, which is unlikely even if method bias is a serious problem in the study. In our sample we found that Harman's single-factor test was used in a majority of the few IS survey research studies that dealt analytically with method bias.

Table 4. Methods for Addressing Item Effects

Source of Item Effect	Suggested Action	Reference
social desirability effects	Assure respondents of the anonymity and confidentiality of their responses	
scale formats/anchors	Use scale formats that range from negative to positive [e.g., -3 to +3] where the two poles refer to opposite attributes, and all positive scales [e.g., 1 to 7] where the scale refers to degrees of the same attribute	[Schwarz, 1996; Schwarz et al. 1991]
scale formats/anchors	Use open-ended responses for questions about frequency and then standardize later. Words such as "frequently" and "occasionally" have different meanings to different respondents and in different contexts	[Menon et al., 1995; Schwarz 1999]
item priming effects, item embeddedness, context-induced moods, grouping of items or constructs	Randomizing the order of items both within instruments and across respondents may prevent systematic effects of these types	[Tourangeau & Rasinski 1988]
item embeddedness	Avoid using occasional reverse-coded items among positively-worded questions. Although some researchers have used reverse-coded items to deter acquiescent responses, they may also cause significant error	[Harrison et al. 1993]
multiple item characteristic and context effects	Use Lessler and Forsyth's [1996] coding system[1] to identify items that may be affected by the context of the survey	[Lesser & Forsyth 1996]

Partial correlation procedures are those that use some measure of what is assumed to be the cause of common method variance as a covariate in the statistical analysis. One partial correlation procedure involves measuring specific causes of method variance such as affective states or social desirability and then partialling out their effects on the other study variables. This technique is limited because it only addresses one or a few potential causes of common method bias. Another weakness is that this technique focuses on the construct level and thus cannot tell

---

[1] Drawing on a cognitive model of survey response, this method guides trained coders through an item-by-item review of question features that may contribute to response error due to item context (i.e. the codes indicate whether the question asks for a current or past behavior or attitude or interferes with the respondent's capacity to understand the question). The frequency distribution of codes is tabulated, then studied to guide evaluation, testing or revision of the survey instrument.

us anything about the many potential sources of common method bias that occur at the item level.

A second partial correlation procedure is performed by including a variable in the study that is theoretically unrelated to any other study variables. Any observed relationship between this variable and the other study variables is assumed to be caused by common method variance, so the average correlation between this "marker" variable and other study variables is partialled out of the analysis. Similar to the previous technique, this technique is limited because it can only account for some of the potential causes of method bias. For example, this technique could not account for covariance caused by social desirability bias, implicit or explicit theories held by the respondent, or the consistency motif. Another weakness of this approach is that it assumes that common method variance can only inflate relationships among study variables. As we have seen, this is not the case; common method variance can also make relationships among study variables seem smaller than they should be. Finally, this method focuses on the construct level, thus ignoring potential causes of common method bias that occur at the individual item level.

A third partial correlation technique is to assume that the first unrotated factor that emerges from an exploratory factor analysis is caused by common method bias and partial out its scale score. One problem with this technique is that the first unrotated factor may not just represent method bias; it may also represent true relationships among study variables. Another weakness is that, like the other partial correlation techniques, it focuses on the construct level and thus ignores measurement error and biases that may affect individual items. Finally, a weakness this technique shares with all the partial correlation techniques is that it assumes that common method bias affects all study constructs in the same way.

Because of these problems with Harman's single-factor test and the partial correlation procedures, Podsakoff et al. [2003] recommend statistical approaches which model a method factor or factors as latent variables. They describe two single latent method factor approaches and two multiple latent method factor approaches.

In the single unmeasured latent method factor approach, a single common methods factor is included in the model as a first-order factor for which all items serve as indicators. Items are allowed to load on their theoretical constructs, as well as on this latent factor, and the significance of the structural parameters is examined both with and without the latent factor in the model. In this way, the variance of the responses to a specific measure is partitioned into three components: (1) trait, (2) method, and (3) random error. Consequently, this method allows the researcher to control for any systematic variance among the items and obtain more accurate estimates of the structural parameters. The strengths of this approach are that it does not require the researcher to specify the source of method bias, and it does not assume that method bias affects all the study constructs in the same way. One disadvantage of this approach is that the methods factor may reflect variance due to relationships between study constructs other than the one hypothesized. Another potential disadvantage is that including the methods factor may cause the model to be under-identified, making it impossible to estimate all the model's parameters [Podsakoff et al. 2003].

Another approach is the single measured latent method factor approach, in which a single proposed source of method bias, such as social desirability or positive affectivity, is measured and included in the model. The items for other model constructs are allowed to load on the methods factor as well as other constructs. This method improves upon partial correlation procedures by allowing error in the methods factor to be estimated and by modeling the effects of methods bias at the item level rather than at the construct level. In addition, it does not assume that the method factor has the same effect on all study variables. The most important disadvantage of this method is that it requires the researcher to specify and measure the most important source of methods bias [Podsakoff et al. 2003].

Because methods bias can stem from many different sources, it would seem that this approach is not broadly useful. However, it may prove useful when the researcher has reason to expect that there is a dominant source of method bias that affects the study.

Multiple first-order latent methods factors (e.g. social desirability, positive affectivity) can also be measured and included in the model, with items from other constructs expected to be affected by each measured method factor allowed to load on that factor. This approach has all the advantages of the single latent method factor approaches; it allows error in the methods factors to be estimated and models the effects of methods bias at the item level rather than at the construct level. In addition, it does not assume that the methods factors have the same affect on all study variables. The primary disadvantage of this method is that the researcher must have a good understanding of what factors could cause method bias in the study and be able to measure them [Podsakoff et al., 2003].

Podsakoff et al [2003] also discuss the multitrait-multimethod (MTMM) procedure of Campbell and Fiske [1959], which involves constructing a correlation matrix to facilitate the assessment of construct validity. The MTMM procedure involves measuring each of several concepts (called *traits* by Campbell and Fiske) by each of several methods. By examining different portions of the resulting correlation matrix (i.e. the diagonal correlations within or across different blocks of the correlation matrix), a researcher can assess various dimensions of construct validity. For example, if the correlations of two traits measured by the same method are high, it indicates that measuring different things with the same method results in correlated measures. Or, in more straightforward terms, you have a strong "methods" factor. The MTMM matrix can be used to assess the true relationships between the traits, even when method variance and measurement error are present [Bagozzi et al. 1991]. The method that Campbell and Fiske proposed for evaluating MTMM data has been criticized, however, because it does not allow specific estimates of the amount of variance attributable to each of the three sources – underlying trait, method effects, and random error [Schmitt, Cole, and Saari 1977].

To overcome this limitation in the Campbell-Fiske procedure for analyzing MTMM data, other researchers have applied confirmatory factor analysis [CFA] to data. Widaman [1985] developed a nested models procedure to be used to both test for the presence of trait and method variance in MTMM data as well as estimate their magnitude. This procedure specifies four hierarchically nested models and performs chi-square goodness of fit tests to determine the presence or absence of trait variance and method variance. The first model is the *null model*, in which variance in the measures is explained by random error only. The second model is the *trait-only model*, in which the variance in the measures is explained by the trait factors and by random error. The third model is the *method-only model*, in which variance in the measures is explained by method factors and random error. The fourth model is the *trait and method model*, in which variance in the measures is explained by the traits, methods, and random error. Trait variance is present if the trait-only model has better fit than the null model and the trait and method model has better fit than the method-only model. Similarly, method variance is present if the method-only model has better fit than the null model and the trait and method model has better fit than the method-only model.

In addition to testing for the presence of trait and method variance, this method also allows for estimating their magnitude. The square of the factor loadings for the method factors indicates the percentage of the variance in the measure due to methods, and the square of the trait loadings indicates the percentage of variance in the measure due to the traits [Widaman 1985]. This procedure was used by Cote and Buckley [1987] in their meta-analysis of 70 datasets from 64 published studies in a variety of social science disciplines which found evidence of significant method variance in most of the studies.

Bagozzi and Yi [1991] point out that CFA analysis of MTMM data relies on the assumption that measure variation is a linear combination of traits, methods, and error. This assumption does no harm when the effects of common methods do not vary by trait. However, there may be situations in which methods and traits can interact in a multiplicative fashion, which would invalidate this

assumption [Campbell and O'Connell 1967, 1982]. In such cases Bagozzi and Yi recommend the Direct Product Model (DPM), proposed by Swain [1975] and extended by Browne [1984, 1989] as a better method for analyzing the MTMM data. However, the DPM method will give misleading results if the relationship between trait and method effects is additive [Bagozzi et al. 1991].

VI. APPROACH TO ASSESSING METHOD EFFECTS IN IS SURVEY RESEARCH

The methodology used in making our empirical assessment of method bias in IS survey research is adapted from that used by Smith [2002] in his study of nonresponse bias reporting in leading social science journals and that used by King and He [2006] in the IS context. It reflects the notion that "Every academic field is marked by its literature" [McLean 1996, p 151].

Three journals, often considered to be "A-level" in the IS field, were selected for assessment— *Management Information Systems Quarterly (MISQ)*, *Information Systems Research (ISR)*, and the *Journal of Management Information Systems (JMIS)* [Vessey, Ramesh, and Glass 2002]. Other "A journals," such as *Management Science (MS)* and *Organization Science (OS)*, were initially intended to be included in the study, but were subsequently excluded because they publish papers on a wide range of topics beyond IS and because the judgmental assessment of multiple raters of which articles qualified to be IS articles was unreliable. This decision was at odds with the method used by Vessey et al. [2002] who "... coded only those articles (from *MS* and *Decision Sciences*) that we considered to be IS articles." [p. 136]. Our different judgment is probably, in part, due to the rapid expansion of the scope of the IS field since their 1995-99 assessment.

Thus, for purposes of this study, we choose to operationally define "IS survey research" as "those survey research papers that appear in one of the three A-rated solely IS journals." We believe that this pragmatic approach is defensible despite the fact that some top IS papers appear in journals that are not solely IS journals and that some excellent IS papers appear in IS journals that are not top-rated by all. We expect that studies appearing in these three top journals reflect the enacted "best research practices" of the field.

We selected the time period January 1999 to September 2005[2] somewhat arbitrarily, because we wished to make the assessment on a recent and not historical, basis. All issues of each of the three journals for this time period were inspected, and a total of 128 quantitative empirical survey studies were carefully reviewed.

DEVELOPING A CODING SCHEME

After reviewing the potential sources of method bias and various approaches to address each bias, we came up with an initial coding scheme. A pilot test was conducted to test this classification scheme using 20 papers randomly selected from the sample. Two researchers independently coded the pilot dataset and compared their codings. The consensus rate was 65 percent. The two researchers then got together and reconciled these discrepancies. A third researcher joined the process whenever the two coders could not agree on a particular question until they reached 100 percent agreement. The coding scheme was refined based on the pilot study and a second pilot test was run using another 20 papers randomly selected from the sample. This time the two independent coders reached a 95 percent consensus rate and the remaining discrepancies were quickly resolved. Some new subcategories were added to the initial coding scheme and other subcategories were dropped or combined.

This final coding scheme, as shown in the appendix, was then applied to all 128 papers in the sample. A relational database was created to store the coded results. After two researchers

---

[2] *MISQ*, 23(1)-29(3); *ISR*, 10(1)-16(3); *JMIS* 15(4)-22(2)

independently coded all the papers, they reached a consensus rate of 92 percent. The remaining 8 percent disagreement was resolved after they met and compared their codings.

VII. CONCLUSIONS

The results, shown in detail in the appendix, are discussed below.

COMMON SOURCE BIAS

The results indicate that there is potentially a great deal of common source bias in IS survey research. In almost 83 percent (106 of 128) of the studies only a single respondent was used to respond to items for all the study variables.

In about 34 percent (44 of 128) of the papers, only one respondent was used for all variables in situations in which common source bias could have been avoided if multiple respondents had been used. Although over 58 percent (62 of 106) of the papers that utilized a single respondent were judged to have done so out of necessity, in the vast majority (45 of 62) of these instances, no mention was made of possible common source bias as a limitation of the study. In only about 24 percent (25 of 106) of the studies with a single respondent for all variables was some action taken to attempt to assess or ameliorate the bias. In 11.3 percent of these (12 of 106), the potential for method bias is mentioned but no action was taken.

In our dataset, only 17 percent (22 of 128) of the papers directly addressed common source bias by collecting data from multiple sources for all or some of their key variables. Of those, only five papers collected data from multiple sources for all variables.

Table 5 summarizes the detailed results of our assessments of common rater effects.

Table 5. Common Source Bias Analysis Results

	USR: 62 [48.4%][†]	**ASR**: 44 [34.4%][†]	**MRA**: 5 [3.9%][†]	**MRB**: 17 [13.3%][†]	**Total**: 128
No Mention	45 [72.6%][*]	24 [54.6%][*]	n/a	n/a	69 [53.9%][†]
Discussed but no action	4 [6.5%][*]	8 [18.2%][*]	n/a	n/a	12 [9.4%][†]
Statistical methods	2 [3.2%][*]	5 [11.4%][*]	1 [20%][*]	3 [17.6 %][*]	11 [8.6%][†]
Non-Statistical methods	3 [4.8%][*]	6 [13.6%][*]	3 [60%][*]	1 [5.9 %][*]	13 [10.2%][†]
Time Dimension Introduced	10 [16.1%][*]	2 [4.6%][*]	0 [0%][*]	2 [11.8 %][*]	14 [10.9%][†]

† The percentage here refers to the ratio of the number of papers in this category relative to the total number of papers.

* The percentage here refers to the ratio of the number of papers that use a particular method to address common source bias relative to the total number of papers in this category.

USR: Unavoidable use of single respondent for all items

ASR: Avoidable use of single respondent for all items

MRA: Multiple respondents across all independent and dependent variables

MRB: Different respondents for independent and dependent variables

These results indicate that common source bias is a significant and largely unaddressed issue in IS survey research. They demonstrate that common source bias may exist in most survey studies and that in the vast majority of studies with this potential bias, little is done to alleviate the potential problem.

One way of alleviating some of the problems associated with common source bias is to use objective data for some of the variables. Forty-six of 128 studies (about 36 percent) used some objective measures.   Table 6 shows the detailed assessments concerning the use of demographic or objective data.

Table 6. Use of Demographic or Objective Data

Category	USR	ASR	MRA	MRB	Total
# of Papers in Category	62 [49.2%][†]	44 [33.6%][†]	5 [3.9%][†]	17 [13.3%][†]	128
Papers that use Demographic or Objective Data for Some Study Variables	14 [22.6%][*]	21 [47.2%][*]	1 [20%][*]	10 [58.8%][*]	46 [35.9%][†]

[†] The percentage refers to the ratio of the number of papers in this category relative to the total number of papers

[*] The percentage here refers to the ratio of the number of papers that use objective measures relative to the total number of papers in this category

INAPPROPRIATE SURVEY DESIGN

A second method bias source that is problematic in IS survey research has to do with other aspects of the survey design, such as all respondents having the same background when this was not required by the objectives of the study. About 20 percent of the studies had this characteristic and did not provide any justification for it.

Similarly, almost 74 percent of the surveys were conducted at one point in time; only about 26 percent introduced "temporal distance" either because it was required for the purpose of the study (e.g., "before and after") or as a part of the design with the explicit objective of reducing the potential for method bias. Table 7 shows the details related to how time was handled in the studies.

Most studies (nearly 90 percent) used a single medium to collect data; however, only a very few (about 4 percent) used a single location – both potential sources of bias. An "ideal" way to collect data to avoid method bias is to collect it at different times or using different media, even if not required by the objectives of the research. Only 11 percent of the studies introduced a temporal dimension in collecting data and only 10 percent introduced some combination of media (e.g. online and paper survey).

Table 7. Time Handling in the Survey Studies

Category	Definition	# of Papers [%]
SCO	*If surveys completed all at same time [e.g. in a classroom setting]*	8 [6.25%]
SAO	*If surveys are administered once [e.g. a one-part mail survey] but cannot determine exactly when surveys are completed*	86 [67.2%]
TDI	*Temporal distance introduced as part of the research design*	14 [10.94%]
TDR	*Temporal distance required for the purpose of the study [e.g a paper that studies change in personal belief and attitude]*	10 [7.8%]
Unknown	*The paper did not indicate how surveys are administered*	10 [7.8%]

INAPPROPRIATE WORDING AND STRUCTURE OF THE SURVEY INSTRUMENT

With regard to the wording and structure of survey instruments, the vast majority of papers (87.5 percent) provided all of the items. However, most did not make clear whether the items were actually presented to respondents in the order presented in the paper. Table 8 shows the summary of item availability.

Table 8. Item Availability Results

Category	Definition	# of Papers [%]
AL	*If all the items used in the study were included in the paper*	112 [87.5%]
EX	*If only example items used in the study were provided in the paper*	10 [7.8%]
NO	*If no items were included in the paper*	6 [4.7%]

Of those studies that provided all of the items, about 25 percent of the studies had items and relationships that appeared to have great potential for implicit theory bias or social desirability bias (24 of 122). It was our judgment that these studies focused on situations in which it was virtually certain that significant relationships would be found for one of those reasons.

Other problems with survey items were found to be present in a small, but significant, proportion of the papers. These problems include item priming and context-induced moods. The full information necessary to replicate a study was often not provided. For instance, only about 64 percent provided scale anchors. In order for reviewers to assess whether the use of common scale formats and/or anchors introduces artifactual covariation, this information must be provided.

METHODS USED TO DETECT/AMELIORATE METHOD BIAS

Researchers who do not design studies to minimize method bias should use statistical techniques to detect and assess such biases. The use of statistical techniques to detect method bias was not extensive, and the statistical technique most often used was Harman's single-factor test, which is a very insensitive test. Non-statistical methods such as a contrasting vignette study design may also be used to address method bias [Jarvenpaa and Staples 2001].

SUMMARY AND RECOMMENDATIONS

The overall method bias problem is perhaps best illustrated by the fact that only 62.5 percent of the studies even mentioned the possibility of method bias. However, since about 18 percent not only mentioned this as a potential problem but attempted to do something about it, it is clear that better quality can be introduced into studies by improved study design and additional testing and discussion of items and instruments.

Overall, these data suggest that there is a significant issue with method bias in IS survey research. This is so potentially serious that it might negate the validity and/or credibility of a significant proportion of the research that has been done in the field.

Since survey research represents more than half of all IS research [King and He 2005], these analyses further suggest that it is likely that the issue of method bias pervades the field. Thus, our response to Crampton and Wagner's [1994] admonition to identify fields in which method bias may exist is, "We have identified the enemy and he/she is us!"

To overcome this, researchers must first be made aware of the potential seriousness of method bias. Additional education is suggested since we believe that one of the reasons for the neglect of method bias is that many researchers do not completely understand its many potential sources. That is one of the primary purposes of this paper. We believe that the data that we have provided form the empirical basis substantiating a need for a warning to IS researchers concerning the potential seriousness of method bias. Once researchers are made aware of this,

they can develop methods and practices for eliminating method bias, to the degree that is feasible, from their work.

IS researchers should also focus on the development of new nonstatistical techniques to minimize method bias [eg., Jarvenpaa and Staples 2001] and novel statistical techniques to detect and assess method bias [eg., Keil et al. 2000].

In the short run, perhaps the best way to accomplish increased awareness and change is for the editors and reviewers for IS journals to make the addressing of method bias be a requirement for the consideration of papers for publication. First, journals should insist that authors provide the actual survey instruments so that the possibility of some varieties of method bias that depend on sequence of items can be assessed by reviewers. Screening reviews should check for potential sources of method bias and editors should ask authors to address these issues in their papers in revisions made prior to detailed review.

This might even be supported by a checklist that is provided to associate editors to enable them to efficiently review for method bias. Another step might be to make such a checklist available to authors and ask them to certify that they have given attention to the various forms of method bias and to include some discussion of what they did in the paper. This could lead to improvements in the general state of awareness of method bias and possibly, to the development of new techniques for detecting and ameliorating method bias. In this fashion, reviewers will have the best possible information concerning whether method bias is likely to exist and what has been done, if anything, to address it.

In all papers in which perceptual measures are used, we believe that the minimum criterion for reviewing a paper should be that level of potential method bias be discussed. Associated with this might be a requirement that there be discussion of why a potential source of method bias was not "designed out" of the study. As well, statistical checks for method bias—particularly for common source bias—should be required.

Since there are always tradeoffs between the economics of conducting research and the rooting out of potential method bias, such criteria would enable authors to argue that they have made good faith attempts to minimize such bias. For instance, we believe that any paper that uses a single respondent for both independent and dependent variables should contain an explanation of why the study was designed in this manner.

We believe that these data concerning the potential for serious method effects suggest the need for changes in the criteria that are used to judge papers for publication. Only when researchers are fully informed about method effects and authors are not able to ignore method effects, as they often do now, will papers with potentially serious method effects be dealt with adequately.

REFERENCES

Bagozzi, R. P. and Y. Yi. (1991). "Multitrait-Multimethod Matrices in Consumer Research," *Journal of Consumer Research,* (17)4, pp. 426-439.

Bhattacherjee, A. and G. Premkumar. (2004). "Understanding Changes in Belief and Attitude toward Information Technology Usage: A Theoretical Model and Longitudinal Test," *MIS Quarterly*, (28)2, pp 229-254.

Bouchard, T. J. (1976). *Field Research Methods: Interviewing, Questionnaires, Participant Observation, Systematic Observation, Unobtrusive Measures*. New York: Wiley.

Browne, M. W. (1984). "The Decomposition of Multitrait-Multimethod Matrices," *The British Journal of Mathematical and Statistical Psychology*, (37), pp. 1-21.

Browne, M. W. (1989). "Relationships between an Additive Model and a Multiplicative Model for Multitrait-Multimethod Matrices," In R. Coppi & S. Bolasco (Eds.), *Multiway Data Analysis* (pp. 507-520). Amsterdam: Elsevier.

Burke, M. J., A. P. Brief, and J. M. George. (1993). "The Role of Negative Affectivity in Understanding Relations between Self-Reports of Stressors and Strains: A Comment on the Applied Psychology Literature" *Journal of Applied Psychology*, (78)3, pp. 402-412.

Burstin, K., E. B. Doughtie, and A. Raphaeli. (1980). "Contrastive Vignette Technique: An Indirect Methodology Designed to Address Reactive Social Attitude Measurement," *Journal of Applied Social Psychology*, (10)2, pp. 147-165.

Campbell, D. T. and D. W. Fiske. (1959). "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," *Psychological Bulletin*, (56) March, pp. 81-105.

Campbell, D. T. and E. J. O'Connell. (1967). "Methods Factors in Multitrait-Multimethod Matrices: Multiplicative Rather than Additive?" *Multivariate Behavioral Research*, (2), pp. 409-426.

Campbell, D. T. and E. J. O'Connell. (1982). "Methods as Diluting Trait Relationships Rather than Adding Irrelevant Systematic Variance," In D. Brinberg & L. Kidder (Eds.), *Forms of Validity in Research* (pp. 93-111). San Francisco: Jossey-Bass.

Cote, J. A. and M. R. Buckley. (1987). "Estimating Trait, Method, and Error Variance: Generalizing Across 70 Construct Validation Studies," *Journal of Marketing Research*, (24)3, pp. 315-318.

Crampton, S. M. and J. A. Wagner. (1994). "Percept-Percept Inflation in Microorganizational Research: An Investigation of Prevalence and Effect," *Journal of Applied Psychology*, (79)1, pp. 67-76.

Davis, F. D. (1989). "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, (13)3, pp. 318-340.

Devaraj, S., M. Fan, and R. Kohli. (2002). "Antecedents of B2C Channel Satisfaction and Preference: Validating e-Commerce Metrics," *Information Systems Research*, (13)3, pp. 316-333.

Fiske, D. W. (1982). "Convergent-Discriminant Validation in Measurement and Research Strategies" In D. Brindberg, & L. Kidder [Eds.], *New Directions for Methodology of Social and Behavioral Science: Forms of Validity in Research*: pp. 72-92. San Francisco: Jossey-Bass Publishers.

Ganster, D. C., H. W. Hennessey, and F. Luthans. (1983). "Social Desirability Response Effects: Three Alternative Models," *Academy of Management Journal*, (26)2, pp. 321.

Gattiker, T. F. and D. L. Goodhue. (2005). "What Happens after ERP Implementation: Understanding the Impact of Inter-Dependence and Differentiation on Plant-Level Outcomes," *MIS Quarterly*, (29)3, pp. 559-585.

Guilford, J. P. (1954). *Psychometric Methods*. New York: McGraw-Hill.

Harrison, D. A. and M. E. McLaughlin. (1993). "Cognitive Processes in Self-Report Responses: Tests of Item Context Effects in Work Attitude Measures," *Journal of Applied Psychology*, (78)1, pp. 129-140.

Harrison, D. A., M. E. Mclaughlin, and T. M. Coalter. (1996). "Context, Cognition, and Common Method Variance: Psychometric and Verbal Protocol Evidence," *Organizational Behavior & Human Decision Processes*, (68)3, pp. 246-261.

Hinkin, T. R. (1995). "A Review of Scale Development Practices in the Study of Organizations," *Journal of Management*, (21)5, pp. 967-988.

Jarvenpaa, S. L. and D. S. Staples. (2001). "Exploring Perceptions of Organizational Ownership of Information and Expertise," *Journal of Management Information Systems*, (18)1, pp. 151-183.

Johns, G. (1994). "How Often Were You Absent? A Review of the Use of Self-Reported Absence Data," *Journal of Applied Psychology*, (79)4, pp. 574-591.

Karimi, J., T. M. Somers, and Y. P. Gupta. (2004). "Impact of Environmental Uncertainty and Task Characteristics on User Satisfaction with Data," *Information Systems Research*, (15)2, pp. 175-193.

Karimi, J., A. Bhattacherjee, Y. P. Gupta, and T. M. Somers. (2000). "The Effects of MIS Steering Committees on Information Technology Management Sophistication," *Journal of Management Information Systems*, (17)2, pp. 207-230.

Keil, M., J. Mann, and A. Rai. (2000). "Why Software Projects Escalate: An Empirical Analysis and Test of Four Theoretical Models," *MIS Quarterly,* (24)4, pp. 631-664.

King, W. R. and J. He. (2005). "External Validity in IS Survey Research," *Communications of the AIS, (*16), pp. 880-894.

Kline, T. J. B., L. M. Sulsky, and S. D. Rever-Moriyama. (2000). "Common Method Variance and Specification Errors: A Practical Approach to Detection," *Journal of Psychology*, (134)4, pp. 401.

Lee, H. and B. Choi. (2003). "Knowledge Management Enablers, Processes, and Organizational Performance: An Integrative View and Empirical Examination," *Journal of Management Information Systems*, (20)1, pp. 179-228.

Lesser, J. T. and B. H. Forsyth. (1996). *A Coding System for Appraising Questionnaires*. San Francisco: Jossey-Bass.

Luong, A. and S. G. Rogelberg. (1998). "How to Increase Your Survey Response Rate," *The Industrial-Organizational Psychologist*, (36)1, pp. 61-65.

McLaughlin, M. E. (1999). "Controlling Method Effects in Self-Report Instruments," *Research Methods Forum*, (4) Summer.

McLean, G. N. (1996). "Refereed Journals: The Cornerstone of a Developing Profession" In E. Holton [Ed.], *Proceedings of the Academy of Human Resource Development Annual Conference*. Minneapolis.

Menon, G., P. Raghubir, and N. Schwarz. (1995). "Behavioral Frequency Judgments: An Accessibility-diagnosticity Framework," *Journal of Consumer Research* (22)2, pp. 212.

Nederhof, A. J. (1985). "Methods of Coping with Social Desirability Bias: A Review," *European Journal of Social Psychology*, 15, pp. 263-280.

Peterson, R. A. (2000). *Constructing Effective Questionnaires*. Thousand Oaks, CA: Sage.

Podsakoff, P. M. and D. W. Organ. (1986). "Self-Reports in Organizational Research: Problems and Prospects," *Journal of Management*, (12)4, pp. 531-544.

Podsakoff, P. M., S. B. MacKenzie, J.-Y.Lee, and N. P. Podsakoff. (2003). "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology*, (88)5, pp. 879.

Ravichandran, T. and A. Rai. (1999). "Total Quality Management in Information Systems Development: Key Constructs and Relationships," *Journal of Management Information Systems*, (16)3, pp. 119-155.

Richman, W. L., S. Weisband, S. Kiesler, and F. Drasgow. (1999). "A Meta-Analytic Study of Social Disability Distortion in Computer-Administered Questionnaires, Traditional Questionnaires, and Interviews," *Journal of Applied Psychology*, (84)5, pp. 754-775.

Ryan, S. D., D. A. Harrison, and L. L. Schkade. (2002). "Information-Technology Investment Decisions: When Do Costs and Benefits in the Social Subsystem Matter?" *Journal of Management Information Systems*, (19)2, pp. 85-127.

Salancik, G. R. (1984). "On Priming, Consistency, and Order Effects in Job Attitude Assessment: With a Note on Current Research," *Journal of Management*, 10, pp. 250-254.

Schmitt, N., B. W. Cole, and B. B. Saari. (1977). "A Review and Critique of Analyses of Multitrait-Multimethod Matrices" *Multivariate Behavioral Research*, (12)October, pp. 447-478.

Schwarz, N. and H. Bless. (1992). *Constructing Reality and its Alternatives: Assimilation and Contrast Effects in Social Judgment*. Hillsdale, NJ: Erlbaum.

Schwarz, N. (1996). *Cognition and Communication: Judgmental Biases, Research Methods, and the Logic of Conversation*, Lawrence-Erlbaum, Mahwah, New Jersey.

Schwarz, N. (1999). "Self-Reports: How the Questions Shape the Answers," *American Psychologist* (54), pp. 93-105.

Schwarz, N., B. Knäuper, H.-J. Hippler, E. Noelle-Neumann, and L. Clark. (1991). "Rating Scales: Numeric Values May Change the Meaning of Rating Scales," *Public Opinion Quarterly* (55)4, pp. 570-582.

Smith, T. W. (2002). "Reporting Survey Nonresponse in Academic Journals," *International Journal of Public Opinion Research*, (14)4, pp. 469-474.

Smither, J. W., R. Buda, and H. Collins. (1989). "When Rate Satisfaction Influences Performance Evaluations: A Case of Illusory Correlation," *Journal of Applied Psychology*, (74)4, pp. 599.

Son, J. Y., S. Narasimhan, and F. J. Riggins. (2005). "Effects of Relational Factors and Channel Climate on EDI Usage in the Customer-Supplier Relationship," *Journal of Management Information Systems*, (22)1, pp. 321-353.

Strack, F. and L. L. Martin. (1987). *Thinking, Judging, and Communicating: A Process Account of Context Effects in Attitude Surveys*. New York: Springer-Verlag.

Sudman, S., N. M. Bradburn, and N. Schwarz. (1996) *Thinking About Answers: The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey Bass.

Swain, A. J. (1975). *Analysis of Parametric Structure for Variance Matrices.* Unpublished dissertation, Department of Statistics, University of Adelaide.

Thatcher, J. B. and P. L. Perrewé. (2002). "An Empirical Examination of Individual Traits as Antecedents to Computer Anxiety and Computer Self-Efficacy," *MIS Quarterly*, (26)4, pp. 381-396.

Thatcher, J. B., L. P. Stepina, and R. J. Boyle. (2002). "Turnover of Information Technology Workers: Examining Empirically the Influence of Attitudes, Job Characteristics, and External Markets," *Journal of Management Information Systems*, (19)3, pp. 231-261.

Thong, James Y. L. (1999). "An Integrated Model of Information Systems Adoption in Small Businesses," *Journal of Management Information Systems*, (15)4, pp. 187-214.

Tourangeau, R. (1992). *Attitudes as Memory Structures: Belief Sampling and Context Effects*. New York: Springer-Verlag.

Tourangeau, R. and K. A. Rasinski. (1988). "Cognitive Processes Underlying Context Effects in Attitude Measurement," *Psychological Bulletin*, (103), pp. 299-314.

Tourangeau, R., L. J. Rips, and K. A. Rasinski. (2000). *The Psychology of Survey Response*. Cambridge, England: Cambridge University Press.

Vessey, I., V. Ramesh, and R. L. Glass. (2002). "Research in Information Systems: An Empirical Study of Diversity in the Discipline and Its Journals," *Journal of Management Information Systems* (19)2, pp. 129.

Wainer, H., and G. L. Kiely. (1987). "Item Clusters and Computerized Adaptive Testing: A Case for Testlets" *Journal of Educational Measurement*, 24, pp. 185-201.

Watson, D. and L. A. Clark. (1984). "Negative Affectivity: The Diposition to Experience Negative Aversive Emotional States," *Psychological Bulletin*, (96), pp. 465-490.

Widaman, K. F. (1985). "Hierarchically Nested Covariance Structure Models for Multitrait-Multimethod Data," *Applied Psychological Measurement*, (9),pp. 1-26.

Winkler, J. D., D. E. Kanouse, and J. E. Ware, Jr. (1982). "Controlling for Acquiescence Response Set in Scale Development," *Journal of Applied Psychology*, (67)5, pp. 555.

APPENDIX: CODING SCHEME AND RESULTING DATA

In this section, we provide a detailed description of the coding scheme as well as the results of its application. The coding scheme involves three categories, each with multiple subcategories to identify the different sources of method biases. The first category, *common source bias* (five subcategories) codes the papers with respect to their susceptibility to common rater effects (see Table 1), and actions taken to avoid such effects. The second category, *potential bias resulting from inappropriate survey design* (four subcategories), codes the papers with respect to their susceptibility to measurement context effects, and actions taken to prevent measurement context effects. The third category, *potential bias resulting from inappropriate wording and structure of the instrument* (three subcategories with ten sub-subcategories), codes the papers with respect to their susceptibility to item characteristic and item context effects (see Tables 2 and 3). Finally, a fourth category is also included to summarize how papers in our sample deal with various sources of method bias.

CATEGORY 1: COMMON SOURCE BIAS

The first category deals with whether there is potential bias in the paper caused by the use of a single respondent for multiple items. For papers that do use single respondents to respond to all the survey items, we further distinguished whether the use of a single respondent is avoidable or unavoidable. For papers that don't have just a single respondent, we also classified whether the study utilized multiple respondents across all variables or different respondents [sources] for different sets of variables. This classification leads to five sub-categories. The first four are mutually exclusive categories related to the strategies used for collecting data from respondents.

Unavoidable Use of Single Respondent for All Items (USR)

We coded this item as true when it was necessary to use a single respondent for all measurement items in a study because of the nature of the constructs being measured. If a research model describes relationships among individuals' beliefs, attitudes, or intentions, only the individual is qualified to respond to the measurement items, and thus the use of a single respondent in such studies is unavoidable.

For example, Thatcher and Perrewé [2002] examined personal innovativeness in IT, trait anxiety, and negative affectivity as antecedents of computer anxiety and computer self-efficacy. Because each of these constructs describes individual attitudes or traits, we judged that the individual was the only feasible source from which to measure these constructs and coded the article as exhibiting the unavoidable use of a single respondent for all items. All TAM-related survey studies also belong to this category.

In our dataset, 62 out of 128 papers (48.4 percent) fall into this category. The presence of common source bias is considered to be largely unavoidable in these papers. However, these studies differ considerably in terms of the ways they acknowledge and/or address the potential bias. Of these 62 papers, 45 (72.6 percent) do not mention the potential common source bias that could exist in their studies, nor do they take any steps to detect or minimize the impact of the potential common source bias. Of the remaining 17 papers, four acknowledge the possible existence of common source bias but take no actions to address the issue. The other 13 papers use various approaches to address the potential common method bias: 10 papers introduce a time dimension into their studies (i.e. collect data at different points of time). Two papers use statistical methods to detect the presence of measurement bias [one uses Harman's single-factor test and the other uses ANOVA]. Three papers use non-statistical methods [such as multiple studies, e.g. Bhattacherjee and Premkumar 2004, or a contrastive vignette, e.g., Jarvenpaa and Staples 2001] to alleviate the impact of potential common source bias.[3]

Avoidable Use of Single Respondent for All Items [ASR]

We coded this item as true when a single respondent was used for all measurement items in a study, and this was judged to be avoidable.

Survey studies are subject to the problem of ASR if aggregate-level data (e.g., those representing organizational- or team-level constructs) are collected from a single respondent. For example, Karimi et al. [2000] examined the relationship between the presence and roles of IT steering committees and the level of IT management sophistication within firms. For each of the 213 firms in the sample, a corporate-level IS executive was the sole respondent. ASR is deemed to be a more severe problem in such a study as it could have been avoided, or minimized, in the research design phase.

In our dataset, 44 out of 128 (34.4 percent) papers fall into the ASR category, of which 24 papers fail to acknowledge or address even the potential presence of common source bias. Of the remaining twenty papers, eight discuss the potential for common source bias in their studies, but do not further examine this issue. Two studies collected data at different time points to minimize common source bias; five papers used statistical methods to estimate the measurement bias (i.e. two-sample Kolmogorov-Smirnov Z test, the multitrait-multimethod (MTMM) technique and ANOVA); six papers used nonstatistical methods to alleviate the impact of potential common source bias.

Multiple Respondents across Independent and Dependent Variables (MRA)

We coded this item as true when a study had multiple respondents for each data point, but each respondent was used to measure both the independent and dependent variables. The aggregation of data from the same respondent may introduce considerable common source bias, although the use of multiple respondents can alleviate this problem to some extent.

For example, Lee and Choi (2003) examined the relationships between knowledge management enablers, knowledge management processes, and several measures of organizational performance, including knowledge management satisfaction, return on assets, return on sales,

---

[3] Note that these numbers do not add up to 13, as a paper might use multiple approaches to address the potential common source bias.

and organizational effectiveness. The unit of analysis was the organization. Responses were received from 426 managers in 58 firms. Each respondent provided perceptions of items measuring both the independent and dependent variables. Multiple responses from a single firm were aggregated for use as the organizational indicator. Inter-rater reliability and agreement were assessed to support the appropriateness of aggregating the responses. This study was coded as having multiple respondents across independent and dependent variables. In our dataset, only five papers (3.9 percent) adopted this strategy to reduce common source bias in their study.

Multiple Respondents for Independent and Dependant Variables (MRB)

We coded this item as true when a study utilized different respondents for independent and dependent variables. For example, Karimi et al. [2004] examined the impacts of environmental uncertainty and task characteristics on user satisfaction with data. The respondents were paired samples of 77 CEOs and 166 senior managers from 77 firms. CEOs responded to items measuring environmental uncertainty. Managers responded to items measuring user satisfaction (with data, IS, and IS support) and task characteristics. Because not all the independent and dependent variables were measured from the same respondent, this study was classified as multiple respondents between independent and dependent variables. This approach can significantly reduce the probability of introducing common source bias as the correlations of responses from different groups of respondents are minimized.

In our dataset, 17 papers (13.3 percent) were judged to meet this criterion. Of these 17, one paper used a non-statistical procedure to reduce potential common source bias (meeting with respondents to discuss questions requiring a convergent view). Three papers further used some statistical techniques to test for the presence of common source bias (i.e. Harman's single-factor test or a correlation test). The data for these subcategories are given in Table 5.

Some Variables Demographic or Objective

In addition to those four mutually exclusive subcategories, we also coded whether or not some of the key variables were objective or demographic variables. The use of demographic and objective data is considered one of the ways to alleviate the impact of common method bias. This subcategory is not mutually exclusive with the other subcategories.

We coded this item as true if any of the key variables in the study were demographic or objective. Note that this applies only to key variables in the study; not to control variables. This category was included because Cote and Buckley [1987] found that concrete concepts are easier to measure and exhibit less method variance than more abstract constructs. Moreover, other researchers suggest that using variables that are demographic or objective in nature can limit the effects of method bias [Podsakoff and Organ 1986; Podsakoff et al. 2003].

For example, Thong [1999] tested a model of information systems adoption in small businesses. The survey respondents were CEOs from 166 small businesses. These CEOs responded to items measuring the independent variables, including CEO's innovativeness and knowledge, characteristics of the firm's information systems, characteristics of the firm, and the level of competition in their business environment. The CEOs also responded to items measuring the dependent variables representing the level of information systems adoption in the firm. So, this study was coded as having an avoidable single respondent for all items. However, the measures for the dependent variables were judged to be objective in nature. The first was a measure of whether or not the business was computerized. For this measure, the respondent was presented with a list of computer applications (such as accounting, inventory control, EDI, personnel and payroll, etc.) and asked to indicate which were in use in their company. The business was considered as computerized if it used at least one of the applications. The second dependent variable was a measure of the number of personal computers in use at the business and the number of software applications in use at the business.

In our dataset, we found that 46 studies (35.9 percent) adopted some objective measures for their key variables. These efforts should help to alleviate the confounding effect of common source

bias in these studies. The distribution of these papers in each of these four subcategories is provided in Table 6.

CATEGORY 2: POTENTIAL BIAS RESULTING FROM INAPPROPRIATE SURVEY DESIGN

Another potential bias in survey studies stems from inappropriate research design which arises in the selection of the survey sample and the administration of the questionnaires. We classified the sources of such bias into various subcategories.

Single Respondent Background

We coded this category as true if all the respondents in the study share the same background, but such was not required for the purposes of the research study. For, example, if all the respondents had the same professional background or were employees from the same department of an organization, we coded this category as true. If the author explained why a specific sample group is used (e.g., to control for company-specific effect or to examine the research question under a particular setting) and the common background would not bias the result, we did not code the paper as falling into this category. Twenty-six out of 128 papers (20.3 percent) in our dataset were coded for administering surveys to respondents of the same background without any statistical or theoretical reasoning to support doing so.

Time Handling

Collecting survey data all at once might produce an artifactual covariance that picks up some time-specific effect. On the contrary, a longitudinal study might reduce the potential for such a common source/method bias. In our analysis, we distinguished between gathering all data at one time, such as by putting all data on one questionnaire, and purposely introducing a temporal distance between measures into the research design to minimize common method bias. Table 7 shows the coding categories that were developed to capture these differences. We found that only 14 papers (10.9 percent) in our sample adopt a temporal dimension into their survey design to alleviate potential method bias.

Common Location

We coded this item as true if all respondents filled out the survey at the same location, which could create a response bias. For example, responses obtained in the workplace might capture some unanticipated effects such as work pressure and time constraints, which could bias the research findings. In the dataset, five studies (3.9 percent) collected survey data at the same location.

Common Medium

This item was coded as true if all surveys were delivered via a single medium, such as mail survey, online survey, phone survey, etc. In our dataset, 115 studies (89.8 percent) used a single medium for the survey instrument. A medium specific effect might confound the research results and, whenever possible, we recommend using multiple survey media to minimize such risks.

CATEGORY 3: POTENTIAL BIAS RESULTING FROM INAPPROPRIATE WORDING AND STRUCTURE OF THE INSTRUMENT

The third major category examines whether the design of the survey instrument itself might lead to potential method bias. It should be pointed out here that many of the codings in this major category were based on properties of the survey items and scales, and these were not available for every paper coded. For that reason, we also classified these papers in terms of the availability of items, scales and anchors used in the questionnaire.

Item Availability

Table 8 shows the codes that were used to categorize the availability of questionnaire items. The number in the parentheses refers to the number of papers that fall into that category. Note that if all items were available, we assumed that they were in the same order as they appeared on the instrument, unless there was evidence to the contrary.

For the papers that were coded as "AL" and "EX," we further analyzed the content of the items. The following seven sub-subcategories were used to identify some potential biases that could stem from the wording of the questionnaire items.

Implicit Theories

One of the reasons that the use of a single respondent for all items can lead to bias is that respondents may hold their own ideas about how the constructs measured in the study relate to each other, and their responses may primarily reflect these ideas. We coded this binary item as true if we judged that respondents would probably have their own theories about how the constructs in the study related to each other. For example, Thatcher et al. [2002] tested a model linking organizational commitment and turnover intention among information technology workers. The model included constructs for attitudes and job characteristics (such as organizational commitment, job satisfaction, and perceived job characteristics), perceptions of external markets) such as perceived competitiveness of pay and perceived job alternatives), and turnover outcomes (turnover intention and turnover). We coded this article as true for implicit theories because we believe it is likely that most people implicitly believe that factors such as job satisfaction, job characteristics, competitiveness of pay, and job alternatives are clearly and directly related to turnover. In our dataset, six papers out of 122 in which items are available (4.9 percent) were identified as having this problem. To avoid such a bias, we recommend that the items that measure outcome variables should be presented earlier in the survey than those that measure antecedents.

Probable Social Desirability Bias

Respondents to a survey may respond in what they believe to be a socially acceptable or appropriate manner. Thus, if there are items on a survey that reflect behaviors, attitudes, or perceptions that are particularly desirable or undesirable socially, we coded this binary category as true. For example, Ravichandran and Rai [1999] examine the relationship among key quality management constructs. In their study, responses from senior IS executives were used to measure several constructs related to quality management in systems development in their organizations, including IS management commitment to quality, quality policy and goals, quality orientation of reward schemes, commitment to skill development, and several others. We judged that most IS executives would feel that it is desirable to strive for quality in their IS development practices, and thus would feel some degree of social pressure to indicate this in their responses, regardless of whether or not they actually actively follow these quality practices. As such, we coded this article as true for probable social desirability bias. Twenty-four of 122 papers (19.7 percent) were identified as suffering from potential social desirability bias. Efforts to reduce social desirability bias include neutral wording of the items and intermixing of the items that are used to measure such constructs.

Use of Reverse-Coded Items

If the study used any reverse-coded items we coded this binary category as true. Twelve papers (9.8 percent) were found to use reverse-coded items. There is some disagreement in the literature concerning whether the use of reverse-coded items has a net positive or negative impact on the accuracy of results [Harrison and McLaughlin 1993]. Some researchers feel that reverse-coding forces respondents to focus on the directionality of each specific item; the contrary view is that if respondents do not notice the reverse-coding, inaccuracies are likely to be introduced by reverse-coding.

The following subcategories assess whether the order and the layout of the items would be likely to create some bias in responding to items. However, these statistics should be interpreted with caution as we assume that the order of items provided in the paper is the same as the actual order of items in the questionnaire given to the respondents.

Item Priming

We coded this binary category as true if we judged that earlier items on the survey might affect the responses to later items by making certain aspects of the research phenomena more salient to the respondents. For example, in Gattiker and Goodhue's [2005] study of ERP implementation the first item on the survey is "The information from the ERP system has numerous accuracy problems that make it difficult for employees to do their jobs." We coded this study as true for the item priming category because we judged that this first item may cause negative aspects of the ERP implementation to be more salient in respondents' minds as they complete the rest of the items, potentially biasing their responses. Five of 122 papers (4.1 percent) where coded to belong in this subcategory. The easiest way to avoid this bias is to place such items at the end of the questionnaire.

Item Embeddedness

We coded this binary category as true if we judged that the survey contained neutral items embedded in the context of positively or negatively worded items, and we believed that the neutral items could take on the evaluative properties of the positively or negatively worded items. We found that most papers tend to use neutral wording in their questionnaires and among the few papers that do use strongly positive or negative wording, no neutral wording items were intermixed. Therefore we did not code any of the papers as true for this category and concluded that it is not a major source of bias in IS research.

Context-Induced Moods

This binary category was coded as true if we judged that earlier items on a survey would likely induce a certain mood in the respondent that would affect their responses to later items. At first glance this category looks similar to the item priming category. The distinction between them is that this category describes an effect on the respondent's affective state, whereas the item priming category describes an effect on the respondent's cognitive state. For example, an earlier item that assessed the frequency of overtime could induce a mood of stress or unhappiness which could affect subsequent items that measure work satisfaction. Ten papers of 122 (8.2 percent) were coded as true in this subcategory. One way to avoid this bias is to separate items that can influence each other as much as possible. Items that have strong potential to affect the respondent's affective state should also be placed at or near the end of the survey questionnaire.

Grouping of Items or Constructs

This binary category was coded as true if items from the same construct were grouped together. The majority of the studies in our sample (118 of 122, 96.7 percent) were found to group related items together. Only four studies (3.3 percent) intermix items from different constructs in their surveys to prevent bias that may result from inflated correlations among grouped items.

Scale Availability

If the scales for all the items used in the study were included in the paper we coded this binary category as true. Ninety-seven of 128 (75.8 percent) papers in our sample provide scales for all items used in their studies.

Common Scale Formats

If the scales for all the items used the same format we coded this as true. Of the 97 papers that provide scales for all items, 67 studies (69.1 percent) adopt the same scale format for all the items.

Scale Length

This binary category was coded as true if the ratio of number of items per construct between any two constructs was 2 or greater. A long scale length may extend the time taken to complete that section of the survey and make items less accessible in short-term memory. We found that 31 studies (32.0 percent) used different scale length in their questionnaires.

Anchor Availability

If the anchors for the scales of all the items used in the study were included in the paper we coded this binary category as true. Eighty-two out of 128 papers (64.1 percent) in our sample provide the anchors for the scales of all items used in their studies.

Common Scale Anchors

If the anchors for the scales of all the items used in the study were the same we coded this binary category as true. Of the eighty-two papers that included scale anchors, 41 papers (50 percent) belong to this subcategory.

The use of common scale anchors can result in a high correlation among items due to format similarity and we recommend that variation should be introduced in the scale anchors to minimize such biases.

CATEGORY 4: METHODS USED TO ADDRESS POTENTIAL BIASES

This assessment examines whether any actions were taken by the studies in our dataset to address the potential for method bias, including common source bias and any of the method biases discussed above. It has four subcategories.

No Discussion of Method Bias

We coded this as true if there was no discussion in the paper concerning method bias. Eighty of 128 papers (62.5 percent) in our sample fail to mention the possibility of any method bias in their studies.

MRA and MRB

As discussed above, 22 papers (17.2 percent) collected data from multiple sources for either all variables or a subset of the variables, as a major approach to avoiding common source bias.

Method Bias Discussed but No Action Taken

We coded this as true if the article mentioned common source bias as a potential problem, but did not take any steps to prevent or detect it. Twelve papers (9.4 percent) acknowledged that method bias may exist in their studies but took no further actions to detect or reduce the potential bias. The other 23 papers (18.0 percent) use either statistical or nonstatistical approaches (or both) to minimize the confounding effect of potential method bias. These approaches are discussed in more detail below.

Nonstatistical Techniques Used to Control Method Bias

We coded this binary category as true if the authors reported using any non-statistical technique to control method bias. Some nonstatistical techniques for controlling method bias are reflected in other coding categories, such as the use of multiple respondents for the independent and

dependent variables, and the introduction of a time delay between the measurement of the independent and dependent variables. This category was included to record any other non-statistical technique used by researchers to control method bias.

An example of such a technique is used by Jarvenpaa and Staples [2001] in their study of the effects of several different factors on perceptions of organizational ownership of information and expertise. They measured the dependent variable, views of organizational ownership of information and knowledge, using a contrastive vignette technique [Burstin et al. 1980] in which the survey items measuring these views assess the respondent's attitudes toward what is presented in a short vignette. The authors argue that this technique may help prevent the effect of social desirability bias. We believe that it also may help to prevent common method bias caused by respondents' implicit theories because it introduces cognitive distance between the measurement of the independent and dependent variables. That is, since the dependent variable measures are explicitly attached to a vignette the respondents might not make a close connection between the dependent variable measures and the independent variable measures. A similar technique is used by Ryan et al. [2002] in their study of whether or not social subsystem costs and benefits are considered in IT investment decisions.

Another technique is exemplified by Bhattacherjee and Premkumar [2004], who used two longitudinal studies in different contexts to test their hypothesized model.

A total of thirteen papers (10.9 percent) were found to use such non-statistical methods to address potential method bias.

Statistical Techniques Used to Test for or Control Method Bias

We coded this binary category as true if the authors reported using any statistical technique to test for the presence of method bias or control method bias. We found that 11 studies (10.2 percent) used various statistical techniques to address potential method bias:

For example, Harman's single-factor test was performed by Devaraj et al. [2002] and Son et al. [2005]. In this test, all the variables are subjected to an exploratory factor analysis. The unrotated factor solution is examined to check for the presence of a single factor that accounts for all the covariance between variables or a general factor that explains a majority of the covariance among the variables [Podsakoff et al. 2003].

Another interesting technique was utilized by Keil et al. [2000] in their study of factors contributing to software project escalation. Their study used a single respondent to select either an escalated or nonescalated project and then rate behavioral factors associated with the project. They admit that there is a possibility that the single respondent imputed the behavioral factors based on whether the project was escalated or not, which would be an example of our "implicit theories" category, and indeed this paper was coded as true for that category. However they performed a one-way ANOVA between the escalation-related variables for the escalated and nonescalated projects, and found several variables for which there was no significant difference in means between the escalated and nonescalated projects, despite the fact that all the variables had been selected because they are expected predictors of project escalation. The authors argued that the selectivity in the factors associated with escalated projects indicates that the threat of method bias is minimal.

ABOUT THE AUTHORS

William R. King holds the title of university professor in the Katz Graduate School of Business at the University of Pittsburgh. He has published more than 300 papers and 15 books in the areas of Information Systems, Management Science, and Strategic Planning. He has served as founding president of the Association for Information Systems (AIS), President of TIMS (now INFORMS) and Editor-in-Chief of the MIS Quarterly. Recently, he was given a Leo Lifetime Exceptional Achievement Award by AIS.

Charles Z. Liu is a Ph.D. candidate in Information Systems at the Katz Graduate School of Business, University of Pittsburgh. He received the B.A. degree from the Department of Economics at Xiamen University, China and the M.A. degree from the Department of Economics at Tulane University. His current research interests include technology adoption in digital goods markets, the economics of networks and strategic issues in electronic markets. His work has been presented at several conferences such as *ICIS*, *WISE* and *INFORMS*. As a Ph.D. student, Mr. Liu won the 2006 Katz Student Research Grant and was a Doctoral Consortium Fellow at *ICIS* 2006.

Mark H. Haney is a Ph.D. candidate in Management Information Systems at the Katz Graduate School of Business, University of Pittsburgh. He has an MBA (Finance) from Case Western Reserve University, an M.A. in Chinese language and literature from Ohio State University, and a B.A. in economics and Chinese from Ohio State University. His research explores the management of the information systems development process, global information systems, and knowledge management. He has presented at the *IFIP 8.6* conference and published in the *Journal of Strategic Information Systems* and *OMEGA*.

Jun He is an assistant professor of MIS at the University of Michigan-Dearborn. He has an MBA from Tsinghua University and a PhD degree from the University of Pittsburgh. His research interests include systems design and development, knowledge management, and methodological issues. He has presented a number of papers at meetings of the Association for Computing Machinery (ACM) and the Americas' Conference on Information Systems (AMCIS), published in *Communications of the Association for Information Systems*, *Information & Management*, and in a book of *Current Topics in Management*.