MIS Quarterly

RESEARCH ESSAY

ESTIMATING THE EFFECT OF COMMON METHOD VARIANCE: THE METHOD–METHOD PAIR TECHNIQUE WITH AN ILLUSTRATION FROM TAM RESEARCH

By: **Rajeev Sharma**
Department of Information Systems
The University of Melbourne
Victoria 3010
AUSTRALIA
rajeevs@unimelb.edu.au

Philip Yetton
Australian School of Business
The University of New South Wales
Sydney NSW 2052
AUSTRALIA
p.yetton@unsw.edu.au

Jeff Crawford
Collins College of Business
The University of Tulsa
800 S. Tucker Drive
Tulsa, OK 74104
U.S.A.
jeff-crawford@utulsa.edu

Appendix A

Techniques to Control for Common Method Variance in Mono-Method Research Designs: A Brief Overview

Researchers have developed a number of statistical techniques to control for the effect of CMV in mono-method research designs.[1] Two techniques that have been frequently employed in IS research are the Harman single-factor test and the marker variable technique (see, for example, Jarvenpaa and Majchrzak 2008; Pavlou et al. 2007).

The Harman single-factor test requires loading all the measures in a study into an exploratory factor analysis, with the assumption that the presence of CMV is indicated by the emergence of either a single factor or a general factor accounting for the majority of covariance among measures (Podsakoff et al. 2003, p. 889). Podsakoff et al. characterize the Harman single-factor test as a diagnostic technique that "actually does nothing to statistically control for (or partial out) method effects" (p. 889). Further, they argue that the emergence of multiple factors does not indicate the absence of CMV and recommend against the use of this test.

[1]See Podsakoff et al. (2003) for an extensive review of techniques available to control for the effects of CMV.

In contrast with the Harman single-factor test, the marker variable technique (Lindell and Whitney 2001) attempts to control for CMV by including "a measure of the assumed source of method variance as a covariate in the statistical analysis" (Podsakoff et al. 2003, p. 889). Application of the marker variable technique requires the inclusion in the study of a variable that is theoretically unrelated to at least one of the focal variables. The correlation observed between the marker variable and the theoretically unrelated variable is interpreted as an estimate of CMV (Lindell and Whitney 2001). This is also assumed to be the extent of CMV contaminating every correlation in the study. Consequently, partialling out the correlation of the marker variable results in correlation values that are not contaminated by CMV (Lindell and Whitney 2001). The technique can also be employed in a *post hoc* manner by assuming the lowest correlation in a correlation matrix to be the magnitude of CMV and partialling it out of the analysis (Lindell and Whitney 2001). Malhotra et al. (2006) employed the marker variable technique in a *post hoc* manner on a sample of IS studies and concluded that CMV does not pose a threat to published findings within the IS domain.

Podsakoff et al. identify a number of conceptual and empirical problems with the marker variable technique. In particular, they argue that it ignores some of the most powerful sources of method biases, including, for example, those arising from implicit theories (p. 893). Further, it assumes that the method factor has a constant correlation with all manifest variables, an assumption that is likely to have a substantive effect on findings resulting from the application of this technique (Podsakoff et al. 2003, p. 893; Richardson et al. 2009).

Importantly, as both Straub and Burton-Jones (2007) and Sharma et al. (2007) argue, the marker variable technique does not control for common rater problems, particularly those including implicit theories and the consistency motif. These are likely to be the important underlying issues for TAM studies, a limitation acknowledged by Malhotra et al. The conceptual overlap between PU and U virtually ensures a high spurious component in observed correlations when self-report measures are employed to measure both variables (Straub and Burton-Jones 2007). These sources of CMV are not captured by a marker such as age of respondent because these sources of CMV are not simply a matter of similar instrument format.

The analysis in the "Discussion" section of the paper shows that, when the PU-U relationship is estimated from measures that are highly susceptible to CMV, CMV may explain more of the observed relationship than does the true relationship. Straub and Burton-Jones argue that the combination of a common rater, providing self-report data, collected on perceptually anchored scales to investigate an intentions-based theory is highly problematic and that such research designs expose the findings to the classic mono-method threat identified, for example, by Cook and Campbell (1979).

In addition, the marker variable technique is not well specified in practice (Sharma et al. 2007). Lindell and Whitney (2001) identify multiple criteria for selecting the marker variable. However, they neither provide a priori theoretical justification for selecting one over the other, nor provide any practical guidelines for doing so. Consequently, the validity of findings arising from the application of this technique is questionable[2] (Sharma et al. 2007).

Finally, Richardson et al. (2009) investigate the validity of the marker variable technique, simulating the values of the marker variable correlation under a number of different conditions. They report that both the marker variable technique and its CFA-based extensions exhibit low accuracy rates. They conclude that the application of these techniques poses a substantial risk of false conclusions and that they are inappropriate in most circumstances.

References

Cook, T. D., and Campbell, D. T. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Studies*, Chicago: Rand McNally.

Jarvenpaa, S. L., and Majchrzak, A. 2008. "Knowledge Collaboration among Professionals Protecting National Security: Role of Transactive Memories in Ego-Centered Knowledge Networks," *Organizational Science* (19:2), pp. 260-276.

Lindell, M. K., and Whitney, D. J. 2001. "Accounting for Common Method Variance in Cross-Sectional Research Designs," *Journal of Applied Psychology* (86:1), pp. 114-121.

Malhotra, N. K., Kim, S. S., and Patil, A. 2006. "Common Method Variance in IS Research: A Comparison of Alternative Approaches and a Reanalysis of Past Research," *Management Science* (52:12), pp. 1865-1883.

Pavlou, P. A., Liang, H., and Xue, Y. 2007. "Understanding and Mitigating Uncertainty in Online Exchange Relationships: A Principal-Agent Perspective," *MIS Quarterly* (31:1), pp. 105-136.

[2]Other partial correlation techniques have also been proposed in the literature. For instance, one technique attempts to control for the effect of CMV by including a variable in the survey instrument to capture social desirability and then partialing out the effect of social desirability from the analysis (Podsakoff et al. 2003). These techniques are conceptually similar to the marker variable technique and are subject to the same validity threats as the marker variable technique.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. 2003. "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology* (88:5), pp. 879-903.

Richardson, H. A., Simmering, M. J., and Sturman, M. C. 2009. "A Tale of Three Perspectives: Examining Post Hoc Statistical Techniques for Detection and Correction of Common Method Variance," *Organizational Research Methods* (doi: 10.1177/1094428109332834), March 11, pp. 1-39.

Sharma, R., Yetton, P., and Crawford, J. 2007. "Common Methods Bias: Reports of its Demise Are Highly Exaggerated," in *Proceedings of the 28th International Conference on Information Systems*, Montreal, December 9-12.

Straub, D. W., and Burton-Jones, A. 2007. "Veni, Vidi, Vici: Breaking the TAM Logjam," *Journal of the Association for Information Systems* (8:4), pp. 223-229.

Appendix B

Sample and Analysis of Validity Threats

Table B1. Studies Included in the Meta-Analysis and Study Data

Authors & Study Details* *Study details are provided where multiple data points have been included from a study.	**Sample Size**	**PU-Usage Correlation**	**Susceptibility of Method–Method Pair to CMV**
1. Adams, D. A., Nelson, R. R., and Todd, P. A. (1992), *MIS Quarterly* (16:2) • Study 1 – E-mail	116	0.35	Low
2. Adams, D. A., Nelson, R. R., and Todd, P. A. (1992), *MIS Quarterly* (16:2) • Study 1 – Voice Mail	68	0.45	Low
3. Adams, D. A., Nelson, R. R., and Todd, P. A. (1992), *MIS Quarterly* (16:2) • Study 2 – Harvard Graphics – Usage (Hours)	73	0.25	Low
4. Adams, D. A., Nelson, R. R., and Todd, P. A. (1992), *MIS Quarterly* (16:2) • Study 2 – Lotus 1-2-3 – Usage (Hours)	73	0.34	Low
5. Adams, D. A., Nelson, R. R., and Todd, P. A. (1992), *MIS Quarterly* (16:2) • Study 2 – Word Perfect – Usage (Hours)	73	0.14	Low
6. Adams, D. A., Nelson, R. R., and Todd, P. A. (1992), *MIS Quarterly* (16:2) • Study 2 – Harvard Graphics – Usage (Frequency)	73	0.42	High
7. Adams, D. A., Nelson, R. R., and Todd, P. A. (1992), *MIS Quarterly* (16:2) • Study 2 – Lotus 1-2-3 – Usage (Frequency)	73	0.48	High
8. Adams, D. A., Nelson, R. R., and Todd, P. A. (1992), *MIS Quarterly* (16:2) • Study 2 – Word Perfect – Usage (Frequency)	73	0.40	High
9. Adekoya, A. A. (1993), Unpublished Doctoral Dissertation, Syracuse University	102	0.13	High
10. Al-Khaldi, M. A., and Wallace, R. S. O. (1999), *Information & Management* (36:4)	151	0.22	Medium
11. Barki, H., and Huff, S. (1985), *Information & Management* (9:5)	42	0.39	Low
12. Baroudi, J. J., Olson, M. H., and Ives, B. (1986), *Communications of the ACM* (29:3)	200	0.28	High
13. Bhattacharjee, A., and Premkumar, G. (2004), *MIS Quarterly* (28:2) • CBT Study – Usefulness at time 2; Satisfaction at time 2	175	0.68	Very high
14. Bhattacharjee, A., and Premkumar, G. (2004), *MIS Quarterly* (28:2) • CBT Study – Usefulness at time 3; Satisfaction at time 2	172	0.64	Very high
15. Bhattacharjee, A., and Premkumar, G. (2004), *MIS Quarterly* (28:2) • RAD Study – Usefulness at time 2; Satisfaction at time 2	77	0.55	Very high
16. Bhattacharjee, A., and Premkumar, G. (2004), *MIS Quarterly* (28:2) • CBT Study – Usefulness at time 1; Satisfaction at time 2	175	0.33	Very high
17. Bhattacharjee, A., and Premkumar, G. (2004), *MIS Quarterly* (28:2) • CBT Study – Usefulness at time 1; Satisfaction at time 3	172	0.59	Very high
18. Bhattacharjee, A., and Premkumar, G. (2004), *MIS Quarterly* (28:2) • RAD Study – Usefulness at time 1; Satisfaction at time 2 **Note:** This data point was excluded from the final analysis because it was determined to be an outlier (see the "Sample" section of the paper for additional details).	88	-0.32	Very high

Table B1. Studies Included in the Meta-Analysis and Study Data (Continued)

Authors & Study Details* *Study details are provided where multiple data points have been included from a study.	**Sample Size**	**PU-Usage Correlation**	**Susceptibility of Method–Method Pair to CMV**
19. Chen, L-D. (2000), Unpublished Doctoral Dissertation, University of Memphis	46	0.36	High
20. Dahmer, B. L. (1994), Unpublished Doctoral Dissertation, Vanderbilt University	344	0.11	Very low
21. Davis, F. D. (1989), *MIS Quarterly* (13:3) • Study 1 – E-mail	109	0.56	High
22. Davis, F. D. (1989), *MIS Quarterly* (13:3) • Study 1 – XEDIT	75	0.68	High
23. Ferguson, C. (1997), *Accounting & Finance* (37:1)	157	0.22	Low
24. Floyd, S. W. (1986), Unpublished Doctoral Dissertation, University of Colorado	73	0.41	Very low
25. Glorfeld, K. D. (1994), Unpublished Doctoral Dissertation, University of Arkansas	106	0.31	Low
26. Green, C. W. (1998), *Small Group Research* (29:1)	125	0.52	High
27. Guimaraes, T., Igbaria, M., and Lu, M. (1992), *Decision Sciences* (23:2)	118	0.48	Very high
28. Guimaraes, T., Yoon, Y., and Clevenson, A. (1996), *Information & Management* (30:3)	114	0.22	Very high
29. Guthrie, C. (2001), Proceedings of AMCIS, Boston • Study 1 – Administrative services	49	0.19	High
30. Guthrie, C. (2001), Proceedings of AMCIS, Boston • Study 2 – Collaborative services	49	0.34	High
31. Guthrie, C. (2001), Proceedings of AMCIS, Boston • Study 3 – Class preparation services	49	0.28	High
32. Habelow, E. M. (2000), Unpublished Doctoral Dissertation, Temple University	106	0.58	High
33. Heilman, G. E. (1997), Unpublished Doctoral Dissertation, University of Arkansas	140	0.57	High
34. Horton, R. P., Buck, T., Waterson, P. E., and Clegg, C. W. (2001), *Journal of Information Technology* (16:4) • Study 1 – Self-reported usage	386	0.34	Low
35. Horton, R. P., Buck, T., Waterson, P. E., and Clegg, C. W. (2001), *Journal of Information Technology* (16:4) • Study 2 – Measured usage (time 3)	65	0.17	Very low
36. Horton, R. P., Buck, T., Waterson, P. E., and Clegg, C. W. (2001), *Journal of Information Technology* (16:4) • Study 2 – Self-reported usage	65	0.23	Low
37. Howard, G. S., and Mendelow, A. L. (1991), *Decision Sciences* (22:2)	422	0.36	Low
38. Igbaria, M. (1993), *OMEGA* (21:1)	519	0.24	Medium
39. Igbaria, M., Guimaraes, T., and Davis, G. B. (1995), *Journal of Management Information Systems* (11:4) • Sample S1 – Perceived usage	107	0.40	High
40. Igbaria, M., Guimaraes, T., and Davis, G. B. (1995), *Journal of Management Information Systems* (11:4) • Sample S2 – Perceived usage	105	0.39	High
41. Igbaria, M., and Iivari, J. (1995), *OMEGA* (23:6)	450	0.45	High
42. Igbaria, M., Parasuraman, S., and Baroudi, J. J. (1996), *Journal of Management Information Systems* (13:1)	471	0.40	High
43. Igbaria, M., Zinatelli, M., Cragg, P., and Cavaye, A. L. M. (1997), *MIS Quarterly* (21:3)	358	0.42	Medium
44. Igbaria, M., and Maansaari, J. (1997), Proceedings of ICIS, Atlanta	44	0.36	High
45. Karpur, S. R. (1992), Unpublished Doctoral Dissertation, Temple University	289	0.04	Low
46. Keil, M., Beranek, P. M., and Konsynski, B. R. (1995), *Decision Support Systems* (13:1) • Old CONFIG	129	0.43	Low
47. Keil, M., Beranek, P. M., and Konsynski, B. R. (1995), *Decision Support Systems* (13:1) • New CONFIG	177	0.42	Low

Table B1. Studies Included in the Meta-Analysis and Study Data (Continued)

Authors & Study Details* *Study details are provided where multiple data points have been included from a study.	**Sample Size**	**PU-Usage Correlation**	**Susceptibility of Method–Method Pair to CMV**
48. Kim, I. (1996), Unpublished Doctoral Dissertation, University of Nebraska–Lincoln • Object-oriented analysis and design	109	0.41	High
49. Kim, I. (1996), Unpublished Doctoral Dissertation, University of Nebraska–Lincoln • Object-oriented programming	109	0.40	High
50. King, M. K. (1997), Unpublished Doctoral Dissertation, Harvard University • Pharmco B	44	0.44	High
51. King, M. K. (1997), Unpublished Doctoral Dissertation, Harvard University • Pharmco A	50	0.55	High
52. Limayem, M., Cheung, C. M. K., and Chan, G. W. (2003), Proceedings of ICIS, Seattle	371	0.61	Very high
53. Maish, A. M. (1979), *MIS Quarterly* (3:1)	47	0.27	High
54. Manson, D. P. (1998), Unpublished Doctoral Dissertation, The Claremont Graduate University • AUDIT Company	38	0.48	Low
55. Manson, D. P. (1998), Unpublished Doctoral Dissertation, The Claremont Graduate University • SCIENCE Company	317	0.24	Low
56. Manson, D. P. (1998), Unpublished Doctoral Dissertation, The Claremont Graduate University • INSURANCE Company	129	0.07	Low
57. Parasarathy, M., and Bhattacharjee, A. (1998), *Information Systems Research* (9:4)	145	0.44	High
58. Rai, A., Lang, S. S., and Welker, R. B. (2002), *Information Systems Research* (13:1)	274	0.66	Very high
59. Roberts, P., and Henderson, R. (2000), *Interacting with Computers* (12:5)	108	0.33	Medium
60. Romano, C. A. (1993), Unpublished Doctoral Dissertation, University of Maryland at Baltimore	229	0.66	Very high
61. Russo, N. L. (1993), Unpublished Doctoral Dissertation, Georgia State University	30	0.34	Low
62. Shih, H-P. (2004), *Information & Management* (41:6)	203	0.59	Very high
63. Srinivasan, A. (1983), Unpublished Doctoral Dissertation, University of Pittsburgh	28	0.39	Very high
64. Swanson, E. B. (1987), *Decision Sciences* (18:1)	171	0.20	High
65. Szajna, B. (1996), *Management Science* (42:1) • Post-implementation usefulness, actual usage	61	0.06	Very low
66. Szajna, B. (1996), *Management Science* (42:1) • Post-implementation usefulness, self-report usage	61	0.38	High
67. Szajna, B. (1996), *Management Science* (42:1) • Pre-implementation usefulness, self-report usage	61	0.09	High
68. Szajna, B. (1996), *Management Science* (42:1) • Pre-implementation usefulness, actual usage	61	0.12	Very low
69. Taylor, S., and Todd, P. A. (1995), *Information Systems Research* (6:2)	786	0.16	Very low
70. Thompson, R. L., Higgins, C. A., and Howell, J. M. (1991), *MIS Quarterly* (15:1)	212	0.38	Medium
71. Thompson, R. L., Higgins, C. A., and Howell, J. M. (1994), *Journal of Management Information Systems* (11:1)	219	0.51	Medium
72. Whang, J. (1992), Unpublished Doctoral Dissertation, The University of Nebraska, Lincoln	106	0.44	High
73. Wober, K., and Gretzel, U. (2000), *Journal of Travel Research* (39:2)	77	0.15	Very low
74. Yang, H-D., and Yoo, Y. (2004), *Decision Support Systems* (38:1)	211	0.31	High
75. Yoon, Y., Guimaraes, T., and O'Neal, Q. (1995), *MIS Quarterly* (19:1)	69	0.31	Very high
76. Zmud, R. W. (1984), *Management Science* (30:6)	47	0.56	High

Validity Threat from Inclusion of Multiple Studies from a Publication

Multiple data points reported in a publication should be included only when there is justification for treating them as independent (Lipsey and Wilson 2001). Otherwise, the assumption of independence underpinning regression and other statistical analyses is violated and may result in spurious results. Multiple data points in a study can be considered independent for a variety of reasons. If the multiple data points relate to independent samples of respondents, they are independent. Multiple data points relating to the same set of respondents but captured at different points in time can also be considered independent. Multiple data points relating to the sample of respondents but based on different stimuli can also be considered independent. Where independence cannot be established, statistical independence can potentially be violated when the multiple effect sizes refer to the same conceptual relationship. The sample employed in this illustration includes 11 publications that contribute multiple data points. We describe some of these cases below, reasons for considering them independent, and also analyze the potential validity threat posed by their inclusion.

Adams et al. (1992) contribute eight data sets gathered from two independent samples of respondents. The first sample contributes two data points, one each for two different technologies. These two data points are independent as they are based on different stimuli. The second sample contributes six data points. Respondents present their responses to three different technologies and provide two different measures of usage for each technology. Responses to the three different technologies are independent as they pertain to different stimuli. The independence of the two measures of usage is not as clear cut. However, the high level of variability in the correlations for the two different measures of usage suggests that the assumption of independence is not violated.

Bhattacharjee and Premkumar (2004) contribute six data points based on two independent samples of respondents: two data points come from the RAD study while the other four come from the CBT study. The multiple data points within both independent samples of respondents were collected at different points in time, ensuring independence of data points.

Davis (1989) contributes two data sets, based on two different technologies involving some overlap of respondents. These two data points are independent as they are based on different stimuli. Instead of including the pooled correlation from this study ($r = 0.63$) with a sample size of 184, we have included two data points ($r = 0.56$ and $r = 0.68$, with sample sizes of 109 and 75, respectively) from this publication to reflect the independence of the data points and the variability in correlations.

Guthrie (2001) contributes three data points based on responses to three different technologies, but with the same respondents. These data points are independent as they are based on different stimuli. The variability of the three correlation values is quite high ($r = 0.19$, 0.28, and 0.34), suggesting that the assumption of independence is not violated.

Horton et al. (2001) contribute three data points, based on two independent samples of respondents. The two data points from one sample of respondents are collected differently: one is system-captured while the other is self-report. The two usage measures are collected from independent sources, ensuring independence of the data points.

Manson (1998) contributes three data points, based on three independent samples drawn from three organizations. These data points are independent.

Overall, potential nonindependence of data points does not pose a substantial validity threat to the findings of this study. The independence of multiple data points from a single publication is clear cut in most cases and the number of publications where the independence of data points may be doubtful is small compared to the overall sample size. Further, since the objective of this illustration is to evaluate the effect of different operationalizations, it is not meaningful to aggregate effect sizes across operationalizations or to select a data point randomly from the multiple operationalizations. Lipsey and Wilson (2001, p. 78) argue that the most inclusive approach employed here "permits the fullest empirical examination of the relationship between the particular ways in which a construct is operationalized and the nature of findings from different studies." Consistent with the objective of this study, the most inclusive approach is employed.

The effect of the violation of the independence assumption by including multiple data points from a single publication is a question that can be resolved empirically. An empirical examination of the effect of including multiple data points from a publication finds that this does not pose a threat to the validity of findings reported in Tables 5 and 6. We repeated the analysis excluding every data point that came from a publication that contributed more than one correlation, leaving a sample size of 39 data points. Both H1 and H2 are supported in this subsample. Further, the values of parameters in this subsample are very similar to those for the full data set reported in Tables 5 and 6. Within this smaller subsample, CMV[MMP-I] is significant ($p < 0.05$) and explains 61.8 percent of the variance, very similar to that explained in the full data set (56.09 percent, Table 5). The means for the categories of usage (Table 6) are also virtually identical (0.16 vs. 0.16 for Very Low; 0.25 vs. 0.29 for Low; 0.41 vs. 0.42 for High; and 0.60 vs. 0.59 for Very High). The findings of the study are not subject to a validity threat arising from the inclusion of multiple data points from a single publication.

Reliabilities of Measures

Table B2 reports the summary statistics of the reliabilities of perceived usefulness and use reported in the primary studies. Reliability for perceived usefulness is reported in 61 of the 75 data sets. The mean reliability of perceived usefulness was used to correct the reported correlation in studies not reporting reliability values. However, reliability for usage is available for only 26 of the 75 data sets. Accordingly, the findings reported in Tables 3, 5 and 6 do not correct for the reliability of usage.

Analysis of Validity Threats Arising from Nonavailability of Reliability Usage

Since reliability values for the system-captured and behavioral continuous categories are not available, there still remains a potential validity threat to the findings reported in Tables 5 and 6. For this threat to be material, the mean reliability for system-captured usage measures would need to be 0.27 for the average PU-U correlation for this category to be equal to that of the perceptually anchored category (r = 0.59, Table 6). Similarly, the mean reliability value for behavioral continuous measures (r = 0.29, Table 6) would need to be 0.49 for the average PU-U correlation for this category to be equal that of the perceptually anchored category (r = 0.59). While these values appear implausible,[3] this is an important empirical question that should be addressed in future research.

To more formally address this validity threat, Tables B3 and B4 report the results of the random effects ANOVA (H1) and planned contrast analysis (H2) for three cases. The first case is for the subsample of 26 data points for which reliability values for both perceived usefulness and usage are available. The second case extends the first and includes the subsample of 50 data points in the three categories (behaviorally anchored, mixed, and perceptually anchored) that report reliability of usage. The mean reliability of usage for each category is employed to correct the observed correlation for these data points. The third case extends the second and includes the complete set of 75 data points. An assumed low reliability of 0.75 is employed to correct the observed correlations in the system-captured and behavioral continuous categories. The dependent variable in each case is the PU-U correlation corrected for the reliabilities of both PU and U. Each data point is weighted by its inverse variance weight[4] (Hunter and Schmidt 2004).

Table B2. Summary of Reliabilities of Perceived Usefulness and Usage Measures

Measure	Perceived Usefulness	Usage
Total number of data sets	75	75
Number of data sets employing multi-item measures and reporting reliability (system-captured, behavioral continuous, behaviorally anchored, mixed, perceptually anchored)	61 (3, 15, 6, 26, 11)	26 (0, 0, 5, 12, 9)
Number of data sets employing one-item measures or with unavailable information	14	49
Maximum reliability	0.99	0.97
Minimum reliability	0.68	0.51
Mean reliability (system-captured, behavioral continuous, behaviorally anchored, mixed, perceptually anchored)	0.91 (0.94, 0.93, 0.82, 0.91, 0.90)	0.84 (NA, NA, 0.74, 0.83, 0.93)
Number of data sets reporting reliability below 0.80	4	6

Note: NA = Not Available

[3]This analysis assumes that the reliability of perceptually anchored measures is 1.00. When more realistic values for the reliability of perceptually anchored measures are employed, the reliability of system-captured and behavioral continuous measures will need to be even lower.

[4]The inverse variance weight employed in the findings reported in Tables B3 and B4 is calculated as per the formula recommended by Hunter and Schmidt (2004, pp. 122-132). The analysis was repeated employing the formula for inverse variance weights recommended by Lipsey and Wilson (2001). The results are virtually identical in both cases and are available on request from the authors.

Table B3. Analysis of Validity Threat Arising from Nonavailability of Reliabilities of Usage: Results of Random Effects ANOVA

Factor	Base Case (Table 5): Variance Explained (Eta Squared) (p value, two-tail)	Case 1: Variance Explained (Eta Squared) (p value, two-tail)	Case 2: Variance Explained (Eta Squared) (p value, two-tail)	Case 3: Variance Explained (Eta Squared) (p value, two-tail)
Publication Type	0.65% (p = 0.276)	3.72% (p = 0.300)	1.83% (p = 0.230)	0.56% (p = 0.348)
Respondent Type	2.71% (p = 0.029)	0.77% (p = 0.635)	6.50% (p = 0.027)	3.09% (p = 0.027)
PU Operationalization Type	4.27% (p = 0.007)	5.09% (p = 0.228)	6.97% (p = 0.022)	3.34% (p = 0.024)
CMV[MMP-I]	56.09% (p = 0.000)	24.68% (p = 0.041)	30.35% (p = 0.000)	51.28% (p = 0.000)
Error	36.28%	65.74%	54.35%	41.73%
Number of data points	75	26	50	75

Table B4. Analysis of Validity Threat Arising from Nonavailability of Reliabilities of Usage: Results of Planned Contrast Analysis

Category of Usage Measure	Base Case (Table 6: Contrast (p value, two-tail)	Case 1: Contrast (p value, two-tail)	Case 2: Contrast (p value, two-tail)	Case 3: Contrast (p value, two-tail)
Behavioral Continuous (CMV[MMP-I] = Low) vs. System-Captured (CMV[MMP- I] = Very Low)	0.29 vs. 0.16 ($p \leq 0.02$)	NA	NA	0.34 vs. 0.18 ($p \leq 0.05$)
Behaviorally anchored (CMV[MMP-I] = High) vs. Behavioral Continuous (CMV[MMP-I] = Low)	0.42 vs. 0.29 ($p \leq 0.01$)	NA	NA	0.46 vs. 0.34 ($p \leq 0.01$)
Perceptually anchored (CMV[MMP-I] = Very High) vs. Behaviorally Anchored (CMV[MMP-I] = High)	0.59 vs. 0.42 ($p \leq 0.01$)	0.59 vs. 0.44 ($p \leq 0.01$)	0.61 vs. 0.46 ($p \leq 0.01$)	0.61 vs. 0.46 ($p \leq 0.01$)
Number of data points†	69	21	44	69

†This analysis excludes the six data points in the mixed category.

An inspection of Table B3 shows that the magnitude of between-studies variance explained by CMV [MMP-I] is significant in each of the three cases. Support for H1 is not subject to a validity threat arising from excluding reliability of usage in the main analysis. Similarly, an inspection of Table B4 shows that the contrasts are significant in each of the three cases. Support for H2 is also not subject to a validity threat arising from excluding reliability of usage in the main analysis (Tables 5 and 6).

To further investigate this validity threat, a weighted least squares regression for Case 3 in Tables B3 and B4 is presented in Table B5. An inspection of Table B5 finds that the regression coefficient for CMV [MMP-I] is significant ($\beta = 0.74$, $p < 0.05$, $n = 75$). H1 and H2 are supported even when reliability of U is included in the analysis. A similar analysis for Case 1 above (including only the 26 data points for which reliabilities of both PU and U are reported) also finds that the regression coefficient for CMV [MMP-I] is significant ($\beta = 0.60$, $p < 0.05$, $n = 26$). H1 and H2 are supported for the subset of studies for which reliability values are available for both PU and U.

The analyses reported in Tables B3, B4, and B5 support the conclusion that the findings reported in Tables 5 and 6 are not subject to a validity threat arising from the nonavailability of reliability values of usage.

Table B5. Analysis of Validitiy Threat Arising from Nonavailability of Reliabilities of Usage: Results of Regression Analysis†

Variable	Standardized Regression Coefficient (Beta)	Standard Error	t	Significance
Intercept	$b_0 = 0.04$	0.06	0.70	p = 0.49
Publication Type	0.07	0.03	0.81	p = 0.42
Respondent Type	0.18	0.03	2.09	p = 0.04
PU Operationalization Type	0.19	0.04	2.22	p = 0.03
CMV[MMP-I]	0.74	0.02	9.11	p = 0.00

†See Appendix C for details of regression analysis.

References

Hunter, J. E., and Schmidt, F. 2004. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (2nd ed.), Newbury Park, CA: Sage Publications.

Lipsey, M. W., and Wilson, D. B. 2001. *Practical Meta-Analysis*, Thousand Oaks, CA: Sage Publications.

Appendix C

Estimating the Magnitude of a Focal Correlation Controlling for CMV

The intercept in a regression analysis with CMV[MMP] as the independent variable, publication type, respondent type, and PU type as the control variables, and observed correlation as the dependent variable, can be interpreted as an estimate of the focal construct-construct correlation controlling for CMV:

$$r_{obsi} = r_{cc} + \hat{\beta} \times \text{CMV[MMP]}_i + \hat{\beta}_k \times (\text{Control Variable})_k + \text{error}$$

where

r_{obsi} is the observed correlation corrected for measurement errors in Study i

CMV[MMP]_i is the CMV[MMP] for Study i

Control variables are the variables hypothesized to influence the magnitude of the focal correlation (r) being examined. In this study, the control variables included are publication type, respondent type and PU type

r_{cc} is the intercept of the regression equation, which is interpreted as an estimate of the construct-construct correlation controlling for CMV

Each study in the regression is weighted by its inverse variance weight.

The CMV[MMP] scale is rank-ordered with zero corresponding to the extent of CMV with objectively measured variables. This allows non-parametric statistical tests to be conducted with the rank-ordered scale (Babakus et al. 1987; Conover and Inman 1981; Labovitz 1970). It also enables the interpretation of the intercept in a regression model as the magnitude of the focal correlation controlling for CMV (Labovitz 1970).

Table C1 presents the results of the above regression analysis (n = 75). The criterion variable in the regression is the PU-U correlation corrected for the reliability of PU. The inverse variance weight employed for each study is its sample size adjusted for the reliability of PU (Hunter and Schmidt 2004, p. 122-132). This is the data set employed in the findings reported in Tables 5 and 6. The intercept is 0.01 (ns). We conclude that the magnitude of the PU-U correlation controlling for CMV and other measurement artifacts is 0.01 (ns).

Table C1. Regression Analysis to Estimate of PU-Usage Correlation Controlling for CMV†				
Variable	Standardized Regression Coefficient (Beta)	Standard Error	t	Significance
Intercept	$b_0 = 0.01$	0.06	0.17	p = 0.86
Publication Type (0 = Dissertation & conference proceeding; 1 = Journal)	0.07	0.03	0.90	p = 0.37
Respondent Type (0 = Student; 1 = Nonstudent)	0.15	0.03	1.83	p = 0.07
PU Operationalization Type (0 = Non-Davis; 1 = Davis' PU instrument)	0.21	0.03	2.69	p = 0.01
CMV[MMP-I]	0.77	0.01	10.11	p = 0.00

†The results are very similar when inverse variance weights recommended by Lipsey and Wilson (2001) are employed ($\hat{\beta}$ for CMP[MMP-I] = 0.780, $p \leq 0.05$, and Intercept = -0.02, ns).

However, the intercept obtained needs to be corrected for various measurement artifacts, as discussed by Hunter and Schmidt (2004), Le et al. (2009) and Schmidt et al. (2003). These corrections are necessary before valid estimates can be obtained for the magnitude of a focal correlation controlling for CMV and other measurement errors. However, given the magnitude of the intercept obtained (r = 0.01, ns) it is unlikely that corrections for other measurement artifacts would raise the estimate of the intercept to any substantive magnitude in this case.

The magnitude of the intercept is sensitive to a number of assumptions underpinning the regression analysis. First, it assumes the CMV[MMP-I] scale to be an interval scale. Violation of this assumption could result in different estimates of the intercept. However, the errors resulting from treating ordinal scales as interval scales in a regression are usually small and valid conclusions can be drawn from regressions conducted with ordinal scales (Labovitz 1970). Regression analysis is usually robust to violations of assumptions regarding the measurement scales and the distribution of errors (Babakus et al. 1987).

Second, the estimate of the intercept is sensitive to the assumption that a linear relationship fits the data. The sensitivity of the estimate of the intercept to a violation of this assumption can be estimated by including quadratic and/or other terms in the regression analysis. A regression analysis including the quadratic term finds the value of the intercept to be 0.019, virtually identical to that reported in Table C1. However, these issues need to be investigated in future research.

Third, the estimate of the intercept is sensitive to two assumptions regarding system-captured measures. One is that correlations involving system-captured measures are independent of any CMV-based biases. Therefore, in the above illustration, the CMV influence on the observed correlation is assumed to be zero when usage is measured on a system-captured scale. The other assumption concerns the errors in system-captured measures. If the reliability of system-captured measures is substantially lower than of other measures, it could have a significant impact on estimates of the slope and intercept in the regression. (See Appendix B for an analysis of this issue.)

To analyze the above validity threats, Table C2 reports the results of a regression analysis for three cases. Case 1 is the base case reported in Table C1. The sample for Case 2 is the same as Case 1, except that it excludes the seven system-captured data points. This addresses the validity threat that system-captured measures are driving down the value of the intercept. The sample for Case 3 is the same as Case 1, except that the dependent variable is the PU-U correlation corrected for the reliabilities of both PU and U. Case 3 assumes a low reliability of 0.75 for system-captured and behavioral continuous measures of U and substitutes the mean value of reliability of U for other data points where reliability of U is not available. An inspection of the results finds that the magnitudes of the intercept in the three cases are r = 0.01 (ns), -0.05 (ns) and 0.04 (ns). The conclusions drawn in the "Discussion" section of the paper regarding the magnitude of the PU-U correlation are not sensitive to the above validity threats.

Inspecting Table C2 for Case 2, the standardized regression coefficient for CMV[MMP-I] is significant ($\hat{\beta} = 0.73$, $p \leq 0.01$). The hypotheses are supported even when system-captured measures are excluded from the analysis. Conclusions from a random effects ANOVA and contrast analysis are identical, and are available on request from the authors.

Table C2. Sensitivity Analysis for Estimates of the Intercept			
Factor	**Case 1 (Table C1): Standardized Regression Coefficient, Beta (p value, two-tail)**	**Case 2: Standardized Regression Coefficient, Beta (p value, two-tail)**	**Case 3: Standardized Regression Coefficient, Beta (p value, two-tail)**
Intercept	b_0 = 0.01 (p = ns)	b_0 = -0.05 (p = ns)	b_0 = 0.04 (p = ns)
Publication Type	0.07 (p = ns)	0.07 (p = ns)	0.07 (p = ns)
Respondent Type	0.15 ($p \leq 0.10$)	0.21 ($p \leq 0.05$)	0.18 ($p \leq 0.05$)
PU Operationalization Type	0.21 ($p \leq 0.01$)	0.26 ($p \leq 0.01$)	0.19 ($p \leq 0.05$)
CMV[MMP-I]	0.77 ($p \leq 0.01$)	0.73 ($p \leq 0.01$)	0.74 ($p \leq 0.01$)
Number of data points	75 (Observed correlation corrected for reliability of PU)	68 (Excludes system-captured measures)	75 (Observed correlation corrected for reliabilities of PU and U

References

Babakus, E., Ferguson, C. E., Jr., and Joreskog, K. G. 1987. "The Sensitivity of Confirmatory Maximum Likelihood Factor Analysis to Violations of Measurement Scale and Distributional Assumptions," *Journal of Marketing Research* (24:2), pp. 222-228.

Conover, W. J., and Inman, R. L. 1981. "Rank Transformations as a Bridge between Parametric and Nonparametric Statistics," *The American Statistician* (35:3), pp. 123-129.

Hunter, J. E., and Schmidt, F. 2004. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (2nd ed.), Newbury Park, CA: Sage Publications.

Labovitz, S. 1970. "The Assignment of Numbers to Rank Order Categories," *American Sociological Review* (35:2), pp. 515-524.

Le, H., Schmidt, F. L., and Putka, D. J. 2009. "The Multifaceted Nature of Measurement Artifacts and its Implications for Estimating Construct-Level Relationships," *Organizational Research Methods* (12:1), pp. 165-200.

Lipsey, M. W., and Wilson, D. B. 2001. *Practical Meta-Analysis*, Thousand Oaks, CA: Sage Publications.

Schmidt, F. L., Le, H., and Ilies, R. 2003. "Beyond Alpha: An Empirical Examination of the Effects of Different Sources of Measurement Error on Reliability Estimates for Measures of Individual Differences Constructs," *Psychological Methods* (8:2), pp. 206-224.

Appendix D

Comparison with Marker Variable Technique

The technique proposed here estimates the effect of CMV as a function of the methods employed to measure the independent and dependent variables. It assumes that increasing the CMV shared by method–method pairs increases the spurious correlation between measures. This assumption is consistent with the MTMM technique and descriptions of the effect of CMV on observed correlations (Doty and Glick 1998; Podsakoff et al. 2003). In contrast, the marker variable technique (Lindell and Whitney 2001; Malhotra et al. 2006) assumes that the effect of CMV is independent of the methods employed in the focal correlation being investigated.

The marker variable technique is underpinned by the major assumption that the method factor has a constant effect on all measured items (Lindell and Whitney 2001; Malhotra et al. 2006). Under that assumption, the lowest (or second lowest) correlation in the full correlation matrix reported in a study is an unbiased proxy for CMV. Note, this technique does not consider the methods employed to measure the focal variables when estimating the expected CMV effect on the focal correlation.

While a critical comparison of the two techniques is beyond the scope of this discussion,[5] we evaluate the validity of Lindell and Whitney's above claim, which is critical to the validity of the MVT. Following this claim, we hypothesize that

HD1: The CMV-corrected correlation between perceived usefulness and usage, obtained by employing the marker variable technique, is not a function of the method–method pair employed.

To test the above hypothesis, following Malhotra et al. (2006), we calculated the CMV-adjusted correlation between perceived usefulness and usage for each study included in Sharma et al. (2007), using the second lowest correlation in the complete correlation matrix reported in the study as the marker variable. The mean values of the CMV-corrected correlations for the four levels of CMV[MMP-I] are 0.08, 0.12, 0.31, and 0.33, respectively. A random effects ANOVA with CMV[MMP-I] as the predictor variable finds that CMV[MMP-I] explains 27.5 percent of the variance in CMV-adjusted correlations ($F = 6.7$, $p \leq 0.05$). Significant method effects remain even after the correlations are corrected for CMV employing MVT.

We conclude that the "constant method effect" assumption underpinning the MVT is not empirically supported. Contrary to Malhotra et al.'s speculation that violation of the constant method effect assumption does not have a material impact on conclusions, the above analysis shows that the impact is both substantial and significant. Importantly, the marker variable technique under-corrects when the potential CMV-based validity threat is high.

An examination of the marker variables identified in each study suggests that this under-correction may be due, in part, to the fact that the marker variable correlation is often between measures that are subject to low CMV, even when the focal correlation is subject to potentially high CMV. Examples of such low CMV measures in the sample analyzed include variables such as a categorical variable capturing highest level of education (Manson 1998), number of years using structured programming (Kim 1996), and number of years working in the tourism industry (Wober and Gretzel 2000). The pattern of corrections resulting from applying MVT indirectly rewards poor measurement techniques subject to high CMV by under-correcting the reported correlations. Consistent with Richardson et al. (2009), we find that the MVT does not offer a valid correction for the effect of CMV and we recommend against its use.

References

Doty, D. H., and Glick, W. H. 1998. "Common Methods Bias: Does Common Methods Variance Really Bias Results," *Organizational Research Methods* (1:4), pp. 374-406.

Kim, I. 1996. *The Effects of Individual, Managerial, Organizational, and Environmental Factors on the Adoption of Object Orientation in U.S. Organizations: An Empirical Test of the Technology Acceptance Model*, Unpublished doctoral dissertation, The University of Nebraska-Lincoln.

Lindell, M. K., and Whitney, D. J. 2001. "Accounting for Common Method Variance in Cross-Sectional Research Designs," *Journal of Applied Psychology* (86:1), pp. 114-121.

Malhotra, N. K., Kim, S. S., and Patil, A. 2006. "Common Method Variance in IS Research: A Comparison of Alternative Approaches and a Reanalysis of Past Research," *Management Science* (52:12), pp. 1865-1883.

Manson, D. P. 1998. *Determinants and Consequences of User Groupware Customization*, Unpublished doctoral dissertation, The Claremont Graduate University.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. 2003. "Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies," *Journal of Applied Psychology* (88:5), pp. 879-903.

Richardson, H. A., Simmering, M. J., and Sturman, M. C. 2009. "A Tale of Three Perspectives: Examining Post Hoc Statistical Techniques for Detection and Correction of Common Method Variance," *Organizational Research Methods* (doi: 10.1177/1094428109332834), March 11, pp. 1-39.

Sharma, R., Yetton, P., and Crawford, J. 2007. "Common Methods Bias: Reports of its Demise Are Highly Exaggerated," in *Proceedings of the 28th International Conference on Information Systems*, Montreal, December 9-12.

Wober, K., and Gretzel, U. 2000. "Tourism Managers' Adoption of Marketing Decision Support Systems," *Journal of Travel Research* (39:2), pp. 172-181.

[5]See Podsakoff et al. (2003, p. 893) for a critique of the marker variable technique. In particular, Podsakoff et al. argue that "it fails to control for some of the most powerful causes of common methods biases (e.g., implicit theories, consistency motif, social desirability)." Podsakoff et al. also question the validity of the "constant method effect" assumption and argue that this assumption is not likely to hold for many types of methods biases, such as those arising from implicit theories and social desirability.

Appendix E

Estimating CMV[MMP-I] in Method–Method Pairs

A number of alternative assumptions could be employed to rank order CMV[MMP] in Table 7. Two assumptions are examined here. One is the assumption that the magnitude of CMV in a correlation is constrained by the measurement method more susceptible to method variance. This results in the following rank ordering (0.0 = Very low, 3.0 = Very high):

$CMV_{CC} = 1.0$
$CMV_{BB} = CMV_{BC} = 2.0$
CMV_{PP} $CMV_{PB} = CMV_{PC} = 3.0$
$CMV_{OO} = CMV_{CO} = CMV_{BO} = CMV_{PO} = 0.0$

The other assumption is that the rank ordering of CMV[MMP] for each cell is a product of the rank ordering of the susceptibility of each method to method variance. This gives a finer-grained rank ordering as follows (0.0 = Very low, 9.0 = Very high):
$CMV_{CC} = 1.0$
$CMV_{BC} = 2.0$
$CMV_{PC} = 3.0$
$CMV_{BB} = 4.0$
$CMV_{PB} = 6.0$
$CMV_{PP} = 9.0$
$CMV_{OO} = CMV_{CO} = CMV_{BO} = CMV_{PO} = 0.0$

The differences in rank orderings for the two assumptions examined here and the assumption presented in the "Discussion" section of the paper are limited to three off-diagonal cells, BC, PC, and PB. Further research is needed to resolve this ambiguity.

Future research can employ the above technique to evaluate the effect of CMV in other research domains. Further, the matrix in Table 7 can be extended to include other methods employed in specific research domains being investigated. Measurement methods not included in Table 1 can be ranked based on an evaluation of the susceptibility of methods to CMV, as illustrated in the development of Table 2.

For instance, consider a method–method pair in which one variable is a self-report measure captured on a perceptually anchored scale (Agree – Disagree) and the other variable is measured by trained raters on a behaviorally anchored scale (Not at all – Very Often) rating a piece of text generated by the respondent capturing his/her behavioral responses to a certain stimulus. For this case, the "Susceptibility to CMV due to data source" would be rated low as the two measures are being rated independently by two different sources. However, the text being rated by trained raters is generated by the respondent and can introduce bias in the two ratings; the "Susceptibility to CMV due to response format" would be rated low as the two measures are being independently rated but in identical response formats; and "Susceptibility to CMV due to abstractness of measures" would be rated very low as the rater is rating behavioral responses. The overall CMV[MMP-I] would be between very low (for perceptually anchored with system-captured measures) and low (for perceptually anchored with behaviorally anchored measures) (see Table 2).