What Do You Mean, My Statistical Results Are Incorrect? The Impact of Multicollinearity and Measurement Error in Tests of Statistical Significance

Completed Research

Dale Goodhue
University of Georgia
dgoodhue@uga.edu

Will Lewis
william.w.lewis@gmail.com

Ron Thompson
Wake Forest University
thompsrl@wfu.edu

Abstract

When two or more predictor variables are highly correlated and contain measurement error (M+ME), regression and Partial Least Squares (PLS) beta coefficients and t statistics can be inaccurate (Goodhue, Lewis and Thompson, 2017). Corrections to the misleading values can be made by using the application created by Goodhue et al., or by correcting correlations for attenuation before running a regression. We examined research articles in three Management Information Systems journals over multiple years and discovered that, of the regression and PLS papers that reported correlations, about one-half (48%) were operating in the danger zone. In the 10 papers that provided sufficient information to use the Goodhue et al. (2017) application we found that, on average, the t statistics were biased by about 0.70. This suggests that when using regression or PLS, researchers should check for the M+ME bias and correct for it when found.

Keywords

Multicollinearity, measurement error, statistical significance, false positives, regression, PLS

Introduction

In a Journal of Applied Psychology paper (Takeuchi, Bolin and Lin, 2015), researchers set out to examine the impact of collectivistic cultural tendencies on organizational citizenship behavior (OCB). They suggested that employees from interdependent cultures like Taiwan would care relatively more about their group's well-being than their own. They hypothesized that those driven by "prosocial values motivations" would engage in more social and psychological behavior that supported organizational task performance.

The researchers surveyed 379 supervisor-subordinate pairs from 13 Taiwanese financial institutions, finding support for the above hypothesis (with $B = .11$, $t = 2.03$). The authors point out that three predictor variables are interrelated, with one correlation of $\rho = .54$, but they do not express concern for the impacts that multicollinearity may have had on their regression results or discuss the variance inflation factor (VIF). The Cronbach's alpha reliabilities for the two highly correlated constructs were .93 and .88. Based on their regression analysis, the authors conclude that organizations " … may be able to increase the occurrence of OCB when they hire employees who value helping others … ".

At first glance, the above anecdote appears to describe a successful research effort, where statistical significance is found for a hypothesized relationship. But what these, or any other researchers using regression analysis at the time, were not aware of is that when two predictor variables in an OLS (ordinary least squares) regression[1] equation are highly correlated and measured with error, the path estimates (beta coefficients) and t statistics will likely be biased (Goodhue, Lewis, and Thompson, 2017) (or GLT, 2017). While many researchers believe that sufficiently low variance inflation factor (VIF) values will ensure at least unbiased t statistics, GLT (2017) demonstrated that this isn't the case. They termed this situation the "Multicollinearity and Measurement Error (or M+ME) blind spot." This condition can produce inaccurate or misleading results for researchers using regression or any Partial Least Squares (PLS) technique that does not correct for measurement error for causal path analysis.

GLT (2017) not only showed the mathematics behind the M+ME blind spot, but they showed with an extensive Monte Carlo simulation that under conditions of measurement error and high correlation between predictor variables, regression and PLS path estimates and t statistics are biased, and can create excessive false positives (path estimates that are shown as being statistically significant but are not).

Fortunately there is a way to correct for the "blind spot" in a conceptually straightforward way, as we explain later. Returning to the previously described research study (Takeuchi et al.), when we corrected those results for M+ME bias, instead of the published result of $\beta = .11$, $t = 2.03$, $p<.05$, we found that the correct values were $\beta = .089$, $t = 1.34$, $p =.181$ (two-tailed) or $p = .091$ (one-tailed). These corrected results are not statistically significant. Because of the combination of a high correlation between predictor variables and measurement error, the authors were unknowingly publishing misleading results.

We are not suggesting that the authors who wrote the paper were in any way negligent in their research methods. Clearly, until very recently, virtually all researchers were unaware of the impact of M+ME. The authors of the above paper followed generally accepted practices and interpreted their results using generally accepted guidelines. With a better understanding of the M+ME impact, however, we believe that when there are high correlations and measurement errors, researchers should test for M+ME bias and correct for it when found.

The purpose of this paper is to determine how large a problem M+ME might be in recent (and future!) MIS research. First, we provide a brief overview of the key findings from GLT (2017), including why VIFs should not be relied upon to protect against multicollinearity. Next we share findings from a review of three Management Information Systems (MIS) academic journals to provide context on how frequently (or infrequently) researchers are operating under conditions where M+ME could be a concern. We then discuss how M+ME might also be a concern for researchers employing single-item indicators, and explore one published study in greater detail to illustrate the potential impact. Finally, we conclude with a call for researchers to be more careful to report correlations, reliabilities and t statistics (in addition to path values and asterisks indicating significance levels) in their publications, so that reviewers and readers will be able to assess the conditions for M+ME.

The M+ME Blind Spot: Why VIFs Don't Tell the Whole Story

When high correlations between predictor constructs exist, researchers frequently rely on the variance inflation factor (VIF) to protect against excessive false positives. The VIF increases the standard deviations of path estimates (Johnston 1972, Neter and Wasserman 1974, Pedhazur 1997), protecting researchers from concluding that biased regression path estimates are statistically significant. In addition, the VIF has long served as an indicator for when multicollinearity in an OLS regression is high enough to cause additional problems. In such a case, researchers might decide to modify their regression equation (e.g., by removing one of the correlated variables, or combining variables).

As noted by GLT (2017), Johnston (1972, pp. 161-163) showed that when constructs in a regression are correlated, there can be bias in path estimates. Green and Kiernan (1989) carried the analysis further, showing that when regression predictor constructs were correlated and measured with error, the bias would increase with increasing correlation and increasing measurement error. GLT (2017) added the

[1] For simplicity, we use the term "regression" here to refer to multiple regression using ordinary least squares. We are focusing on constructs measured either with reflective indicators or single indicators.

recognition that when regression predictor constructs are measured with error, the VIF is biased. If correlated constructs are measured with error, the calculated correlation will be attenuated by that measurement error (Nunnally and Bernstein 1994). Therefore the VIF correction (based on the calculated correlation) will be underestimated. The combination of biased path estimates and biased variance inflation factors leads to the M+ME blind spot. Putting together the equations for path bias and VIF bias, GLT (2017) demonstrated mathematically that when there are two highly correlated predictor variables measured with error in a regression, both path estimates and t statistics will often be biased. Although PLS differs from regression in the way it weights indicators to develop construct scores and correlations, it still uses ordinary least squares to determine path estimates. GLT (2017) explained why PLS would suffer the same problems as regression in the presence of correlated constructs measured with error[2]. GLT (2017) used Monte Carlo simulation to demonstrate the impact of the biases on false positives in both regression and PLS. Note that M+ME is not a concern for covariance-based structural equation modeling (CB-SEM), since CB-SEM incorporates measurement error in its estimations GLT, 2017).

As one example described in GLT (2017), a regression simulation was carried out with 500 samples of n=300 with an underlying correlation of .60 between one predictor construct with a zero effect and another with a large effect, both measured with .80 reliability. This simulation produced 65 false positives (13%) for the zero effect construct (see GLT (2017), Figure 4A.) Note that the VIF for the underlying correlation of .60 was 1.56, far less than any suggested danger level provided in the literature. Among other things, this demonstrates persuasively that relying on rules of thumb for the VIF as a way to protect against the dangers of multicollinearity is ill-advised. Additional simulations in GLT (2017) showed graphically the impact of greater correlations and smaller reliability on excessive false positives in regression, PLS and CB-SEM. The three panels of Figure 1 on the next page reproduce GLT's (2017) Figure 3. Figure 2 (GLT's 2017 Figure 7) on the following page, shows that the problem is also exacerbated when the path value of the true predictor in a correlated pair is relatively large, as suggested by Green and Kiernan's equations (1989). GLT (2017) also showed that M+ME bias is distinct from problems caused by a lack of measurement discriminant validity.

In addition to demonstrating the existence of M+ME bias, and the factors that contribute to it, GLT (2017) also proposed a correction application (available at www.misq.org/online-supplements-2017 September) that took as input the correlation between two highly correlated predictor constructs, their beta values, t statistics, and reliabilities, and calculated corrected beta and t statistic values. Though not stated explicitly in GLT (2017), it is apparent that this application assumes that the bulk of the problem in any given regression is the single high correlation between one pair of predictor constructs. This was the case with virtually all the simulations GLT (2017) ran. Therefore, when they used their correction application with the datasets producing the results in Figure 2, the false positives were reduced to an acceptable percent, as shown in Figure 3 below.

GLT (2017) did suggest that when there were more than one pair of high correlations in a single regression equation, the correction application should not be used. But they were vague about how high any "additional" correlations would have to be to cause problems with the application's results. Their work shows that the correction application works well when the bulk of the problem comes from a single high correlation, but is less accurate when there is more than one pair of highly correlated constructs.

A much more general correction approach for M+ME problems (including those with multiple high correlations) is available. The biased correlation matrix can be corrected for measurement error (Cohen et al., 2003; Spearman 1904) using Nunnally and Bernstein's (1994) well known equation for the underlying correlation, where α1 and α2 are the measurement reliabilities of the two constructs, ρ_{Overt} is the calculated correlation observed by a researcher (which may be attenuated by measurement error) and $\rho_{\text{Underlying}}$ is an overt correlation that has been corrected for random measurement error:

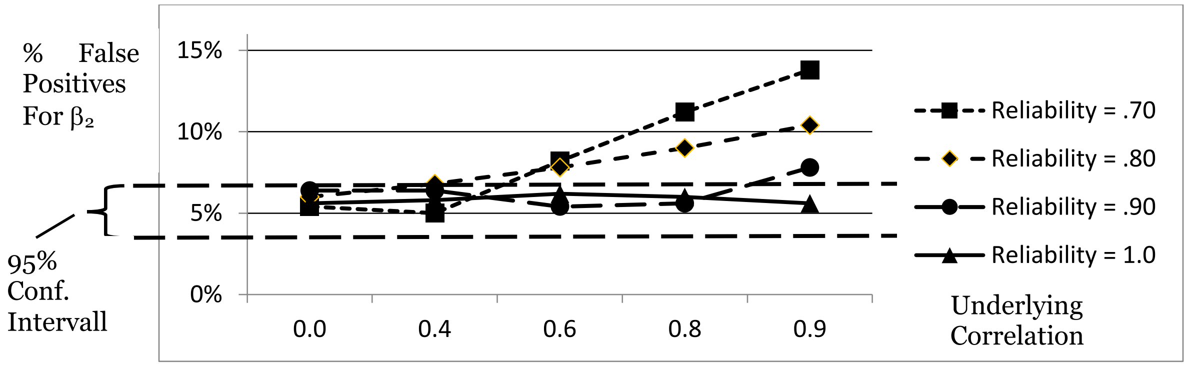
---

[2] We note that "consistent PLS" appears to use a correlation matrix corrected for measurement error. Although it is another matter, we also note that even consistent PLS does not address the underlying problem of capitalization on chance inherent in the PLS approach for weighting indicators (Goodhue, Lewis, and Thompson 2015, Rönkkö and Evermann 2013).

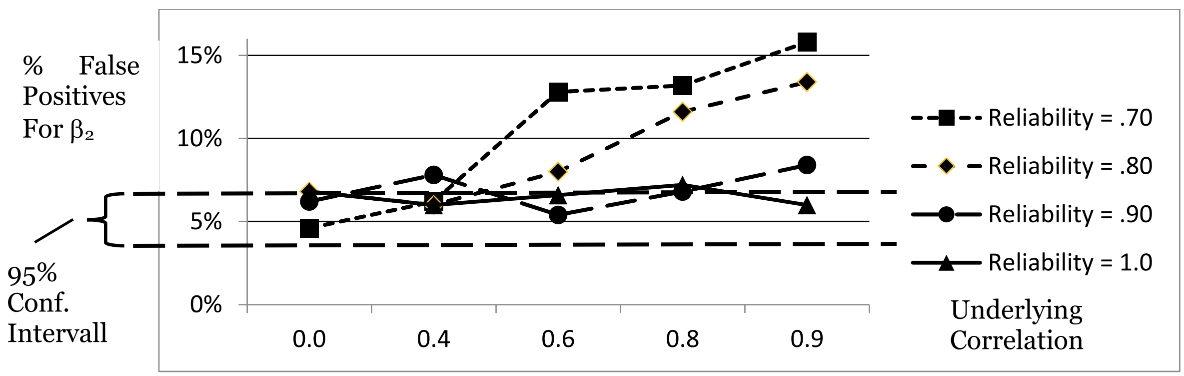
$$\rho_{\text{Underlying}} = \rho_{\text{Overt}} / \left((\alpha 1 * \alpha 2)^{1/2} \right)$$

Because the corrected correlation matrix contains "correlations corrected for measurement error", it can be used as the basis for a regression (or PLS) analysis, and the M+ME bias will be avoided.

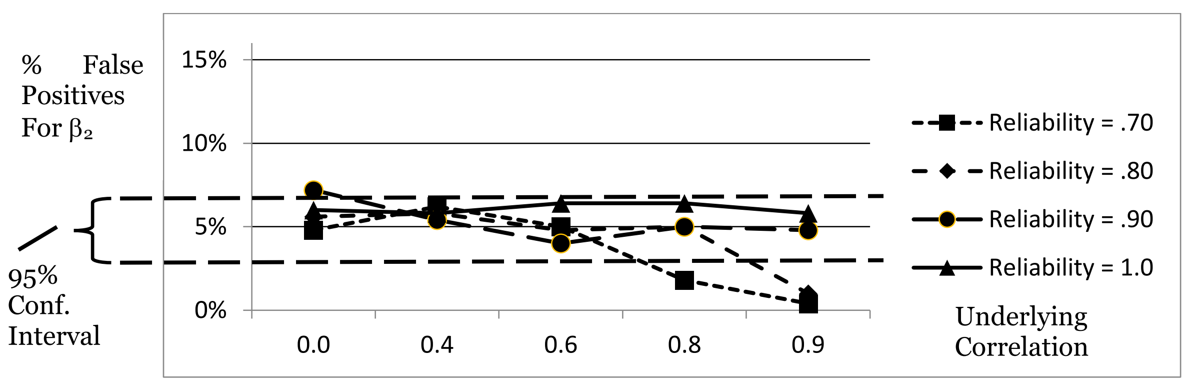
For those interested, the Appendix shows an Excel spreadsheet that automates the calculations for the correlation matrix correction. Once the correlation matrix has been corrected for measurement error, it can be used as the starting point for any regression or PLS analysis. Of course care must be taken to correct the revised regression or PLS results for sample size. For example, SAS will carry out the regression based on a correlation matrix, but unless otherwise specified, will assume a sample size of 10,000. T statistics based on a sample size of 10,000 must be adjusted for the actual sample sizes.



Panel 1A. Regression – % of False Positives at Different Levels of ρ_{12}



Panel 1B. PLS – % of False Positives at Different Levels of ρ_1



Panel 1C. CB-SEM – % of False Positives at Different Levels of ρ_{12}

Figure 1. False Positives From High Correlations and Measurement Error

(500 samples of N=100, β_1=.292 (Reprinted with permission from Goodhue et al. 2017's Figure 3)

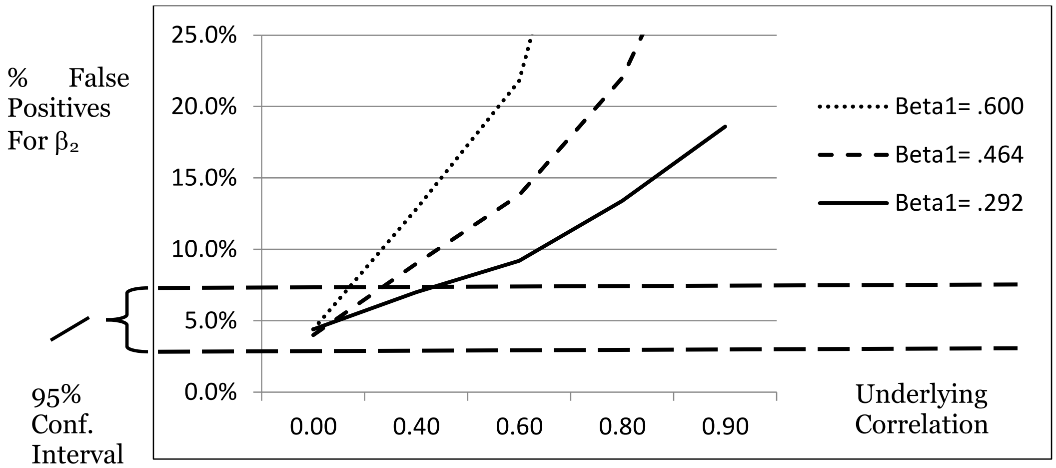


Figure 2. Regression False Positives as β_1 increases (from .292 to .600).
(Alpha=.80, N= 200) Reprinted with permission from GLT 2017's Figure 7).

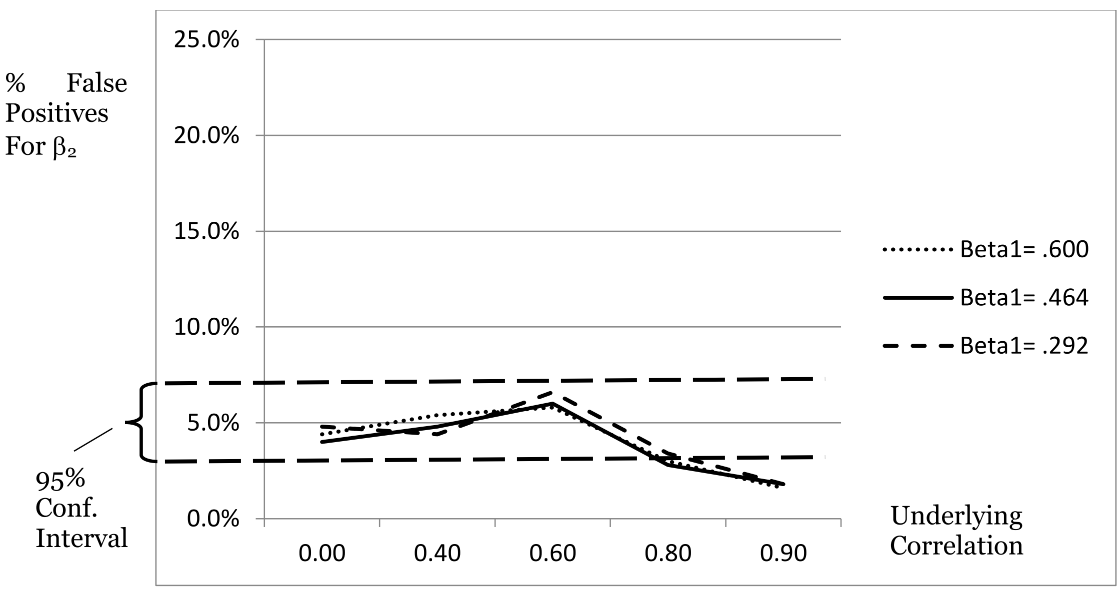


Figure 3. Regression Results after Correcting the Analyses in Figure 2

(Alpha=.80, N = 200). (Reprinted with permission from GLT 2017's Figure 9).

Since it is easier to use, the GLT (2017) correction application offers a quick and dirty way to decide whether M+ME problems might be affecting statistical results. For analysis where M+ME bias seems consequential, the more accurate but burdensome correction approach (corrected correlation matrices) should be used.

How Common Is M+ME? Review of MIS Articles

To answer the above question, we looked at three highly respected MIS journals (MIS Quarterly, Information Systems Research, and the Journal of Management Information Systems), each over a period of two to three years. Table 1 provides an overview of the results of our literature search. Note that columns B through H start with the total number of papers, and each column reduces that count showing the resulting number that met all the requirements in earlier columns. We will use the Totals row to illustrate how to read Table 1. For these three journals in the specified years, Columns B and C show that there were 254 papers total, 113 of which used Regression or PLS. Column D shows that about a third of the Regression and PLS papers (32%) did not publish correlation tables, so we will exclude them from our analysis, leaving 77. Column E shows that of the 77 regression and PLS papers that did report correlations, about half (37 of the 77) had at least one pair of highly correlated constructs (which we are defining as being greater than .50) occurring in the same regression equation, and are thus operating in what GLT (2017) call the "M+ME Danger Zone".

Of the 37 papers operating in the danger zone[3], Column F shows that only 16 had multiple items measures (which allow researchers to estimate reliabilities), and only 12 (in Column G) of those papers reported reliabilities for the relevant constructs. Finally, of those 12, Column H shows that only 10 reported enough information for us to also know the t statistics for the relevant regression paths.

A. Journal/Years	B. # of Papers Published	C. # Regression/PLS	D. # Correlation Matrix	E. # I.V.'s corr. > .50 [DANGER ZONNE]	F. # Multi- Item Meas.	G. # Reliabilities Provided	H. # t-stats Available [Enough Info To Check For M+ME]	I. Weighted Average Change In t-Stat	J. # Changed Type A	K. # Changed Type B
MISQ 2014-15	99	33	28	12	4	2	1	2.178	1	0
ISR 2014-15	92	23	22	12	4	4	4	.636	0	1
JMIS 2012-14	63	57	27	13	8	6	5	.460	1	0
Totals:	**254**	**113**	**77**	**37**	**16**	**12**	**10**	**.729**	**2**	**1**

Table 1. Summary Statistics for M+ME Conditions – Multiple Item Measures

For these 10 with complete enough information, we used GLT 2017's correction application to get a quick indication of papers where reported statistical results for hypothesized paths might be questionable (i.e., where the level of statistical significance changed when the correction was applied). For these 10 the weighted average change in the t statistic for the high correlation construct paths was .729, a considerable amount. Columns J and K show the numbers of papers where the correction caused changes in the level of statistical significance for hypothesized paths. Type A is a change in the significance of a hypothesized path from "significant" to "not significant" or vice versa, while Type B is a change in the level of significance in significant paths, such as from .01 to .05. These three articles were worthy of further attention.

Table 2 compares the published path estimate results for those three papers with the results from two different correction approaches for the two highly correlated constructs. Cells where the corrected results are different from the published results are shown in bold and slightly larger type. In two (the first and third papers), at least one hypothesized path that was reported as statically significant was not significant when the corrections were applied. In the second paper the level of statistical significance in a hypothesized path was changed, in this case from $p<.01$ to $p<.05$.

---

[3] Note that the authors plan on expanding the review of articles published in these and other journals; what is reported here is the subset of the review that was completed by the deadline for submitting this article.

In two of the papers (the second and the third, marked by "##") in addition to the focal high correlation, there was also a second high correlation in the same regression equation. This suggests that the GLT (2017) correction application should not be relied upon. For the second and third papers, therefore, only the corrected correlation matrix results should be deemed correct. Even so, in the first two examples, the two correction approaches give very close to the same results for the t statistics, and both show a change in the level of significance. Because we cannot use the Corrected Correlation Matrix approach for the third paper, due to the interaction effect in that regression, the GLT Correction Application results should be viewed with caution. In the third paper, the difference between the problem correlation results for the two different correction approaches is likely due to the influence of a second high correlation.

Journal and Source of Statistical Results	Problem Correlation	Construct One			Construct Two		
		Path	t Stat	p<	Path	t Stat	p<
JMIS As Published	0.68	0.50	8.34	.000	0.18	**3.04**	**.003**
GLT Correction Application	0.75	0.59	7.03	.000	0.14	**1.63**	**.104**
Corrected Correl. Matrix	0.75	0.60	9.52	.000	0.12	**1.81**	**.071**
ISR As Published	0.61	1.47	**2.49**	**.007**	2.16	3.43	.001
GLT Correction Appl. ##	0.64	1.48	**2.25**	**.013**	2.26	3.22	.001
Corrected Correl. Matrix	0.64	0.16	**2.34**	**.0101**	0.24	3.38	.001
MISQ As Published	0.51	0.59	***	.001	0.12	*	**.050**
GLT Correction Appl. ##	0.67	0.89	8.60	.000	-0.03	**-0.65**	**.519**
Corrected Correl. Matrix	0.62	*	*	*	*	*	*

Table 2. Comparing the Two Different Correction Approaches

(* Because the full regression in the third paper contains an interaction term (and the interaction term is not included in the published correlation table), we cannot recover a corrected regression for that paper.)

Single Item Measures

GLT (2017) focused on analyses where the measurement error can be estimated, where multiple items are used as a proxy for any given predictor variable. When single item variables are employed, researchers don't necessarily assume that the variables are measured without error, but since there is no easy way to know the extent of measurement error or to incorporate that in a regression or PLS analysis, the default situation is to assume no error. In many (if not most) situations, however, single item measures are not completely accurate. Even values reported by providers such as Compustat will have some error, since they are dependent on reporting by many different organizations that will vary in the accuracy of their reported values. Another example might be the use of total IT expenditures by a firm as a predictor of firm performance. No matter what technique is employed to gather data on total IT expenditures, it is highly likely that the resulting measure will contain some error.

In Table 3, we repeat some of the information provided in Table 1, only now we focus on articles where single item measures were used. Since we have no way of accurately knowing what the true reliabilities are for the single item measures that were employed, we decided to use a relatively simple approach for exploring the extent to which M+ME *might* come into play. Specifically, we decided to see what impact M+ME would have if the actual reliability for a single item measure were .90, instead of the perfect 1.00 that is commonly assumed. That is, we applied the GLT (2017) correction application and used .90 as the reliability for both of the correlated predictor variables for analyses using single items in *MIS Quarterly* and *Information Systems Research*. What Table 3 shows is that even assuming a fairly high reliability for single item measures, there is the possibility that many of our papers with single item measures could also be reporting inaccurate levels of statistical significance.

The choice of .90 for reliability for a single item measure in Table 3 is somewhat arbitrary[4]. However, in any given case the GLT (2017) correction algorithm could be used for sensitivity analysis, to gauge how results obtained from PLS or regression might be affected by M+ME.

A. Journal/Year	B. # of Papers Published	C. # Regression/PLS	D. # Correlation Matrix	E. # I.V.'s corr. > .50 [DANGER ZONNE]	F. # Single- Item Meas.	G. # t-stats Available [Enough Info To Check For M+Me]	H. Average Change In t-Stat	I. # Changed Type A	J. # Changed Type B	K. # More than 2 highly correlated constructs
MISQ 2014-15	Same Numbers as Table 1				8	8	.525	2	1	0
ISR 2014-15	Same Numbers as Table 1				8	8	1.164	1	2	3
Totals:	Same Numbers as Table 1				16	16	.845	3	3	3

Table 3. Summary M+ME Results for Single Item Measures, Assuming Reliability = .90

As an example, consider Langer et al., (2014). As part of a larger regression model, team size and the log of the number of function points (logFP) as a measure of software size were predicted to influence the cost performance for software projects. LogFP and team size were correlated at .68, and the sample size was 230. Langer et al. (2014) obtained estimates that were statistically significant for both predictors (beta = -1.111 and t = -7.507 for logFP, and beta = -0.234 and t = - 2.294 for team size). This resulted in an estimate for the impact of team size on cost performance that was statistically significant at $p < .05$. As we noted previously, however, these predictor variables were assumed to be measured without error. By employing the GLT (2017) correction application in sensitively analysis, we tested to see what impact it would have on the results if different assumptions about measurement error were made. By trying different levels of reliability for the predictors, we found that the corrected values for the impact of team size were:

- At α = .98 for both predictors, beta = -.213, t = -1.97; p = .05 (2-sided), p = .025 (1-sided)
- At α = .95 for both predictors, beta = -.173, t = -1.454; p = .147 (2-sided), p = .074 (1-sided)
- At α = .90 for both predictors, beta = -.077, t = -.524, p = .601 (2-sided), p = .300 (1-sided)

As can be readily observed, the estimate (and statistical significance) of team size in this regression equation is quite sensitive to the accuracy of the measure. Once again, we are not suggesting that Langer et al. (2014) did anything incorrect in their analysis. Rather, we are suggesting that if researchers begin to relax the assumption that their single item indicators are measured completely without error and use the GLT (2017) application to perform some sensitivity analysis addressing the assumption of perfect measurement, they might find that in some situations their results (and their subsequent interpretations of those results) could change. At the very least we might be more cautious in interpreting results that depend on perfect or very high accuracy for their statistical significance.

[4] Wanous and Hudy 2001, looking at student evaluations of teaching effectiveness suggest an estimate of a minimum of .80 for group level reliability, and a minimum of .70 for individual level data. Clearly any estimate of reliability for a single item measure will depend on many factors.

Conclusion

It would be possible to interpret the research presented here as a lot of noise but little substance. After reviewing 254 MIS research papers, we found only three where M+ME is clearly a problem.

But consider the following. Though there were many articles (254 total) published, we are only interested in the 113 papers using regression or PLS. Of those 113 that did show correlation matrices (77), about half (37) were operating in the danger zone with correlations above .50. It is not unreasonable to estimate that had all 113 shown correlation matrices, about half of the total (56 or so papers) would have been observed to be operating in the M+ME danger zone.

Only 10 of the 37 clearly in the danger zone reported enough information for us to assess whether results were being affected by M+ME problems. Of those 10 where we had complete enough information, 30% were reporting incorrect levels of statistical significance. What if we applied that 30% ratio of "danger zone" to "incorrect reporting" to all of the hypothetical 56 that probably were operating in the danger zone? In that case we would be talking about perhaps 17 papers reporting erroneous results. That is a number high enough to be concerning. Furthermore, we demonstrated that if we relax the assumption that single item measures are error-free, the possibility of M+ME coming into play in additional papers is also quite real. We argue that researchers employing single-item measures should also consider M+ME, and take corrective action when M+ME danger signals are apparent.

Surely some caveats are in order. First, most of these papers have more than one hypothesis, and we typically only found one error in the papers we labeled as having an error, so we are not saying that 17/113 = 15% of our hypotheses are in error.

Secondly, our literature review was limited to articles published over two years in *MIS Quarterly* and *Information Systems Research*, and three years for the *Journal of Management Information Systems*. We do plan on expanding our literature review to include both more years and also more journals across multiple disciplines (e.g., Organizational Behavior and Marketing). As a prelude to that future study, we looked at the *Journal of Applied Psychology* for 2015 and found that of 117 papers, 33 were using regression or PLS and were operating in the danger zone. Of those, only 12 provided enough information to use the correction application, and of those, two had hypotheses incorrectly reported as statistically significant, and two more had more than two highly correlated predictors in the same regression. So the problem is not unique to MIS.

Third, GLT (2017) did not address other types of regression (e.g., probit, logit, multi-level, etc.) in their assessment. While we expect M+ME should likely be a concern with these techniques as well, we currently have no definitive way of testing that assertion. Regardless of these limitations, the current paper goes further than GLT (2017) by showing that in MIS research, M+ME biases are certainly leading to identifiable papers with incorrect interpretations of results (more specifically, to false positives). Extrapolating our results to the larger set of MIS research suggests a substantial number of papers with correctable errors in statistical results. In addition, this paper suggests that the GLT (2017) correction application, though valuable as a quick and dirty assessment of M+ME problems, should probably be replaced by a more accurate approach: correcting correlation matrices for measurement error.

We wish to reiterate that we are not suggesting that any of the MIS researchers who have published articles that we reviewed did anything wrong. They were following commonly accepted practices and procedures. What we are suggesting is that now that we as a research community know more about the potential impact of M+ME, we need to modify our practices and procedures for handling measurement error going forward. More specifically, we recommend that: (1) reviewers and editors insist that researchers report all necessary information required to assess the potential impact of M+ME; correlations among all variables, path coefficient estimates, t statistics, (or beta standard errors), construct reliabilities, and sample sizes; and (2) when the M+ME conditions are in the danger zone (e.g., where correlations among predictor variables exceed .50), researchers should take steps to assess whether there are M+ME problems, and if there are, to correct their reported results.

REFERENCES

Cohen, J., Cohen, P., West, S., and Aiken, L. 2003. *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3rd Ed.), Routledge: New York, New York.

Goodhue, D., Lewis, W. and Thompson, R. 2017. "A Multicollinearity and Measurement Error Statistical Blind Spot: Correcting False Positives in Regression and PLS," *MIS Quarterly*, (41:3), pp. 667-684.

Goodhue, D., Lewis, W. and Thompson, R. 2015. "PLS Pluses And Minuses In Path Estimation Accuracy," AMCIS 2015 Proceedings.

Green, C., J. and Kiernan, E. 1989. "Multicollinearity and Measurement Error in Econometric Financial Modelling," *The Manchester School* (57:4), pp. 357-369.

Johnston, J. 1972. *Econometric Methods (2nd ed)*, New York, NY: McGraw-Hill.

Langer, N., Slaughter, S., and Mukhopadhyay, R. 2014. "Project Managers' Practical Intelligence and Project Performance in Software Offshore Outsourcing: A Field Study," *Information Systems Research*, (25:2), pp. 364-384.

Neter, J., W. Wasserman and Kutner, M. 1985. *Applied linear statistical models (2nd ed.)*, Homewood, IL: Richard D. Irwin.

Neter, J., and Wasserman, W. 1974. *Applied linear statistical models (1st ed.)*, Homewood, IL: Richard D. Irwin.

Nunnally, J. C., and Bernstein, I.H. 1994. *Psychometric Theory (3rd ed.)*, New York: McGraw-Hill.

Pedhazur, E. J. 1997. *Multiple regression in Behavioral Research (3rd ed.)*, Australia: Wadsworth.

Rönkkö , M. and Evermann, J. 2013 "A Critical Examination of Common Beliefs About Partial Least Squares Path Modeling", *Organizational Research Methods*, (16:3), pp. 425–448.

Spearman, C. 1904. "The Proof and Measurement of Association between Two Things," The American Journal of Psychology, (15:1), pp. 72-101.

Takeuchi, R., Bolino, M.C, and Lin, C.C. 2015. "Too Many Motives? The Interactive Effects of Multiple Motives on Organizational Citizenship Behavior," *Journal of Applied Psychology* (100:4), pp. 1239-1248.

Wanous, J.P., and Hudy, M.J. 2001. "Single-Item Reliability: A Replication and Extension," *Organizational Research Methods*, (4:4), pp. 361-375.

APPENDIX: Correcting a Correlation Matrix for Measurement Error

The spreadsheet below takes the overt correlations shown in E2 to G4, combined with the reliabilities in E1 to G1 and D2 to D4, and produces the corrected correlations in E6 to G8. Note that the Excel equations generating the corrected correlations are shown in cells C&D10, C&D11, and E&F11. This general approach can be expanded to larger correlation matrices.

	A	B	C	D	E	F	G
1				Cronbach's α	0.91	0.89	0.91
2	Overt Correlation Matrix		InfoEXch	0.91	1		
3			Trust	0.89	0.63	1	
4			ConfRes	0.91	0.51	0.56	1
5							
6	Corrected Correlation Matrix			InfoEXch	1		
7				Trust	0.700	1	
8				ConfRes	0.560	0.622	1
9							
10	Equations for the Corrections in E7,E8,F8		=E3/(($D3*E$1)^0.5)				
11			=E4/(($D4*E$1)^0.5)		=F4/(($D4*F$1)^0.5)		

Figure A1. Spreadsheet Example Correcting a Correlation Matrix for Measurement Error