

The authors examine the construct validation results of 70 published data sets. The analysis shows that, on average, traits account for less than 50% of the variance in construct measures. These findings raise questions about the application of statistical techniques that assume minimal measurement error or do not properly model systematic measurement error.

Estimating Trait, Method, and Error Variance: Generalizing Across 70 Construct Validation Studies

Measures of a construct have variance due to the construct or trait being measured and variance due to measurement error. Measurement error can be divided into two components, random error and systematic error (the term "measurement error" refers to the combination of random error and method effect). A major source of systematic measurement error is method effect (Fiske 1982), which is the influence of the measurement instrument on the variance in a measure. Measurement error is omnipresent and widely recognized (e.g., Fiske 1982; Peter 1981). Campbell (1969) has gone so far as to say that measurement error (both random error and method effect) and its confounding influences on research findings cannot be avoided.

Researchers have long recognized the problems caused by measurement error and many have examined the validity of individual measures used in social science research. Numerous social science studies have examined construct validity by multitrait-multimethod (MTMM) analysis (see Phillips 1982, John and Reve 1982, and Seymour and Lessne 1984 for recent examples in the marketing literature; Turner 1981 lists 70 studies using MTMM analysis). The value of previous attempts to determine the validity of various measures has been limited for two reasons. First, no attempt has been made to generalize the results across a large number of samples or constructs. As noted by Cattell et al. (1969), no re-

searcher should rely on the outcome of a single correlation matrix from a single experiment. The deficiencies involved in addressing complex theoretical issues with any one empirical research study include variation due to the sample examined, bias due to the construction and administration of measurement instruments, recording and computational errors, and the lack of external validity (Hunter, Schmidt, and Jackson 1982).

A second limiting factor is the procedures that have been used to analyze data. In most studies, Campbell and Fiske's (1959) methodology has been used to analyze MTMM data. This procedure does not yield specific estimates of trait, method, and random error variance (Schmitt, Coyle, and Saari 1977). Fewer studies have applied confirmatory factor analysis (CFA) to estimate the size of the variance components and, as Widaman (1985) points out, even those studies may not have applied CFA appropriately.

The purpose of our study is to estimate the amount of trait, method, and random error variance in measures across a large sample of studies and constructs in the social sciences, using Widaman's (1985) suggested procedure to overcome the problems with previous single-sample analyses. These estimates make possible a generalized evaluation of the relative amount of trait, method, and random error variance in social science research and provide a more complete picture than the single-sample studies previously reported in the literature. Knowing the extent to which measurement error is a problem also enables us to discuss the implications of applying statistical techniques in which measurement error is assumed to be minimal. Because marketing researchers tend to use ab-

*Joseph A. Cote is Assistant Professor of Marketing, Washington State University. M. Ronald Buckley is Assistant Professor of Management, University of Oklahoma.

stract concepts that may be more prone to measurement error, applying statistical techniques that assume no measurement error may be problematic. The seriousness of this problem can be assessed by using the estimates of measurement error provided by our study.

METHOD

All of the MTMM matrices analyzed were gathered from published sources. Several approaches were used to find published MTMM matrices. One source was Turner's (1981) article listing 70 studies in which the MTMM method was applied. Another source was a computer search using "multitrait multimethod" and "MTMM" as key words. Citations generated by these two procedures were examined to determine whether they contained a published MTMM matrix. Finally, back volumes of 114 social science journals were examined visually (more than 2500 issues).

These procedures uncovered more than 200 studies with published MTMM matrices (some with multiple data sets). However, not all of the matrices found could be examined by means of confirmatory factor analysis. Sixty-four usable studies reporting a total of 70 data sets were examined. These studies cover a wide variety of disciplines, measurement methods, and constructs. A description of the studies can be obtained from the first author.

Confirmatory factor analysis was used to analyze the MTMM matrices. Estimates of trait, method, and random error variance were obtained by the nested models procedure outlined by Widaman (1985). The first step suggested by Widaman is to test for the presence of trait and method variance by using a series of hierarchically nested models. Four basic models were examined to determine whether trait or method variance is present.

Model 1 is a null model in which the variance in the measures is explained only by random error (no trait or method factors).

Model 2 is a trait-only model in which the variance in the measures is explained by trait factors and random error (t trait factors, freely estimated intercorrelations, no method factors).

Model 3 is a method-only model in which the variance in the measures is explained by method factors and random error (m method factors, freely estimated intercorrelations, no trait factors).

Model 4 is a trait and method model in which the variance in the measures is explained by trait factors, method factors, and random error (t trait factors, m method factors, freely estimated intercorrelation among trait factors, freely estimated intercorrelations among method factors, fixed zero intercorrelations between trait and method factors).

The presence of trait factors can be determined by examining the improvement in the chi square goodness-of-fit value caused by adding trait factors to the null model and to the method model. If trait factors are present, the trait model should have a significantly better fit than the null model *and* the trait and method model should have

a significantly better fit than the method model. In other words, model 2 should have a significantly lower chi square goodness of fit than model 1 *and* model 4 should have a significantly lower chi square goodness of fit than model 3. If both of these conditions are met, one can conclude that trait factors are present.

The presence of method factors is determined by using a similar rationale. If method factors are present, the method model (model 3) should have a significantly better fit than the null model (model 1) *and* the trait and method model (model 4) should have a significantly better fit than the trait model (model 2). If both these conditions hold, one can conclude that method factors are present.

Once trait and method variance have been shown to exist, their magnitude can be estimated. The square of the trait factor loadings indicates the percentage of variance in a measure due to traits and the square of the method factor loadings indicates the amount of variance due to methods (Widaman 1985). Fisher's r to z transformation (see Hays 1973) was used to calculate the average amount of variance due to trait, method, and random error across the 70 data sets.

Methodological Limitations

Trait by method intercorrelations were not modeled because estimating them ". . . present[s] both logical and empirical estimation problems of great magnitude" (Widaman 1985, p. 7). In addition, other nonrandom measurement effects (constant tendency, halo, experimenter, or experimental effects) may be present. Without more detailed knowledge about measures and experimental methodology, such effects cannot be modeled accurately and no attempt was made to identify or model them. As a result, the measurement model *may* not fit the data. If the model does not fit the data, the estimates of trait and method variance *may* be biased and the direction of the possible bias cannot be determined. Therefore, we assume that trait by method interactions and other nonrandom measurement effects are minimal and that the estimates of trait and method variance are accurate (Widaman 1985).

RESULTS

Our findings are in agreement with those from studies performed on single samples showing measurement error to be a serious problem. Significant trait variance is found for all but one of the studies examined and significant method variance is found for all but three of the studies. On the average (weighted by sample size and after Fisher's r to z transformation), measures contain 41.7% trait variance, 26.3% method variance, and 32.0% random error variance (see Table 1). The average correlation between traits is .674 and the average correlation between methods is .484. These results indicate that, in addition to being pervasive, measurement errors are relatively large. Most important, trait variance accounts

Table 1
SUMMARY STATISTICS FOR THE ESTIMATION OF TRAIT, METHOD, AND ERROR COMPONENTS^a

Data set	Number of studies	Percentage of variance due to:			Average correlation among:	
		Traits	Methods	Error	Traits	Methods
All studies	70	41.7	26.3	32.0	.674	.484
<i>Breakdown by discipline</i>						
Marketing	6	68.4	15.8	15.8	.486	.315
Psych/sociology	31	36.2	28.9	34.9	.701	.600
Other business	19	43.6	23.8	32.6	.612	.375
Education	14	34.9	30.5	34.6	.744	.504
<i>Breakdown by construct measured^b</i>						
Attitudes	11	29.8	40.7	29.5	.850	.556
Personality	27	39.1	24.7	36.2	.623	.546
Aptitude/achievement	6	39.5	25.1	35.4	.777	.572
Job performance/satisfaction	20	46.5	22.5	31.0	.567	.328

^aAll differences in the reported values are statistically significant ($P < 0.05$) with the following exceptions: measures in psychology and education have the same error variance; measures of personality and aptitude have the same trait, method, and error variance; measures of attitude and aptitude have the same correlation among methods; and measures of attitude and personality have the same correlation among methods.

^bSix studies examined constructs that could not be categorized with those of any other study. These constructs were omitted from this section of the table.

for less than 50% of the observed variance in measures.

Measurement theory indicates that abstract constructs may be more difficult to measure than concrete constructs. To test this belief, we categorized the studies by the type of construct measured (see Table 1). Attitude measures have the least trait variance (29.8%). Personality and aptitude are the second most difficult to measure (39.1% and 39.5% trait variance, respectively) and job performance has the most trait variance (46.5%). These findings confirm the notion that the more abstract constructs are more difficult to measure.

The reported results include measures from a wide variety of social science disciplines. Comparison of measures used in different disciplines may be of interest. On the basis of the journal in which the study was reported, the constructs examined, and content of the article as inferred from the title, six of the studies were categorized as marketing studies, 31 as psychology/sociology studies, 19 as other business areas (primarily management), and 14 as education studies. Measures used in the marketing studies have 68.4% trait variance, 15.8% method variance, and 15.8% random error variance (see Table 1). These findings indicate that measures used by marketing researchers may be better than those generally used in other social science disciplines ($p < 0.05$). However, this result is tempered by the fact that several of the marketing studies were attempts to develop and validate measures. Therefore, they would be expected to have more trait variance than measures typically utilized in research studies.

IMPLICATIONS

In most research it is assumed that the measures used are valid and that measurement error is inconsequential.

Trait variance often is assumed to be so large in relation to measurement error that it is acceptable to apply techniques like regression without correcting for measurement error. Reviews by Wilkie and Pessemier (1973) and Peter (1979) highlight the pervasiveness of this assumption. Twelve multiattribute studies reviewed by Wilkie and Pessemier used regression analysis, yet none investigated measurement reliability or validity, and Peter notes that less than 5% of the studies he reviewed assessed the reliability of the measures employed.

Our findings indicate the general assumption of minimal measurement error is highly questionable. Measurement error, on average, accounts for most of the variance in a measure. This observation raises questions about the practice of applying statistical techniques based on the assumption that trait variance is large in relation to measurement error variance. For example, in classical regression analysis one assumes that the independent variables are measured without error. If only the independent variable is measured with error and the errors and true values are uncorrelated, the bias caused by measurement error for the two-variable case can be calculated by using equation 1 (Johnston 1972, p. 282). The effects of measurement error on multivariate regression cannot be determined easily and vary from case to case.

$$(1) \quad \text{plim } \hat{B} = \frac{B}{1 + \sigma_e^2/\sigma_x^2}$$

On average, the amount of variance in a measure due to traits (σ_x) is 41.7% and the amount due to measurement error (σ_e) is 58.3% (26.3% method variance and 32.0% random error). Inserting these values into equation 1 indicates that the true value for the slope (B) in a

two-variable case would be 2.4 times greater than the estimated value (\hat{B}).

Structural equation models have been offered as a means for modeling measurement error (Bagozzi 1980). However, such models (e.g., LISREL) cannot be applied correctly without properly modeling measurement error. Darden, Carlson, and Hampton (1984) note that marketing researchers tend to use models that specify only latent factors (traits) and random error. Correlated errors (an indication of possible method variance) are added *post hoc* on the basis of the largest first derivative (Darden, Carlson, and Hampton 1984). This *post hoc* process produces a model that does not properly specify method effects. Given the "average" measurement conditions found in the social sciences, using these misspecified models is highly questionable.

The confounding effects of measurement error on data analysis can be controlled by the *proper* modeling of measurement error with structural equations. By using multiple measures, one can estimate the influence of method variance and random error components by analyzing the multitrait-multimethod matrix with confirmatory factor analysis. The researcher then can use the confirmatory factor analysis estimates to model the method effects using *prespecified* correlated measurement errors. Equality constraints can be used to ensure the model is identified (John and Reve 1982).

CONCLUSION

Overall, our study shows that the measures used in social science research have a large error component. Though this problem has been pointed out before in individual studies, our findings demonstrate measurement error to be present over a large variety of samples and constructs. Large measurement error indicates that the statistical techniques currently being used may be misapplied. More effective validation of the constructs used in research is essential. In the future, researchers must be more resolute in their desire to develop construct measures that are valid and free of measurement error.

REFERENCES

- Bagozzi, Richard P. (1980), *Causal Models in Marketing*. New York: John Wiley & Sons, Inc.
- Campbell, Donald T. (1969), "Definitional Versus Multiple Operationalism," *Et Al.*, 2 (March), 14-17.
- and Donald W. Fiske (1959), "Convergent and Discriminant Validation by the Multitrait-Multimethod Matrix," *Psychological Bulletin*, 56 (March), 81-105.
- Cattel, R. B., Karel R. Balcar, J. L. Horn, and J. R. Nesselroade (1969), "Factor Matching Procedures: An Improvement of the *S* Index: With Tables," *Educational and Psychological Measurement*, 29 (Winter), 781-92.
- Darden, William R., S. Michael Carlson, and Ronald D. Hampton (1984), "Issues in Fitting Theoretical and Measurement Models in Marketing," *Journal of Business Research*, 12 (September), 273-96.
- Fiske, Donald W. (1982), "Convergent-Discriminant Validation in Measurement and Research Strategies," in *New Directions for Methodology of Social and Behavioral Science: Forms of Validity in Research*, D. Brindberg and L. Kidder, eds. San Francisco: Jossey-Bass Publishers, 72-92.
- Hays, William H. (1973), *Statistics for the Social Sciences*. New York: Holt, Rinehart and Winston, Inc.
- Hunter, John E., Frank L. Schmidt, and Gregg B. Jackson (1982), *Meta-Analysis: Cumulating Research Findings Across Studies*. Beverly Hills, CA: Sage Publications.
- John, George and Torger Reve (1982), "The Reliability and Validity of Key Informant Data from Dyadic Relationships in Marketing Channels," *Journal of Marketing Research*, 19 (November), 517-24.
- Johnston, J. (1972), *Econometric Methods*. New York: McGraw-Hill Book Company, 282-3.
- Peter, J. Paul (1979), "Reliability: A Review of Psychometric Basics and Recent Marketing Practices," *Journal of Marketing Research*, 16 (February), 6-17.
- (1981), "Construct Validity: A Review of Basic Issues and Marketing Practices," *Journal of Marketing Research*, 18 (May), 133-45.
- Phillips, Lynn W. (1982), "Explaining Control Losses in Corporate Marketing Channels: An Organizational Analysis," *Journal of Marketing Research*, 19 (November), 525-49.
- Schmitt, Neal, Bryan W. Coyle, and Bruce B. Saari (1977), "A Review and Critique of Analyses of Multitrait-Multimethod Matrices," *Multivariate Behavioral Research*, 12 (October), 447-78.
- Seymour, Daniel and Greg Lessne (1984), "Spousal Conflict Arousal: Scale Development," *Journal of Consumer Research*, 11 (December), 810-21.
- Turner, Carol J. (1981), "The Multitrait-Multimethod Matrix: 1967-1980," *JSAS Catalog of Selected Documents in Psychology*, Ms. 2280.
- Widaman, Keith F. (1985), "Hierarchically Nested Covariance Structure Models for Multitrait-Multimethod Data," *Applied Psychological Measurement*, 9 (March), 1-26.
- Wilkie, William L. and Edgar A. Pessemier (1973), "Issues in Marketing's Use of Multi-Attribute Attitude Models," *Journal of Marketing Research*, 10 (November), 428-41.

Copyright of *Journal of Marketing Research* (JMR) is the property of American Marketing Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.