DOCUMENT RESUME

ED 298 144 TM 012 230

AUTHOR Marsh, Herbert W.
TITLE Confirmatory Factor Analyses of Multitrait-Multimethod Data: Many Problems and a Few Solutions.
PUB DATE 22 Jul 88
NOTE 42p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Construct Validity; *Factor Analysis; *Multitrait Multimethod Techniques; Research Methodology; *Research Problems
IDENTIFIERS *Confirmatory Factor Analysis; Parameter Identification

ABSTRACT

During the last 15 years, there has been a steady increase in the popularity and sophistication of the confirmatory factor analysis approach to multitrait-multimethod (MTMM) data. However, important problems exist, the most serious being the ill-defined solutions that plague MTMM studies and the assumption that so-called method factors primarily reflect the influence of method effects. In three different MTMM studies--by T. M. Ostrom (1969), B. M. Byrne and R. J. Shavelson (1986), and H. W. Marsh and R. Ireland (1984)--ill-defined solutions were frequent and alternative parameterizations designed to solve this problem tended to mask the symptoms instead of eliminating the problem. More importantly, so-called method factors apparently represented trait variance in addition to, or instead of, method variance for at least some models in all three studies. Further support for this counter interpretation of method factors was found when external validity criteria were added to the MTMM models and correlated with the trait and so-called method factors. This problem invalidates the traditional interpretation of trait and method factors and the comparison of different MTMM models. A new specification of method effects as correlated uniqueness instead of method factors was found to be less prone to ill-defined solutions and, apparently, to the confounding of trait and method effects. Two tables conclude the document. (Author/TJH)

Confirmatory Factor Analyses of Multitrait-multimethod Data: Many Problems and a Few Solutions

Herbert W. Marsh
University of Sydney, Australia

Revised: 22 July, 1988

Running Head: Multitrait-multimethod data

The author would like to thank Barbara Byrne and Richard Shavelson for providing data used in the present investigation, and to thank Michael Browne, Dennis Hocevar, David Kenny, Jack McArdle, Keith Widaman, and Werner Wothke for comments on earlier drafts of this paper. The contents, however, represent the views of the author. Correspondence regarding this manuscript should be sent to Herbert W. Marsh, Faculty of Education, University of Sydney, NSW, 2006, Australia.

Confirmatory Factor Analyses of Multitrait-multimethod Data:
Many Problems and a Few Solutions

ABSTRACT

During the last 15 years there has been a steady increase in the popularity and sophistication of the confirmatory factor analysis (CFA) approach to multitrait-multimethod (MTMM) data. There exist, however, important problems, the most serious being the ill-defined solutions that plague MTMM studies and the assumption that so-called method factors reflect primarily the influence of method effects. In three different MTMM studies ill-defined solutions were frequent and alternative parameterizations designed to solve this problem tended to mask the symptoms instead of eliminating the problem. More importantly so-called method factors apparently represented trait variance in addition to, or instead of, method variance for at least some models in all three studies. Further support for this counter interpretation of method factors was found when external validity criteria were added to the MTMM models and correlated with the trait and so-called method factors. This problem, when it exists, invalidates the traditional interpretation of trait and method factors and the comparison of different MTMM models. A new specification of method effects as correlated uniquenesses instead of method factors was less prone to ill-defined solutions and, apparently, to the confounding of trait and method effects.

The purpose of this investigation is to demonstrate and critically evaluate recently developed applications of confirmatory factor analysis (CFA) to multitrait-multimethod (MTMM) data. Campbell and Fiske (1959) argued that construct validation requires multiple indicators of the same construct to be substantially correlated with each other but substantially less correlated with indicators of different constructs. They proposed the MTMM design in which each of a set of multiple traits is assessed with each of a set of multiple methods of assessment, and developed four guidelines for evaluating MTMM data. Their MTMM design has become, perhaps, the most frequently employed construct validation design, and their original guidelines continue to be the most frequently used guidelines for examining MTMM data. Important problems with their guidelines are, however, well known (e.g., Althauser & Heberlein, 1970; Alwin, 1974; Campbell & O'Connell, 1967; Marsh, in press; Wothke, 1984) and have led to many alternative analytic approaches (e.g., Browne, 1984; Hubert & Baker, 1978; Jackson, 1969; 1977; Marsh, in press; Marsh & Hocevar, 1983; Schmitt, Coyle, & Saari, 1977; Schmitt & Stults, 1986; Stanley, 1961; Wothke, 1984). Factor analytic approaches (e.g., Boruch & Wollins, 1970; Joreskog, 1974; Marsh, in press; Marsh & Hocevar, 1983; Widaman, 1985) or mathematically similar path-analytic approaches (e.g., Werts & Linn, 1970; Schmitt, Coyle & Saari, 1977) currently appear to be the most popular approach and will be the focus of the present investigation.

A GENERAL MTMM MODEL AND A TAXONOMY OF ALTERNATIVE MODELS

The General MTMM Model.

In the CFA approach to MTMM data a priori factors defined by different measures of the same trait support the construct validity of the measures but a priori factors defined by different traits measured with the same method argue for method effects. For purposes of the present investigation I will emphasize a general MTMM model (Table 1) adapted from Joreskog (1974; also see Marsh & Hocevar, 1983; Widaman, 1985) in which: a) there are at least 3 traits (T=3) and 3 methods (M=3); b) T x M measured variables are used to infer T + M a priori common factors; c) each measured variable loads on the one trait factor and the one method factor that it represents but is constrained so as not to load on any other factors; d) correlations among the trait factors and among the method factors are freely estimated, but correlations between trait and method factors are constrained to be zero. For this model I assume there are at least three traits and three methods, but alternatives have been proposed for studies with only two methods (Kenny, 1979) or only two traits (Marsh & Hocevar, 1983). While some researchers have estimated correlations between trait and method factors there are important logical, interpretive, and pragmatic reasons for fixing these correlations to be zero (see Jackson, 1974; Marsh & Hocevar, 1983; Widaman, 1985). This constraint allows the decomposition of variance into additive trait, method, and error components, and without this constraint the solution is almost always empirically underidentified (also see Widaman, 1985; Wothke, 1984). In justifying this constraint, Joreskog (1971, p. 128) noted that: "This is our way of defining each method factor to be independent of the particular traits that the method is used to measure. In other words, method factors are what is left over after all trait factors have been eliminated."

Insert Table 1 About Here

In the present investigation CFA models were fit with LISREL V (Joreskog & Sorbom, 1981) and three design matrices from LISREL were used to define all the MTMM models. For T=3 traits and M=3 methods (see Table 1) the three design matrices are: a) Lambda Y, a 9 (M x T = number of measured variables) by 6 (M + T = number of factors) matrix of factor loadings; Psi, a 6 (M + T = number of factors) x 6 factor variance-covariance matrix of relations among the factors; and c) Theta, a 9 (M x T = number of measured variables) x 9 matrix of error/uniquenesses in which the diagonal values are analogous to one minus the communality estimates in exploratory factor analyses. All parameters (Table 1) with values of 0 or 1 are fixed and values for other parameters are estimated so as to maximize goodness of fit. Standard errors are estimated for all estimated parameters but not for parameters with fixed values. This model is easily modified to accommodate more traits or methods, to conform to other models and other parameterizations that will be described, or to incorporate unique factors for the measured variables (Rindskopf, 1983).

A Taxonomy of Alternative Models

Researchers have proposed many variations of the general MTMM model to examine inferences about trait or method variance or to test substantive issues specific to a particular study (e.g., Bagozzi, 1978; Joreskog, 1974; Marsh, Barnes & Hocevar, 1985; Marsh & Hocevar, 1983; Widaman, 1985; Wothke, 1984). Widaman (1985) proposed an important taxonomy of such models that systematically varied different characteristics of the trait and method factors. This taxonomy was designed to be appropriate for all MTMM studies, to provide a general framework for making inferences about the effects of trait and method factors, and to objectify the complicated task of formulating models and representing the MTMM data. One purpose of the present investigation is to evaluate the taxonomy in relation to these goals and to describe an expansion of the taxonomy formulated for the present investigation. The expanded taxonomy (see Table 2) represents all possible combinations of 4 trait structures (trait structures 1 - 4) and 5 method structures (method structures A - E). The 4 trait structures posit no trait factors (1), one general trait factor defined by all measured variables (2), T uncorrelated trait factors (3), and T correlated trait factors (4). The 5 method structures posit no method factors (A), one general method factor defined by all measured variables (B), M uncorrelated method variables (C), M correlated method factors (D), and method effects inferred on the basis of correlated uniqueness (E). This taxonomy differs from Widaman's original taxonomy only in the addition of Method structure E.

Insert Table 2 About Here

The general factors posited in method structure B and trait structure 2 may present interpretive or estimation problems. Widaman (1985) avoided some problems by constraining each general factor to be uncorrelated with all other factors and this constraint is used here. The rationale for this constraint is consistent with the requirement that trait and method factors be uncorrelated. Models 1B and 2A are, however, are the same, whereas Model 2B requires one additional, perhaps arbitrary, zero constraint to assure rotational identification. Finally, even for models that contain a general method factor in combination with T trait factors, or a general trait factor in combination with M method factors, the interpretation of the general factor may be problematic.

POTENTIAL PROBLEMS IN THE ESTIMATION AND INTERPRETATION OF MTMM MODELS

Goodness Of Fit

An important, unresolved problem in CFA is the assessment of goodness of fit. To the extent that a hypothesized model is identified and is able to fit the observed data, there is support for the model. The problem of goodness of fit is how to decide whether the predicted and observed results are sufficiently alike to warrant support of a model. Whereas X^2 values can be used to test whether these differences are statistically significant, there is a growing recognition of the inappropriateness of this classical hypothesis testing approach. Because hypothesized models are only designed to approximate reality, all such restrictive models are a priori false and will be shown to be false with a sufficiently large sample size (Cudeck & Browne, 1983; Marsh, Balla & McDonald, in press; McDonald, 1985). Hence, a variety of fit indices have been derived to aid in this decision process such as the X^2/df ratio and the Tucker-Lewis index (TLI; Tucker & Lewis, 1973) that are used here. In simulation studies of more than 30 such indices Marsh, Balla and McDonald (1988) and Marsh, McDonald and Balla (1987) found that both the X^2/df and TLI indices imposed apparently appropriate penalty functions for the inclusion of additional parameters that controlled for capitalizing on chance, whereas the TLI was the only widely used index that was also relatively independent of sample size. The TLI is emphasized in subsequent discussions, but values for other fit indices like the Bentler and Bonett's (1980) index can easily be computed from the results.

Model selection must be based on subjective evaluation of substantive issues, inspection of parameter values, model parsimony, and a comparison of the performances of competing models as well as goodness of fit. In the application of CFA to MTMM data there is an unfortunate tendency to under-emphasize the examination of parameter estimates and to over-emphasize goodness of fit. If a solution is ill-defined, then further interpretations must be made very cautiously if at all. If the parameter estimates for a model make no sense in relation to the substantive, a priori model, then fit may be irrelevant.

As described by Bentler and Bonett (1980), when two models are nested the statistical significance of the difference in the X^2s can be tested relative to the difference in their df. Widaman (1985) emphasized this feature in developing his taxonomy of MTMM models and in comparing the fit of different models. However, the problems associated with the application of the classical hypothesis testing approach also apply to this test of X^2 differences. When the sample size is sufficiently large the saturated model (i.e., a model with df = 0) will perform significantly better than any restricted model (see Cudeck & Browne, 1983) such as those in Table 2, thus making problematic the interpretation of tests between any two restricted models. Furthermore, many important comparisons are not nested and so cannot be made with this procedure (e.g., the trait-only (4A) and method-only (1D) models in Table 2). Because of these problems with the X^2 difference test, a perhaps more useful test is simply to compare the TLIs for competing models.

Insert Table 2 About Here

Poorly Defined Solutions

Poorly defined solutions represent a serious, unresolved problem for CFA that is particularly prevalent in MTMM studies. Poorly defined solutions refer to underidentified or empirically underidentified models (Kenny, 1979; Wothke, 1984), failures in the convergence of the iterative procedure used to estimate parameters, parameter estimates that are outside their permissible range of values (e.g., negative variance estimates called Heywood cases), or standard errors of parameter estimates that are excessively large. Each of these problems is an indication that the empirical solution is poorly defined, even if the model is apparently identified otherwise and even if goodness of fit is adequate (Joreskog & Sorbom, 1981). Such problems are apparently more likely when: the sample size is small; there are few indicators of each latent factor; measured variables are allowed to load on more than one factor; measured variables are highly correlated; there is a lot of missing data and covariance matrices are estimated with pairwise deletion for missing data; and the model is misspecified. Knowingly or unknowingly such problems are usually ignored, and the implications of this practice have not been explored for MTMM studies. Although there is no generally appropriate resolution for such problems, alternative parameterizations of the MTMM model (see below) may eliminate some improper parameter estimates.

There are apparent ambiguities about the identification status of MTMM models. Some researchers (e.g., Alwin, 1974; Browne, 1984; Joreskog, 1974; Schmitt, 1978) suggest that models with correlations between traits and methods are permissible, and Long (1983, p. 55) claimed to prove the identification for this model. However, Bollen and Joreskog (1985) demonstrated that the criteria used by Long were not sufficient to demonstrate identification, and Widaman (1985, p. 7) explicitly eliminated such models from his taxonomy, claiming that they "are very likely not identified."

In order to test the identification status of a model with correlated traits and methods David Kenny (personal communication, 23 January, 1987) used simulated data "to see if LISREL could recover loadings for your model 4D with traits and methods correlated. It did so, but not exactly. It was not clear whether the difference was due to under-identification or rounding error." I also attempted to fit model 4D with correlations between method and trait factors to the simulated population covariance matrix published by Cole and Maxwell (1985) in which the population correlations between trait and method factors were simulated to be zero. Whereas I was able to recover the population values, it took more than 500 iterations. For their sample matrices that included random error, however, the solutions failed to converge after more than 1000 iterations. It appears that whereas the model with correlated trait and method factors may technically be identified, it is unlikely to result in a proper solution for actual data so that it is of little practical use. Because models in Table 2 do not posit traits/method correlations and because all studies considered here have at least 3 traits and 3 methods, these ambiguities will not be examined here but they illustrate that the issue of identification has not been resolved.

Different Parameterizations: Potential Cures For Poorly Defined Solutions

The standard parameterizations. In order for the models in Table 2 to be identified, one parameter for each latent factor must be fixed -- typically at a value of 1.0 (see Joreskog & Sorbom, 1981; Long, 1983). This is usually done by: (a) fixing the factor loading of one measured variable for each latent factor to be 1.0 and estimating the factor variance, or (b) fixing the factor variance of each latent factor to be 1.0 (so that the factor variance/covariance matrix is a correlation matrix) and estimating all the factor loadings. For purposes of the present investigation these will be called the fixed factor loading and fixed factor variance parameterizations, and collectively they will be referred to as the standard parameterizations. So long as the CFA solution is well defined both parameterizations are equivalent, but fixing the factor variances introduces an implicit inequality constraint that restricts the factor variances to be nonnegative. Thus, fixing factor variance estimates may lead to a proper solution when fixing factor loadings does not.

The Rindskopf parameterization. Rindskopf (1983) proposed a solution for negative uniqueness estimates by using M x T additional factors -- one unique factor for each of the M x T measured variables -- to define each uniquenesses. Because the factor loading on each unique factor is the square root of the uniqueness, the uniqueness is implicitly constrained to be non-negative. Joreskog (1981), commenting on the merits of imposing inequality constraints, noted that if a solution is inadmissible, then LISREL will find a solution outside the permissible parameter space whereas the imposition of inequality constraints will produce a solution on the boundary of the parameter space. Joreskog (p. 91) concluded: "In both cases the conclusion will be that the model is wrong or that the sample size is too small." Similarly, Dillon, Kumar and Mulani (1987) noted that in their research the Rindskopf parameterization always resulted in the offending parameter estimate taking on a zero value that resulted in the same solution as simply fixing the parameter to be zero.

Method structure E -- an alternative conceptualization of method variance. Method variance is an undesirable source of systematic variance that distorts correlations between different traits measured with the same method. As typically depicted in MTMM models (i.e., method structures C and D) a single method factor is used to represent the method effect associated with variables assessed by the same method. The effects of a particular method of assessment are implicitly assumed to be unidimensional and the sizes of the method factor loadings provide an estimate of its influence on each measured variable. Hence, method structures C and D restrict method covariance components to have a congeneric-like structure (but see Wothke, 1984). Alternatively, method effects can be represented as correlated uniquenesses (method structure E) and this representation does not assume either the unidimensionality of effects associated with a particular method or a congeneric structure. Kenny (1979; also see Marsh & Hocevar, 1983; Marsh, in press) proposed this method structure for the special case in which there are only two traits, but it is also reasonable when there are more than two traits. Method structure E also resembles McDonald's multi-mode analysis (1985) and Browne's multiple battery analysis (1980).

Method structure E corresponds most closely to method structure C (Table 2) in that the method effects associated with one method are assumed to be uncorrelated with those of other methods. When there are 3 traits and the solutions are well-defined, method structures C and E are merely alternative parameterizations of the same model. When T > 3, however, the number of correlated uniquenesses in method structure E (M x (T x (T-1)/2) is greater than the number of factor loadings used to define method factors in method structure C (T x M). Thus method structure C is a special case of method structure E in which each method factor is required to be unidimensional and this assumption is testable when T > 3.

A particularly important advantage of method structure E is that it apparently eliminates some improper solutions without limiting the solution space or forcing parameter estimates to the boundaries of the permissible space. Because method variance is one source of uniqueness, uniqueness is reflected in both method factors and error/uniquenesses. Improper solutions are frequently due to either negative method factor variances or negative error/uniquenesses, but not both. In method structure E all sources of uniqueness are contained in the diagonal of Theta, and in many cases -- as demonstrated in the present investigation -- this combined influence will not be negative even when method factor variances or uniquenesses are negative for other parameterizations. Thus, even when there are 3 traits so that method structures C and E are equivalent when the solutions are well defined, it is possible that method structure C will results in poorly defined solutions whereas method structure E will not. When there are more than three traits it is possible for method structure E to fit the data better than either method structures C or D, thus calling into question the assumed unidimensionality of method effects in structures C and D.

Problems in the Interpretation of Trait and Method Factors

Widaman's taxonomy and the MTMM models in Table 2 implicitly assume that:

a) method factors represent method variance, b) trait factors represent trait variance, c) a general factor in combination with trait factors represents method variance, and d) a general factor in combination with method factors represents trait variance. For present purposes these assumptions will be referred to as the traditional interpretation of the MTMM models. These assumptions are probably reasonable when correlations among the trait factors and among the method factors are small, but this situation is unusual. These assumptions may not be reasonable when correlations among trait factors and correlations among method factors are substantial. For present purposes I will examine the possibility that so-called method factors actually reflect trait variance, but the problem might also apply to so-called trait factors that actually reflect method variance.

In most MTMM studies the multiple traits are correlated and this may produce a general trait factor that makes ambiguous the interpretation of so-called general method factors or even correlated method factors. When traits are substantially correlated, the so-called general method factor (method structure B) may represent trait variance instead of, or in addition to, method variance. When traits are substantially correlated, each so-called correlated method factors (method structure D) may represent this general trait factor and correlations among the method factors may represent the convergence of this general trait across the methods of assessment. If this problem exists, the traditional interpretation of MTMM models and the comparison of alternative models is unjustified. Hence, tests of this plausible counter interpretation of method factors must be examined.

Results to be discussed here suggest that the traditional interpretation of method factors may be unjustified if: (a) interpretations based on the Campbell-Fiske guidelines and an examination of the MTMM matrix differ substantially from those based on the CFA approach (there are, of course, problems with the Campbell-Fiske approach, but if both the Campbell-Fiske and the CFA approaches lead to consistent conclusions then confidence in these conclusions is increased); (b) substantive theory dictates an expected pattern of correlations among trait factors that is not supported; (c) the substantive nature of the data dictates an expected pattern of correlations among method factor that is not supported (though a priori hypotheses of relations among method factors may be difficult to formulate); (d) Model 4A (trait factors only) and 1D (method factors only) both fit the data reasonably well and Model 4D provides only a modest improvement; (e) the amount of variance explained by trait factors is substantially reduced by the inclusion of method factors; (f) external validity criteria collected in addition to the MTMM variables are more substantially correlated with so-called method factors than with trait factors and there is an a priori basis for assuming the external criteria to be more strongly related to trait factors than method factors (it may be impossible to obtain external validity criteria that are free of all method effects so that the aim is to ensure that any method effects associated with the external validity criteria are unrelated to those associated with the MTMM data; there is still a danger that, unknown to the researcher, the external validity criteria are affected by the same method effects as the original MTMM variables). Whereas each of these indications of potential problems with the interpretation of method effects is fallible, taken together they provide a stronger basis for evaluating these interpretations than do typical application of the CFA approach. They also require that more emphasis be placed on the substantive interpretation of results than is typical in the CFA approach to MTMM data.

APPLICATION OF THE CFA APPROACH IN THREE MTMM STUDIES

The purposes of the present investigation are to evaluate: a) the application of the MTMM taxonomy (Table 2), (b) the problems of poorly defined solutions and parameterizations designed to eliminate them, (c) the merits of method structure E, and (d) the validity of traditional interpretations of trait and method factors. Data come from three MTMM studies: Ostrom (1969), Byrne and Shavelson (1986), and Marsh and Ireland (1984). For all three of these studies there were at least three traits and three methods, and there was at least one external validity criterion in addition to the MTMM data. In the present analysis of each of the studies: a) models in the taxonomy were fit to just the MTMM data; b) the behavior of the solutions was examined; c) the substantive nature of the data and the parameter estimates were used to evaluate alternative interpretations of the method and trait factors; and d) external validity criteria were added to the MTMM models in order to test alternative interpretations of the trait and method factors.

The Ostrom (1969) Study

Description of the Study and Data.

Ostrom (1969) examined the distinction between affective, behavioral, and cognitive components (T1-T3) of attitudes toward the church assessed with four different methods of scale construction (M1-M4). Ostrom also collected additional "overt behavioral indices" and hypothesized that these should be most highly correlated with the behavioral trait component. For purposes of the present investigation one of these, responses to the item "How many days out of the year do you attend church services" was used. Ostrom presented the correlations based on responses by 189 subjects as well as a more detailed account of the theoretical rationale, the 12 MTMM variables, and the external validity criterion. My application of the Campbell-Fiske guidelines suggested strong support for convergent validity. However, support for discriminant validity was problematic and there appeared to be method variance associated with at least M2 and M4.[1] The substantive nature of the data indicates that the traits should be substantially correlated, but there is no a priori basis for positing the relative size of these different correlations. Finally, for models in which the external validity criterion was added, the criterion should be: a) more correlated with specific and general trait factors than with specific and general method factors; and b) most highly correlated with the behavioral trait component.

CFA models similar to those considered here have been applied to this data by Bagozzi (1978), Schmitt (1978), and Widaman (1985). Schmitt (1978) excluded one of the methods and estimated trait/method correlations, and so his results are not comparable. Bagozzi (1978) fit Model 4A to the 12 variables considered here, but an inspection of correlations between the error/uniquenesses led him to eliminate one of the methods from subsequent analyses. It should be noted that such correlated uniquenesses are indicative of a method effect as depicted in method structure E. Widaman (1985) also noted this apparent misinterpretation of method effects and was critical of other conclusions by Bagozzi. Widaman (1985) fit many of the models used here and chose to represent the MTMM data with Model 4D. However, his solution for Model 4D was poorly defined in that an error/uniqueness was estimated to be zero and had a large standard error[2]. None of these previous CFAs of the Ostrom data incorporated the external validity criterion included here.

<u>Behavior of the Solutions For Different Parameterizations.</u>

All models in Table 2 were tested with both the fixed factor loading and the fixed factor variance parameterizations, and the Rindskopf parameterization was used when both standard parameterizations produced poorly defined solutions (Table 3). For the fixed factor loading parameterization, 7 of the 19 models were poorly defined as indicated by a failure to converge or improper solutions. For the fixed factor variance parameterization, 5 of these 7 models were still poorly defined but the problem symptoms were not always the same. When these five models were tested with the Rindskopf parameterization, one solution was improper and the remaining four had error/uniquenesses estimates close to zero with extremely large standard errors. In Model 1D there were factor correlations greater than 1.0 for all three parameterizations, demonstrating that none of the parameterizations protect against this type of improper solution. Whereas the different parameterizations varied in their behavior and manifest symptoms, none eliminated the poorly defined solutions.

Method Structure E

In method structure E correlated uniquenesses are used to represent method effects, and in support of this structure all four solutions based on it are well defined. When there are 3 traits, method structure E is equivalent to method structure C so long as the models are well defined. For the Ostrom data this was demonstrated for Models 1C and 1E, but Models 2C, 3C, and 4C were poorly defined for all three parameterizations. Even though Model 2C failed to converge for either of the standard parameterizations, the parameter estimates for the trait factors and overall fit were nearly the same as for Model 2E. Model 3C converged to an improper solution for the fixed factor variance parameterization, but the parameter estimates for trait factors and overall fit were the same as for Model 3E. Model 4C converged to an improper solution for the fixed factor loading parameterization but parameter estimates for trait factors and overall fit were the same as for Model 4E. The Rindskopf parameterization eliminated improper solutions for Models 2C, 3C, and 4C, but resulted in error/uniqueness estimates of zero with large standard errors. These findings suggest that method structure E is a better representation of method effects than method structure C.

The substantive interpretation of trait and method factors

Interpretations based on just the MTMM Data. Model 4D provides an exceptionally good fit to the data, but there are problems with the solution. First, it is poorly defined for all three parameterizations. Second, the trait factors are very weak in that 7 of the 12 factor loadings are not statistically significant, and this contradicts conclusions based on the Campbell-Fiske guidelines. Models 4C, 4E, and 3D also fit the data very well (TLIs > .98), and Models 4A, 1D and even 2A/1B explain most of the variance (TLIs > .9). In contrast to Model 4D, Models 4E, 4C, and 4A have strong trait factors for which all factor loadings are significant. As noted by Widaman it may be problematic to compare trait and method variance for this data because most of the variance can be explained by either trait or method factors and neither trait nor method factors uniquely explain much variance. Also, since trait factor loadings are so much lower when correlated method factors are included, these so-called method factors may reflect trait variance.

Widaman (1985) chose Model 4D to represent this MTMM data on the basis of fit. However, Model 4E (Table 4) also provides a good fit and has important advantages over Model 4D. First, it is well defined whereas Model 4D is not. Second, the strong trait factors in Model 4E more accurately reflect more parsimonious Model 4E may be preferable.

Summary of Analyses of Ostrom data.

None of the three parameterizations eliminated problems of poorly defined solutions for Ostrom data. The fixed factor loading parameterization was most prone to improper solutions. The Rindskopf parameterization was more likely to converge to proper solutions, but only at the expense of error/uniqueness estimates of zero with extremely large standard errors. In contrast solutions for method structure E were always well defined, suggesting that it might be a more appropriate formulation of method effects. For the Ostrom data most of the variance can be explained in terms of either method factors or trait factors, whereas the inclusion of both trait and method factors produced only a small improvement in fit. Since relatively little variance was uniquely due to either trait or method factors, any conclusions about their relative importance are problematic. Even more serious problems exist in the interpretation of the correlated method factors. These so-called method factors were more substantially correlated with an external validity criterion than were the trait factors, and apparently reflect trait variance instead of, or in addition to, method effects. The solution for Model 4E apparently provides a better representation of the MTMM data than the solution for Model 4D selected by Widaman even though the fit of Model 4D is slightly better. The assumption of uncorrelated method effects in Model 4E is worrisome, but the traditional interpretation of the method factors in Model 4D is clearly unjustified and undermines any comparisons between it and other models. This illustrates the problems associated with using fit as the primary basis for selecting between alternative models instead of substantive interpretations of the parameter estimates.

Byrne and Shavelson (1986) Study.

Description of the Study and Data.

Byrne and Shavelson (1986) examined the relations between three academic self-concept traits (Math, Verbal, and School self-concepts) measured by three different self-concept instruments (M1-M3). School performance measures were also available for English and mathematics. Marsh and Shavelson (1985) reported Math and Verbal self-concepts to be nearly uncorrelated with each other even though both were substantially correlated with School self-concept. They posited two higher-order academic facets -- verbal/academic and math/academic self-concept -- to explain specific facets of academic self-concept. Their research posits a specific pattern of correlations among the trait factors and suggests that two general trait factors may provide a reasonable fit to the Byrne and Shavelson data. For the expanded MTMM models more parsimonious Model 4E may be preferable.

Summary of Analyses of Ostrom data.

None of the three parameterizations eliminated problems of poorly defined solutions for Ostrom data. The fixed factor loading parameterization was most prone to improper solutions. The Rindskopf parameterization was more likely to converge to proper solutions, but only at the expense of error/uniqueness estimates of zero with extremely large standard errors. In contrast solutions for method structure E were always well defined, suggesting that it might be a more appropriate formulation of method effects. For the Ostrom data most of the variance can be explained in terms of either method factors or trait factors, whereas the inclusion of both trait and method factors produced only a small improvement in fit. Since relatively little variance was uniquely due to either trait or method factors, any conclusions about their relative importance are problematic. Even more serious problems exist in the interpretation of the correlated method factors. These so-called method factors were more substantially correlated with an external validity criterion than were the trait factors, and apparently reflect trait variance instead of, or in addition to, method effects. The solution for Model 4E apparently provides a better representation of the MTMM data than the solution for Model 4D selected by Widaman even though the fit of Model 4D is slightly better. The assumption of uncorrelated method effects in Model 4E is worrisome, but the traditional interpretation of the method factors in Model 4D is clearly unjustified and undermines any comparisons between it and other models. This illustrates the problems associated with using fit as the primary basis for selecting between alternative models instead of substantive interpretations of the parameter estimates.

Byrne and Shavelson (1986) Study.

Description of the Study and Data.

Byrne and Shavelson (1986) examined the relations between three academic self-concept traits (Math, Verbal, and School self-concepts) measured by three different self-concept instruments (M1-M3). School performance measures were also available for English and mathematics. Marsh and Shavelson (1985) reported Math and Verbal self-concepts to be nearly uncorrelated with each other even though both were substantially correlated with School self-concept. They posited two higher-order academic facets -- verbal/academic and math/academic self-concept -- to explain specific facets of academic self-concept. Their research posits a specific pattern of correlations among the trait factors and suggests that two general trait factors may provide a reasonable fit to the Byrne and Shavelson data. For the expanded MTMM models that include validity factors, each validity factor should be substantially correlated with the trait factors, particularly the trait factor in the matching content area and, to a less extent, the school factor, and relatively uncorrelated with the method factors. The Byrne and Shavelson study is unusual because there is just a good a priori basis for predicting the structure of the trait factors and also because two of the trait factors are relatively uncorrelated.

My application of the Campbell-Fiske guidelines to the MTMM matrix (Marsh, in press) suggested strong support for convergent and discriminant validity; every convergent validity was substantial, was larger than every heterotrait-heteromethod coefficient, and was larger than nearly every heterotrait-monomethod coefficient. For all three instruments School self-concept was moderately correlated with Math and Verbal self-concepts whereas Math and Verbal self-concepts were nearly uncorrelated with each other. There was evidence of some method effects associated with at least M3 and, perhaps, M2. The Byrne and Shavelson study is an exemplary MTMM study because of the clear support for the Campbell-Fiske guidelines, because of the large sample size (817, after case-wise deletion for missing data), because of the good psychometric properties of the measures, and because of the a priori knowledge of the trait factor structure. All models in Table 2 were fit using the fixed factor loading parameterization, the fixed factor variance parameterization was used for models that were poorly defined with the first parameterization, and the Rindskopf parameterization was used for solutions that were poorly defined by both standard parameterizations.

Insert Table 6 About Here

Behavior of the Solutions Under Different Parameterizations.

Nearly half of the solutions, 9 of 19, were poorly defined for the fixed factor loading parameterization (Table 6); 4 solutions were improper, and 5 of the solutions failed to converge. When these 9 poorly defined solutions were tested with the fixed factor variance parameterization, 2 of the solutions were well defined but the remaining 7 were still poorly defined. For the Rindskopf parameterization only 1 of the 7 problem solutions was improper but the all offending estimates were approximately zero and had large standard errors in the other 6 solutions.

1. All four solutions for method structure E were well defined whereas all four corresponding models for method structure C were poorly defined. When the four solutions for method structure C converged to improper solutions with either of the standardized parameterizations, the parameter estimates for trait factors and the overall fit were the same as for the corresponding

solution for method structure E. For the Rindskopf parameterization, the parameter estimates varied somewhat, the fit was always somewhat poorer, and some parameters had values close to zero with large standard errors. For this application method structure E provides a better representation of method effects than method structure C.

2. Model 1D (correlated methods) produced the same improper solution, factor correlations greater than 1.0, for all three parameterizations. Because none of these parameterizations constrain factor correlations to be less than 1.0, they provide no protection from this problem.

3. Model 3D converged to an improper solution with the fixed factor loading parameterization. Whereas the solution was constrained to be proper by the Rindskopf parameterization, the X^2 was approximately twice as large.

4. For Model 4D the fixed factor loading parameterization resulted in an improper solution whereas the fixed factor variance parameterization produced a well defined solution. The fit of the fixed factor variance solution was somewhat poorer indicating that it apparently imposed a limitation on the solution space.

Substantive Interpretations of Trait and Method Factors

Interpretations based on the MTMM data. Models positing only method factors fit the data poorly. Model 2B with one so called general method factor and one general trait factor did substantially better. However, partly due to the way Model 2B was specified, these two general factors represent the math/academic and verbal/academic factors that were originally posited. The interpretation of either of these as a general method factor is unjustified.

Model 2D (3 correlated M factors and 1 general T factor) provided a reasonable fit to the data, but inspection of the parameter estimates demonstrated interpretational problems. The three School measures should have loaded substantially on the general trait factor, but all three loadings were small (.10 to .22). Loadings for Verbal and Math scales were larger, but in the opposite direction, suggesting that the so-called general trait factor represented a bipolar (verbal vs. math self-concept) factor. For each of the method factors all loadings were substantial and positive, and the three method factors were substantially intercorrelated (.80 to .97). However, these so-called method factors and the high correlations among them seem to represent convergence on general trait factors associated with each of the self-concept instruments. The interpretation of these factors is speculative, though substantively interesting, but the traditional interpretation of the factors is clearly unjustified.

Insert Table 7 About Here

Model 4D, particularly given the large sample size, provides a remarkably good fit to the data (TLI=.99). Parameter estimates (Table 6) indicate that each of the trait-factors is well defined. Consistent with theory and previous research, the School trait factor is substantially correlated with the Verbal and Math factors whereas the Verbal and Math factors are nearly uncorrelated with each other. These trait factors are stronger than the method factors in that all nine measured variables have higher trait factor loadings than method factor loadings, but models without any method factors (e.g., Model 4A) provide a poorer fit to the data.

Models 4C and 4E also provide good fits to the data (TLI=.977). For Models 4C, 4D and 4E the factor loadings are similar for the School and Math traits, but the Verbal trait-factor loadings are smaller for Model 4D. Correlations among the trait factors are similar for all three models. Model 4D may be preferred because it fits the data slightly better, but there is no compelling basis for rejecting the more parsimonious model 4E and the two models lead to similar conclusions.

Relations between MTMM factors and external validity criterion. Two validity factors defined by achievements in English and mathematics were added to selected MTMM models (Table 8). In contrast to the external validity criterion used with the Ostrom data, method effects associated with the self-report measures are unlikely to be related to the achievement test scores. Support for the validity of the interpretation of the MTMM solutions requires each achievement factor to be most highly correlated with its matching trait factor, less correlated with the School trait factor, substantially less correlated with the non-matching trait factor, and relatively uncorrelated with the method factors. The validity of the method factors would also be supported if the hypothesized pattern of correlations between trait factors and validity factors is improved by the addition of method factors.

Five models contain three trait factors and two validity factors in combination with: correlated method factors (Model 4D); uncorrelated method factors (Model 4C); uncorrelated method effects represented as correlated errors (Model 4E); 1 General method factor (Model 4B); or no method factors (Model 4A). For each of these models, there is reasonable support for the predicted pattern of correlations between validity and trait factors. The inclusion of method factors improves support for this hypothesized relation in terms of Verbal self-concept, but has little effect on predictions in relation to Math self-concept or School self-concept. These results provide clear support for the a priori interpretation of the trait factors.

The comparison of Models 4C and 4E is informative. Model 4C is ill-defined but its TLI is much better than Model 4E. In Model 4C the validity factors are correlated with the method factors whereas in Model 4E no correlations were posited between validity factors and the error/uniquenesses that represent method effects. The results from both Models 4C and 4D suggest that the method factors are correlated with validity factors, and this apparently explains the poorer fit of Model 4E. The hypothesized pattern of relations between trait and validity factors is also stronger for Models 4C than Model 4E. However, when selected error/uniquenesses in Model 4E were correlated with the validity factors (Model 4E' in Table 8) support for the posited pattern of correlations and goodness of fit were similar to Model 4C.

Correlations between the trait and validity factors are similar in Models 4C, 4D and 4E' (Table 8). The only substantive difference is that English achievement is somewhat more highly correlated with School self-concept than Verbal self-concept in Model 4D, whereas English achievement is more highly correlated with Verbal self-concept than School self-concept for Models 4E' and 4C. Support for the posited pattern of relations is somewhat weaker for Model 4D even though its fit is best, Model 4C is poorly defined, and the correlations between method effects and validity factors are not easily represented in Model 4E'. There is no compelling reason for rejecting either Model 4D or 4E. In fact, the similar interpretations based on each of these different models suggests that the traditional interpretation of these models is probably justified.

Insert Table 8 About Here

Four models contain three method factors and two achievement factors in combination with: correlated trait factors (Model 4D discussed above), uncorrelated trait factors (Model 4C discussed above), 1 General trait factor (Model 2D), and no trait factors (Model 1D). Neither Models 2D nor 1D contain specific trait factors, and their method factors are substantially and positively correlated with the validity factors. In fact, English achievement is more highly correlated with M3 in Model 2D, and mathematics achievement is more highly correlated with M1 and M3 in Model 1D, than any trait factors in any of the other models. These results provide clear support for the earlier supposition that these so-called method factors contain substantial amounts of trait variance. In Model 2B there are just two general factors that might be interpreted to reflect a general method factor and a general trait factor. However, the correlations between these two general factors and the validity factors demonstrate that they reflect the math/academic and verbal/academic factors originally posited.

Summary of the Analyses of the Byrne and Shavelson Data.

The Byrne and Shavelson study is an exemplary MTMM study that should be well suited to the CFA approach. For this reason it is particularly disappointing that so many of the models from Table 2 resulted in poorly defined solutions. So long as both correlated trait and either correlated or uncorrelated method factors were included in the MTMM model, support for the traditional interpretation of MTMM factors appeared reasonable. However, for models with no trait factors, or only one general trait factor, the interpretation of method factors as representing method variance was clearly unjustified. Instead, the so-called method factors reflected the influence of trait variance. Further support for this counter-interpretation was provided by the substantial correlations between these so-called method factors and the validity factors.

The Marsh and Ireland (1984) Study.

Description of the Study and data.

Marsh and Ireland (1984) asked each of six teachers (M1-M6) to evaluate 139 student essays on six single-item scales of writing effectiveness: Mechanics, Sentence Structure, Word Usage, Organization, Content/ideas and Quality of style (T1-T6). Previous research reported a large general component of writing effectiveness suggesting that trait factors should be substantially correlated. There was no a priori hypothesis about the relative size of different trait correlations, but the traits were roughly ordered from lower-level components to higher-order components according to Foley's (1971) adaptation of the Bloom taxonomy. Unlike the first two MTMM studies, ratings for the different methods were not completed by the same person. The teachers did not know any of the students who had written the essays or, typically, each other, and each teacher performed the rating task independently of the others.[3] In addition to the 36 measured variables that constitute the MTMM data, a school performance measure of writing effectiveness was also available. My application of the Campbell-Fiske guidelines (Marsh & Ireland, 1984) suggested strong support for the convergent validity of all 6 traits. However, there was little support for discriminant validity and some indication of method effects associated with ratings by each of the six teachers.

Behavior of the Solutions.

Because the number of variables in this MTMM study was large only a subset of the MTMM models was tested with the fixed factor variance parameterization (Table 9). Nevertheless, 12 of these 13 models resulted in well defined solutions. Model 4B was improper in that several trait correlations exceeded 1.0, but this type of improper parameter estimate is unlikely to be eliminated by any of the parameterizations. Model 4D was technically improper in that the factor correlation matrix (see Table 9) was not positive definite even though none of the correlations were greater than 1.0. Despite these problems, the solutions for the Marsh and Ireland data appear to be better behaved than for either of the first two MTMM studies.

Insert Tables 9 and 10 About Here

Substantive Interpretation of Trait and Method Factors.

Interpretations based on the MTMM data. The goodness-of-fit statistics (Table 9) demonstrated that much of the variance can be explained by either six correlated method factors (Model 1D) or six correlated trait factors (Model 4A), but that Model 1D fit the data slightly better than 4A. For Model 4D method factor loadings were consistently much larger than trait factor loadings. A superficial inspection of these results might suggest that the ratings reflect primarily method effect, but there are problems with this interpretation. First, it contradicts conclusions based on the Campbell-Fiske guidelines. Second, trait-factor loadings were substantially smaller in Model 4D than 4A. This suggests that the so-called method factors may represent general trait factors associated with each teacher and the high correlations represent agreement across teachers on this general trait.

Correlations between MTMM factors and the validity factor. In order to test the counter interpretation of the method factors, the school performance measure was added to Model 4D. The parameters for the MTMM variables were relatively unaffected by the inclusion of this additional variable. However, the school performance factor was substantially more correlated with the so-called method factors (.56 to .68) than with the trait factors (.14 to .30). Because this pattern of results is so implausible, the traditional interpretation of the so-called method factors in this model must be rejected. Models 2D and 1D also posited correlated method factors and correlations between these method factors and the validity factor were also very high (.53 to .75), whereas the general trait factor in Model 2D was only modestly correlated with the validity factor. In contrast to models with correlated methods, Model 4C posited method factors to be uncorrelated. For Model 4C correlations between the trait factors and the validity factor varied between .71 and .89 whereas 5 of the 6 correlations between method factors and the validity factor were nonsignificant. Even though the fit for Model 4C was somewhat poorer than Model 4D, the substantive interpretation of the solution argues that it better reflected the MTMM data.

Method Structure E.

The first two MTMM studies both contained three trait factors, and for such MTMM studies method structure E results are equivalent to those for method structure C so long as the solutions are well defined. However, when there are more than 3 traits, as here, the two structures are not equivalent. For T = 6, method structure C uses 6 parameters to define each method factor, whereas there are 15 (T x (T-1)/2) correlations among the error/uniquenesses associated with each method. Insofar as method structures C and E are both well defined and fit the data equally well, then the more parsimonious method structure C is preferable. However, the fit of models based on method structure E was much better than those based on method structure C for this data. This suggests that the 15 correlated error/uniquenesses associated with each method effect cannot be explained by a single method factor and that method effects do not have a congeneric-like structure. This is very important in that all the method factors in the entire taxonomy are based on these assumptions. This also explains why Model 4E (TLI=.948) fits the data better than Model 4D (TLI=.935) even though Model 4E posits uncorrelated method effects whereas Model 4D posits correlated method factors.

The superiority of Model 4E over 4C is also shown in the expanded models containing the validity factor. Correlations between trait factors and the validity factor are substantial for both Models 4E and 4C, but are higher for Model 4E. As noted for the first two MTMM studies, correlations between method effects and the validity factor are not easily incorporated into Model 4E. However, correlations between validity and method factors were small and generally nonsignificant for Model 4C. Similarly, inspection of the modification indices provided by LISREL (see Joreskog & Sorbom, 1981) indicated that error/uniquenesses in Model 4E were essentially uncorrelated with the validity factor. For this reason no alternative model corresponding to Model E' in the first two studies was proposed.

SUMMARY AND IMPLICATIONS

Why does one use the kinds of analyses discussed here? One perception in accordance with my own[4] is that the motivation for MTMM analyses has been the desire to establish specific trait representations in measures. Method variance is seen as contaminating that representation. The CFA approach as traditionally applied has modelled trait and method factors as if they were equally important. The approach advocated here places greater emphasis on the interpretation of trait representations. This is accomplished by comparing different models to determine if the introduction of method factors substantially alters the interpretation of trait representations, by introducing an alternative method structure (method structure E) that apparently provides a more accurate representation of the trait representation, and by demonstrating how external validity criteria can be used to test the validity of the traditional interpretations of different models. Whereas the conventional approach was due at least in part to early work by Joreskog, the perspective taken here is consistent with Joreskog's statement that "method factors are what is left over after all trait factors have been eliminated" (1971, p. 128).

Despite the growing enthusiasm for the CFA approach to MTMM data, problems demonstrated here call into question its value, the traditional interpretation of MTMM factors, and the validity of previous MTMM research. The most important of these problems are the technical difficulties in estimating parameters and the interpretation of so-called method effects that apparently represent the effects of trait variance in addition to, or instead of, method variance. So long as problems as basic as these remain unresolved, the promise of the CFA approach to MTMM data cannot fulfilled.

The CFA approach to MTMM data is plagued by technical difficulties in the estimation of parameter values and different parameterizations of MTMM models were proposed to eliminate such problems. The fixed factor loading parameterization was apparently most prone to ill-defined solutions, whereas the Rindskopf parameterization was most likely to converge to proper solutions. However, when error/uniquenesses were negative for the standard parameterizations, the offending parameters were estimated to be close to zero (also see Dillon et al., 1987) with very large standard errors with the Rindskopf parameterization. Hence the Rindskopf parameterization did not solve the the problem, but merely made the manifest symptoms less obvious.

Method structures in Widaman's 1985 taxonomy and those used in most applications of CFA to MTMM data posit a separate method factor associated with each method of assessment. An alternative conceptualization, Method structure E, was formulated in which method effects are represented as correlated error/uniquenesses. Method structure E has three important advantages over method structures C and D. First, models with method structures C and D were frequently ill-defined no matter what parameterization was used, whereas models based on method structure E were always well defined in the present applications. Second, when there were more than three traits, method structure E provided a test of the implicit assumption that all the correlated error/uniquenesses associated with a single method of assessment could be explained in terms of a single method factor. The importance of this second advantage was demonstrated for the Marsh and Ireland data in that Model 4E provided a better fit than the corresponding Models 4C and 4D. Third, method structure E apparently provided a more accurate interpretation of trait variance than alternative models when these interpretations were evaluated in relation to external validity criterion. In this respect, the use of external validity criteria to validate interpretations of the method and trait effects is an important conceptual innovation.

The most serious potential problem with MTMM models is the implicit assumption that so-called method factors represent primarily the effects of method variance. If this assumption is violated, then the interpretation of trait and method factors in most CFA studies and the detailed comparison of nested models proposed by Widaman (1985) may be unjustified. Results from the MTMM studies considered here suggested that this assumption is often implausible. In all three MTMM studies the so-called method factors for at least some of the MTMM models apparently represented trait variance in addition to or instead of method variance (also see Marsh & Butler, 1984, for another compelling example). When there actually are distinct traits that are at least moderately correlated, this phenomenon is most likely in models that posit correlated method factors (method structure D). Using method structure D the problem is likely to be most severe in models that posit no trait factors (1D) and to become less severe as the trait structure proceeds from 1 to 4. The problem will apparently be least likely when method factors were required to be uncorrelated as in method structures C and E.

The emphasis of the present investigation has been on potential problems in the interpretation of so-called method factors that really reflect variance that should be attributed to a general trait effect. This is consistent with Joreskog's (1971) conceptualization of method effects as what is left after trait factors have been removed and my perspective on the the intent of MTMM analyses. It is important to note, however, that the converse phenomenon may also exist. That is, it is possible that so-called trait effects really reflect variance that should be attributed to a general method effect. If an appropriate method structure is not employed, then so-called trait factors may represent method variance in addition to, or instead of, trait variance. An unresolved conceptual and technical problem is how to discriminate between method and trait factors when both are highly correlated. In the extreme, it is easy to imagine the case where a MTMM matrix of correlations produced by highly correlated trait and method factors could be explained by a single factor (Model 1B/2A). Whereas this situation would clearly indicate a lack discriminant validity there would be little basis for determining whether the single factor represented a general trait effect, a general method effect, or a combination of the two.

The taxonomy of MTMM models in Table 2 was based in large part on Widaman's taxonomy. Widaman (1985) used essentially the same CFA approach, many of the same MTMM models, and even analyzed one of the same MTMM studies.

Because Widaman's evaluation of the CFA approach was much more optimistic than mine, it is informative to critically evaluate his findings in relation to the criteria used here. Widaman did not provide a detailed report of the behavior of his CFA solutions, but results reported here indicate that poorly defined solutions occurred for the Ostrom data considered in both studies. Widaman chose to present five MTMM solutions as the most appropriate representations of his MTMM analyses. However, four of these solutions had error/uniquenesses of zero in conjunction with large standard errors whereas the fifth solution required a correlation between two method factors to be 1.0. Wothke (1984) also reported that 21 MTMM matrices -- including the three analyzed by Widaman -- resulted in poorly defined solutions when he fit Model 4D. Apparently, none of the solutions chosen by Widaman was well defined according to criteria used here suggesting that Widaman was also plagued by poorly defined solutions. Widaman did not report a critical evaluation of alternative interpretations of his method factors, but results reported here suggest that this was a problem for the Ostrom data. Using criteria described earlier there is reason to suspect that so-called method factors in at least some of Widaman's results of other MTMM matrices may have also represented trait variance in addition to, or instead of, method variances. In summary, a critical evaluation of Widaman's results provides little basis for optimism about the application of CFA to MTMM data. His results suggest the same sort of problems that were identified here.

RECOMMENDATIONS

Problems with the CFA approach to MTMM data appear to be most serious for MTMM studies in which method effects are substantially correlated and for MTMM models that posit correlated method factors. Campbell and Fiske (1959) originally stressed that the multiple methods should be as distinct as possible, and this advice seems appropriate for the CFA studies as well. The choice of method effects is, however, often dictated by the nature of the study, and the pattern of correlations among method factors may be difficult to determine a priori. Particularly when both traits and methods are substantially correlated, the researcher must critically evaluate the MTMM solutions for alternative interpretations. Because the traditional interpretation of trait and method factors may frequently be unjustified, the burden of proof lies with the researchers to demonstrate that they are justified. This requires that more emphasis be given to the substantive interpretations than has typically been the case in CFA studies.

The use of uncorrelated traits may also be helpful, though this is unusual in MTMM studies. Byrne and Shavelson (1986), however, did consider and method factors (4D and 4C) and method effects represented as correlated uniquenesses (4E) seem most useful, supplemented perhaps by those positing 1 general factor (1A/2B), only trait factors (4A), and only method factors (1D). Particularly when Models 4D and 4E are both well-defined and lead to similar conclusions as with the Byrne and Shavelson (1986) data, then the traditional interpretation of these models is probably justified. In this case it may be reasonable to base inferences about method and trait effects on just these models -- dispensing with other models altogether. Other models from the taxonomy or models idiosyncratic to particular substantive issues may, however, provide useful supplemental information about the data. Because model 4E has not been widely applied elsewhere, it is important to further examine its apparent advantages in other studies. A particularly useful evaluation would be to apply various models -- including method structure E -- to simulated data in which the underlying factor structure was known. Subject to the results of this further research I recommend that Model 4E should be at least one of the MTMM models examined in all applications of CFA to MTMM data.

Footnotes

1 -- Campbell and Fiske (1959, p. 85) stated that "the presence of method variance is indicated by the difference in level of correlations between parallel values of the monomethod block and the heteromethod block, assuming comparable reliabilities among the tests." Marsh (in press) operationalized this statement to provide estimates of the relative size of method effects associated with each method of assessment and discussed limitations in the inferences based upon it.

2 -- Standard errors of estimated parameters that were extremely large were indicated to be 1.0 by Widaman (1985), though the footnote indicating this was mistakenly omitted from the published article (Widaman, personal communication, 3 September, 1987).

3 -- The CFA approach to MTMM data assumes that the different methods represent fixed effects. Whereas this limitation may be reasonable for some applications, it is probably inappropriate for the Marsh and Ireland data where the different raters more realistically constitute a random effects facet (i.e., a sample of potentially much larger sample of raters). I am not aware, however, of any solution to this problem.

4 -- This perspective was expressed by an anonymous reviewer.

REFERENCES

Althauser, R.P. & Heberlein, T.A. (1970). Validity and the multitrait-multimethod matrix. In E.F. Borgotta & W. Bohrnstedt (Eds.), Sociological methodology 1970 (pp. 151-169). San Francisco: Jossey-Bass.

Alwin, D. F. (1974). Approaches to the interpretation of relationships and the multitrait-multimethod matrix. In H. L. Costner (Ed.), Sociological methodology 1973-4 (pp. 79-105). San Francisco: Jossey-Bass.

Bagozzi, R. P. (1978). The construct validity of the affective, behavioral, and cognitive components of attitude by analysis of covariance structures. Multivariate Behavioral Research, 13, 9-31.

Bentler, P. M. & Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. Psychological Bulletin, 88, 588-606.

Bollen, K. A. & Joreskog, K. G. (1985). Uniqueness does not imply identification: A note on confirmatory factor analysis. Sociological Methods & Research, 14, 155-163.

Boruch, R. F., & Wolins, L. (1970). A procedure for estimation of trait, method and error variance attributable to a measure. Educational and Psychological Measurement, 30, 547-574.

Browne, M. W. (1980). Factor analysis of multiple batteries by maximum likelihood. British Journal of Mathematical and Statistical Psychology, 33, 184-199.

Browne, M. W. (1984). The decomposition of multitrait-multimethod matrices. British Journal of Mathematical and Statistical Psychology, 37, 1-21.

Byrne, B. M., & Shavelson, R.J. (1986). On the structure of adolescent self-concept. Journal of Educational Psychology, 78, 474-481.

Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by multitrait-multimethod matrix. Psychological Bulletin, 56, 81-105.

Campbell, D. T. & O'Connell, E. J. (1967). Method factors in multitrait-multimethod matrices: Multiplicative rather than additive? Multivariate Behavioral Research, 2, 409-426.

Cole, D. & Maxwell, S. E. (1985). Multitrait-multimethod comparisons across populations: A confirmatory factor analysis approach. Multivariate Behavioral Research, 4, 389-418.

Cudeck, R. & Browne, M. W. (1983). Cross-validation of covariance structures. Multivariate Behavioral Research, 2, 147-167.

Dillon, W. R., Kumar, A., & Mulani, Narendra. Offending estimates in covariance structure analysis: Comments on the causes and solutions to

Heywood cases. Psychological Bulletin, 101, 126-135.

Foley, J. J. (1971). Evaluation of learning in writing. In B. S. Bloom, J. T. Hastings, & G. F. Maduas (Eds.), Handbook on formative and summative evaluation of student learning. New York: McGraw-Hill.

Hubert, L. J. & Baker, F. B. (1978). Analyzing the multitrait-multimethod matrix. Multivariate Behavioral Research, 13, 163-179.

Jackson, D. N. (1969). Multimethod factor analysis in the evaluation of convergent and discriminant validity. Psychological Bulletin, 72, 30-49.

Jackson, D. N. (1974). Multimethod factor analysis: A reformulation. Multivariate Behavioral Research, 10, 259-275.

Joreskog, K. G. (1971). Statistical analysis of sets of congeneric tests. Psychometrica, 36 409-426.

Joreskog, K. G. (1974). Analyzing psychological data by structural analysis of covariance matrices. In R. C. Atkinson, D. H. Krantz, R.D. Luce, & Suppes (Eds.), Contemporary developments in mathematical psychology (Vol. 2, pp. 1-56). San Francisco: W. H. Freeman.

Joreskog, K. G. (1981). Analysis of covariance structures. Scandanavian Journal of Statistics, 8, 65-92.

Joreskog, K. G. & Sorbom, D. (1981). LISREL V: Analysis of Linear Structural Relations By the Method of Maximum Likelihood. Chicago: International Educational Services.

Kenny, D. A. (1979). Correlation and causality. New York: Wiley.

Long, K. S. (1983) Confirmatory factor analysis: A preface to LISREL. Beverly Hills, CA: Sage.

Marsh, H. W. (in press). Multitrait-multimethod analysis. In J. P. Keeves (Ed.) Educational research methodology, measurement, and evaluation: An international handbook. Oxford: Pergamon Press.

Marsh, H. W., Balla, J. and McDonald (1988). Goodness-of-fit indices in confirmatory factor analysis: The effect of sample size. Psychological Bulletin, 103, 391-410.

Marsh, H. W., Barnes, J., & Hocevar, D. (1985). Self-other agreement on multidimensional self-concept ratings: Factor analysis & multitrait-multimethod analysis. Journal of Personality and Social Psychology, 49, 1360-1377.

Marsh, H. W., & Butler, S. (1984). Evaluating reading diagnostic tests: An application of confirmatory factor analysis to multitrait-multimethod data. Applied Psychological Measurement, 8, 307-320.

Marsh, H. W., & Hocevar, D. (1983). Confirmatory factor analysis of multitrait-multimethod matrices. Journal of Educational Measurement, 20, 231-248.

Marsh, H.W. & Ireland, R. (1984). Multidimensional evaluations of writing effectiveness. (ERIC Document Reproduction Service No. ED 242 785).

Marsh, H. W., McDonald, R. P., & Balla, J. (1987). Goodness of fit in confirmatory factor analysis: The effect of sample size and of over-fitting or under-fitting models. (In Review).

Marsh, H. W., & Shavelson, R. J. (1985). Self-concept: Its multifaceted, hierarchical structure. Educational Psychologist, 20 107-125.

McDonald, R. P. (1985). Factor analysis and related methods. Hillsdale, NJ: Erlbaum.

Ostrom, T. M. (1969). The relationship between the affective, behavioral, and cognitive components of attitude. Journal of Experimental Social Psychology, 5, 12-30.

Rindskopf, D. (1983). Parameterizing inequality constraints on unique variances in linear structural models. Psychometrica, 48, 73-83.

Schmitt, N. (1978). Path analysis of multitrait-multimethod matrices. Applied Psychological Measurement, 2, 157-173.

Schmitt, N., Coyle, B. W., & Sarri, B. B. (1977). A review and critique of analyses of multitrait-multimethod matrices. Multivariate Behavioral Research, 13, 447-478.

Schmitt, N., & Stults, D. M. (1986). Methodological review: Analyses of multitrait-multimethod matrices. Applied Psychological Measurement, 10, 1-22.

Stanley, J. C. (1961). Analysis of unreplicated three way classification with application to rater bias and trait independence. Psychometrica, 26, 205-219.

Werts C. E. & Linn, R. L. (1970). Path analysis: Psychological examples. Psychological Bulletin, 74, 193-212.

Widaman, K. F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. Applied Psychological Measurement, 9, 1-26.

Wothke, W. (1984). The estimation of trait and method components in multitrait-multimethod measurement. (An unpublished Dissertation, University of Chicago).

Table 1

Parameters To Be Estimated For the General MTMM Model (Model 4B)

	Factor Loadings (Lambda Y)						Error/Uniquenesses (Theta)								
Variables	T1	T2	T3	M1	M2	M3	t1m1	t2m1	t3m1	t1m2	t2m2	t3m2	t1m3	t2m3	t3m3
t1m1	LY[a]	0	0	LY	0	0	TE								
t2m1	0	LY	0	LY[a]	0	0	0[b]	TE							
t3m1	0	0	LY	LY	0	0	0[b]	0[b]	TE						
t1m2	LY	0	0	0	LY	0	0	0	0	TE					
t2m2	0	LY[a]	0	0	LY	0	0	0	0	0[b]	TE				
t3m2	0	0	LY	0	LY[a]	0	0	0	0	0[b]	0[b]	TE			
t1m3	LY	0	0	0	0	LY[a]	0	0	0	0	0	0	TE		
t2m3	0	LY	0	0	0	LY	0	0	0	0	0	0	0[b]	TE	
t3m3	0	0	LY[a]	0	0	LY	0	0	0	0	0	0	0[b]	0[b]	TE

Factor Variance/Covariances (PSI)

Factors	T1	T2	T3	M1	M2	M3
T1	1[c]					
T2	PS	1[c]				
T3	PS	PS	1[c]			
M1	0	0	0	1[c]		
M2	0	0	0	PS	1[c]	
M3	0	0	0	PS	PS	1[c]

Note. Confirmatory factor analysis (CFA) models to be considered in this investigation are defined in terms of the three LISREL design matrices presented here. The MTMM problem shown here is has 3 traits factors (T1 - T3) and 3 method factors (M1 - M3) that are defined in terms of the 9 measured variable (t1m1 - t3m3). All parameters with valu , of 0 or 1 are fixed and not estimated whereas all other parameters are estimated without constraint. The parameterization shown here, with factor variances (in PSI) fixed to be 1, is referred to as the fixed factor variance parameterization.

[a] For the fixed factor loading parameterization these factor loadings would be fixed to be 1 and factor variances would be freed. [b] For the fixed factor loading parameterization these factor variances would be fixed to be 1 and no factor loadings would be fixed to 1. [c] For method structure E (see Table 2) these correlations between error/uniquenesses would be estimated, and the method factors and their associated parameters would be eliminated from Lambda Y and PSI.

Table 2

Taxonomy of Structural Models for MTMM Data Adapted From Widaman (1985)

Trait Structure	Method Structure A	B[a]	C	D	E
1[a]	1A Null Model	1B[b] 1 General M-factor	1C M uncorrelated M-factors	1D M correlated M-factors	1E T x M correlated errors
2[b]	2A 1 General T-factor	2B[c] 2 General factors	2C 1 General T-factor. M uncorrelated M-factors	2D 1 General T-factor, M correlated M-factors	2E[d] T x M correlated errors, 1 General T-factor
3	3A T uncorrelated T-factors	3B 1 General M-factor, T uncorrelated T-factors	3C T uncorrelated T-factors M uncorrelated M-factors	3D T uncorrelated T-factors, M correlated M-factors	3E T x M correlated errors, T uncorrelated T-factors
4	4A T Correlated T-factors	4B 1 General M-factor, T correlated T-factors	4C T Correlated T-factors, M uncorrelated M-factors	4D T Correlated T-factors, M correlated M-factors	4E T x M correlated errors, T correlated T-factors

[a] General factors are defined to be uncorrelated with other factors in the model. Although general factors are posited to represent either trait variance or method variance, this assumption will not always be accurate and may be difficult to test. [b] Models 2A and 1B are equivalent, and it is generally not possible to determine whether the one general factor reflects trait variance, method variance or some combination of trait and method variance. [c] Model 2B requires additional constraints that may be arbitrary and that may not provide equivalent solutions. Hence its usefulness may be dubious unless there is an a priori basis for the constraints. [d] Models under method structure E have no method factors. Instead method effects are inferred on the basis of correlated error/uniquenesses (see Table 1). This method structure, particularly when there are three traits, corresponds most closely to method structure C in which there are M uncorrelated methods.

Table 3

Summary of Goodness of Fit and Solution Behavior for the Ostrom Data

	Parameterization														
	Fixed Factor Loadings					Fixed Factor Variances					Rindskopf Parameterization				
Model[a]	χ^2	df	χ^2/df Ratio	TLI[b]	Problem[c]	χ^2	df	χ^2/df Ratio	TLI	Problem	χ^2	df	χ^2/df Ratio	TLI	Problem
Models Without Validity Factors															
1A	1872	66	28.36	.000	---										
1B/2A	141	54	2.61	.941	---										
1C	776	54	14.36	.512	---										
1D	75	48	1.56	.980	2	75	48	1.56	.980	2	75	48	1.56	.980	2
1E	776	54	14.36	.512	---										
2B	187	44	4.26	.881	---										
2C	73	42	1.73	.973	1	73	42	1.73	.973	1	73	42	1.73	.973	5
2D	40	36	1.11	.9[illegible]6	1	39	36	1.09	.997	---					
2E	73	42	1.73	.973	---										
3A	722	54	13.37	.548	---										
3B	112	42	2.68	.939	1	109	42	2.60	.942	---					
3C	608	42	14.47	.508	1	607	42	14.46	.508	4	607	42	14.46	.508	5
3D	44	36	1.23	.991	---										
3E	607	42	14.46	.508	---										
4A	135	51	2.66	.939	---										
4B	57	39	1.47	.983	---										
4C	54	39	1.38	.986	3	54	39	1.38	.986	1	54	39	1.38	.986	5
4D	29	33	0.87	1.005	4	22	33	0.66	1.012	1	29	33	0.87	1.005	5
4E	54	39	1.38	.986	---										
Models With Validity Factors															
1A											2041	78	26.16	.000	---
1D											127	57	2.22	.951	2
2D											53	44	1.20	.992	---
4A											164	61	2.69	.933	---
4B											89	48	1.85	.966	---
4C											57	45	1.27	.989	5
4D											38	39	.97	1.001	5
4E[d]											79	49	1.61	.976	---
4E'[d]											58	46	1.26	.990	---

Note. All models were first tested with both the fixed factor loading and the fixed factor variance parameterizations. However, in all cases in which the fixed factor loading parameterization resulted in a well defined solution the solution for the fixed factor variance parameterization was the same and so it is not presented. When neither of these standard parameterizations resulted in a well-defined solution the Rindskopf parameterization was used. For models that also had a validity factor, only the Rindskopf parameterization was used because the previous analyses showed that this parameterization facilitated convergence for this application.

[a] see Table 2 for a description of the models. [b] TLI = Tucker Lewis Index. [c] Problems: 1 = failed to converge in 250 iterations; 2 = factor correlation > 1.0; 3 = Negative factor variance; 4 = Negative error/uniqueness; 5 = estimates with excessively large standard errors. Problems 2, 3 and 4 were only examined if the solution converged, and problem 5 was examined only if the solution converged and had no out-of-range parameter estimates. [d] Validity factor is correlated with trait factors but not with any error/uniquenesses used to represent method effects in Model 4E whereas two such correlations appear in Model 4E'.

Table 4

Parameter Estimates For Model 4E For the Ostrea Data

	Factor Loadings			Error/Uniquenesses (Theta)											
Variables	T1	T2	T3	t1m1	t2m1	t3m1	t1m2	t2m2	t3m2	t1m3	t2m3	t3m3	t1m4	t2m4	t2m4
t1m1	.80*	0	0	.36*											
t2m1	0	.74*	0	.01	.46*										
t3m1	0	0	.81*	-.01	.06	.34*									
t1m2	.87*	0	0	0	0	0	.25*								
t2m2	0	.90*	0	0	0	0	.05*	.20							
t3m2	0	0	.88*	0	0	0	.04	.06*	.22*						
t1m3	.64*	0	0	0	0	0	0	0	0	.59*					
t2m3	0	.76*	0	0	0	0	0	0	0	-.03	.42				
t3m3	0	0	.88*	0	0	0	0	0	0	-.01	.01	.36*			
t1m4	.84*	0	0	0	0	0	0	0	0	0	0	0	.29*		
t2m4	0	.74*	0	0	0	0	0	0	0	0	0	0	.14*	.44*	
t3m4	0	0	.74*	0	0	0	0	0	0	0	0	0	.16*	.15*	.45*

Factor Variance/Covariances (PSI)

Factors	T1	T2	T3
T1	1		
T2	.96*	1	
T3	.98*	.94*	1

Note. See Table 1 for a definition of the MTMM factors and the variables. All parameters with values of 0 or 1 are fixed whereas all other parameters were freely estimated. Correlated errors/uniquenesses in the Theta matrix are between different traits assessed with the same method and are used to infer method effects.

* $p < .05$.

Table 5

Correlations Between Trait (T1-T3), Method (M1-M4), General (G1-G2) and Validity (V1) Factors for Selected MTMM Models Based on Ostrom study

Models[b]	MTMM Factors[a] T1	T2	T3	M1	M2	M3	M4	G1	G2	V1
Model 4E'										
T1	1									
T2	.96*	1								
T3	.97*	.94*	1							
V1	.68*	.74*	.60*	---	---	---	---	---	---	1
Model 4E										
T1	1									
T2	.96*	1								
T3	.97*	.94*	1							
V1	.69*	.77*	.61*	---	---	---	---	---	---	1
Model 4D										
T1	1									
T2	.91*	1								
T3	.95*	.91*	1							
M1	0	0	0	1						
M2	0	0	0	.91*	1					
M3	0	0	0	.89*	.86*	1				
M4	0	0	0	.75*	.70*	.74	1			
V1	.33*	.38*	.25*	.65*	.65*	.66*	.46*	---	---	1
Model 4B										
T1	1									
T2	.96*	1								
T3	.97*	.94*	1							
G1	0	0	0	---	---	---	---	1		
V1	.69*	.76*	.58*	---	---	---	---	-.21*	---	1
Model 4A										
T1	1									
T2	.98*	1								
T3	.99*	.97*	1							
V1	.67*	.76*	.60*	---	---	---	---	---	---	1
Model 2D										
M1	---	---	---	1						
M2	---	---	---	.97*	1					
M3	---	---	---	.96*	.95*	1				
M4	---	---	---	.88*	.87*	.88*	.			
G2	---	---	---	---	---	---	---	---	1	
V1	---	---	---	.59*	.62*	.59*	.51*	---	.59*	1
Model 1D[c]										
M1	---	---	---	1						
M2	---	---	---	.97*	1					
M3	---	---	---	1.01*	.99*	1				
M4	---	---	---	.89*	.87*	.93*	1			
V1	---	---	---	.70*	.68*	.73*	.58*	---	---	1

Note. Only parameter estimates from the matrix of factor correlations (PSI) are presented.

* $p < .05$

[a] T1 = Trait 1, T2 = Trait 2, T3= Trait 3, M1= Method 1, M2=Method 2, M3= Method 3, G1 = General 1 (general method), G2 = General 2 (general trait), V1 = Validity Criterion 1. [b] See Table 2 for a description of the models. [c] This model resulted in an improper solution in that 1 of the method factor correlations was greater than 1.0. Similar results were found when the validity criterion was not included (see Table 3).

Table 6

Summary of Goodness of Fit and Solution Behavior for the Byrne and Shavelson Data

	Parameterization														
	Fixed Factor Loadings					Fixed Factor Variances					Rindskopf Parameterization				
Model[a]	χ^2	df	χ^2/df Ratio	TLI[b]	Problem[c]	χ^2	df	χ^2/df Ratio	TLI	Problem	χ^2	df	χ^2/df Ratio	TLI	Problem
Without Validity Criteria															
1A	5272	36	146.44	.000	---										
1B/2A	2365	27	87.59	.405	---										
1C	3807	27	141.02	.037	1	3807	27	141.02	.037	1	3934	27	145.71	.005	5
1D	2302	24	95.92	.347	2	2302	24	95.92	.347	2	2302	24	95.92	.347	2
1E	3807	27			---										
2B	534	20	26.70	.823	---										
2C	1698	18	94.34	.358	1	1528	18	84.88	.423	4	1528	18	84.90	.423	5
2D	385	15	25.63	.831	1	323	15	21.52	.859	---					
2E	1528	18	84.90	.423	---										
3A	1107	27	41.01	.726	---										
3B	313	18	17.37	.887	1	310	18	17.21	.889	1	387	18	21.51	.859	5
3C	708	18	39.31	.737	1	707	18	39.26	.737	4	707	18	39.28	.737	5
3D	94	15	6.30	.964	3	135	15	9.03	.945	1	188	15	12.54	.921	5
3E	707	18	39.26	.737	---										
4A	451	24	18.81	.878	---										
4B	112	15	7.49	.955	5										
4C	65	15	4.31	.977	3 4	65	15	4.31	.977	1	66	15	4.38	.977	5
4D	28	12	2.35	.991	1	40	12	3.32	.984	---					
4E	65	12	4.31	.977	---										
With Validity Criteria															
1A						6399	45	142.21	.000	---					
1D						2699	38	71.03	.504	4					
2B						800	36	22.23	.850	---					
2D						469	27	17.35	.884	---					
4A						697	38	18.33	.877	---					
4B						268	27	9.92	.937	---					
4C						132	27	5.73	.967	1	148	23	6.45	.961	5
4D						73	20	3.65	.981	---					
4E[c]						363	29	12.52	.918	---					
4E'[c]						149	26	5.73	.966	---					

Note. All models were first tested with the fixed factor variance parameterization, models with solutions that had problems were then tested with the fixed factor variance parameterization, and if there were still problems with the Rindskopf parameterization. The expanded models with the validity criteria were tested with the fixed factor variance parameterization and then the Rindskopf parameterization was used if a poorly defined solution was obtained.

[a] see Table 2 for a description of the models. [b] TLI = Tucker Lewis Index. [c] Problems: 1 = failed to converge in 250 iterations; 2 = factor correlations > 1.0; 3 = Negative factor variance; 4 = Negative error/uniqueness; 5 = estimates with excessively large standard errors. Problems 2, 3 and 4 were only examined if the solution converged, and problem 5 was examined only if the solution converged and had no out-of-range parameter estimates. [c] In Model 4E, error/uniquenesses used to represent method effects were uncorrelated with the validity factor whereas for Model 4E' selected correlations were estimated.

Table 7

Parameters For Model 4D For the Byrne and Shavelson Data

	Factor Loadings (Lambda Y)						
Variables	T1	T2	T3	M1	M2	M3	Error/ Uniqueness
t1m1	.84*	0	0	.18	0	0	.27*
t2m1	0	.57*	0	.56*	0	0	.35*
t3m1	0	0	.94*	.03	0	0	.11*
t1m2	.68*	0	0	0	.26*	0	.48*
t2m2	0	.70*	0	0	.65*	0	.07[b]
t3m2	0	0	.93*	0	.19*	0	.12*
t1m3	.77*	0	0	0	0	.60*	.03[b]
t2m3	0	.87*	0	0	0	.33*	.15*
t3m3	0	0	.85*	0	0	.25*	.18*

Factor Variance/Covariances (PSI)

Factors	T1	T2	T3	M1	M2	M3
T1	1					
T2	.59*	1				
T3	.60*	.04	1			
M1	0	0	0	1		
M2	0	0	0	.80*	1	
M3	0	0	0	.25	.22	1

Note. The three traits are School self-concept (T1), Verbal self-concept (T2) and Math self-concept (T3) whereas the three methods corresponding to three different self-report instruments used to measure each of these facets of self (M1-M3). All parameters with values of 0 or 1 are fixed and not estimated whereas all other parameters are estimated without constraint. The fixed factor variance parameterization, with factor variances (in PSI) fixed to be 1, was used to estimate parameters.

[a] Because error/uniquenesses were constrained to be uncorrelated in this model the estimates are presented in this form to save space (see Theta matrix in Table 1). [b] Although these error/uniquenesses did not differ significantly from zero, their standard errors were very small.

Table 8

Correlations Between Trait (T1-T3), Method (M1-M3), General (G1-G2) and Criterion (C1-C2) Factors for Selected MTMM Models Based on Byrne and Shavelson Study

Models[b]	MTMM Factors[a]									
	T1	T2	T3	M1	M2	M3	G1	G2	V1	V2
Model 4E'										
T1	1									
T2	.61*	1								
T3	.61*	.05	1							
V1	.56*	.63*	.21*	---	---	---	---	---	1	
V2	.48*	.07*	.57*	---	---	---	---	---	.52*	1
Model 4E										
T1	1									
T2	.61*	1								
T3	.61*	.05	1							
V1	.54*	.43*	.22*	---	---	---	---	---	1	
V2	.46*	.08*	.59*	---	---	---	---	---	.52*	1
Model 4D										
T1	1									
T2	.61*	1								
T3	.61*	.06	1							
M1	0	0	0	1						
M2	0	0	0	.85*	1					
M3	0	0	0	.22	.18	1				
V1	.59*	.57*	.21*	-.13*	-.16*	.10*	---	---	1	
V2	.41*	.06	.56*	.07	.02	.37*	---	---	.52*	1
Model 4C										
T1	1									
T2	.62*	1								
T3	.61*	.06	1							
M1	0	0	0	1						
M2	0	0	0	.85*	1					
M3	0	0	0	.22	.18	1				
V1	.52*	.60*	.21*	-.29*	-.38*	.21*	---	---	1	
V2	.40*	.07	.57*	.07	.02	.31*	---	---	.52*	1
Model 4B										
T1	1									
T2	.66*	1								
T3	.63*	.10*	1							
G1	0	0	0	---	---	---	1			
V1	.54*	.50*	.20*	---	---	---	-.32*	---	1	
V2	.48*	.12*	.59*	---	---	---	-.25*	---	.52*	1
Model 4A										
T1	1									
T2	.66*	1								
T3	.63*	.08*	1							
V1	.58*	.44*	.22*	---	---	---	---	---	1	
V2	.50*	.09*	.60*	---	---	---	---	---	.52*	1
Model 2B										
G1	---	---	---	---	---	---	1			
G2	---	---	---	---	---	---	0	1		
V1	---	---	---	---	---	---	.53*	.21*	1	
V2	---	---	---	---	---	---	.11*	.61*	.52*	1
Model 2D										
M1	---	---	---	1						
M2	---	---	---	.98*	1					
M3	---	---	---	.89*	.82*	1				
G2	---	---	---	0	0	0	---	1		
V1	---	---	---	.54*	.39*	.62*	---	.08	1	
V2	---	---	---	.38*	.30*	.48*	---	-.47*	.52*	1
Model 1D[c]										
M1	---	---	---	1						
M2	---	---	---	1.11*	1					
M3	---	---	---	1.05*	.95*	1				
V1	---	---	---	.36*	.22*	.35*	---	---	1	
V2	---	---	---	.66*	.54*	.70*	---	---	.52*	1

Note. Only parameter estimates from the matrix of factor correlations (PSI) are presented.

* $p < .05$

[a] See Table 4 for a description of the MTMM factors. [b] See Table 2 for a description of the models. [c] This model resulted in an improper solution with 2 of 3 method factor correlations being greater than 1.0.

Table 9

Summary of Goodness of Fit and Solution Behavior for the Marsh and Ireland Data

Model[a]	X^2	df	X^2/df Ratio	TLI[b]	Problem[c]
Without Validity Criteria					
1A	6407	630	10.17	.000	---
1B/2A	2439	593	4.11	.660	---
1C	1991	594	3.36	.742	---
1D	1296	579	2.24	.865	---
1E	1590	540	2.94	.788	---
2C	1199	558	2.15	.875	---
2D	997	543	1.84	.909	---
2E	852	504	1.69	.925	---
4A	2294	479	3.96	.677	---
4B	1776	543	3.27	.752	2
4C	981	543	1.81	.912	---
4D	845	528	1.60	.935	---[d]
4E	722	489	1.48	.948	---
With Validity Criterion					
1C	6593	666	9.90	.000	---
1D	1334	610	2.19	.867	---
2D	1035	573	1.83	.909	---
2E	896	540	1.66	.923	---
4A	2331	610	3.82	.683	---
4B	1814	573	3.17	.757	2
4C	1009	568	1.78	.913	---
4D	873	533	1.64	.928	---
4E	758	520	1.56	.948	---

Note. All models were first tested with the fixed factor variance parameterization.

[a] see Table 2 for a description of the models. [b] TLI = Tucker Lewis Index.

[c] Problems: 1 = failed to converge in 250 iterations; 2 = factor correlations > 1.0; 3 = Negative factor variance; 4 = Negative error/uniqueness; 5 = estimates with excessively large standard errors of estimate. Problems 2, 3 and 4 were only examined if the solution converged, and problem 5 was examined only if the solution converged and had no out-of-range parameter estimates.

Table 10

Correlations Between Trait (T1-T6), Method (M1-M6), General (G1-G2) and Validity (V1) Factors for Selected MTMM Models Based on Marsh and Ireland Data

Factors in MTMM Models

Model 4E

	T1	T2	T3	T4	T5	T6	M1	M2	M3	M4	M5	M6	G1	G2	V1
T1	1														
T2	98*	1													
T3	92*	96*	1												
T4	93*	94*	93*	1											
T5	90*	91*	93*	96*	1										
T6	94*	96*	95*	98*	98*	1									
V1	81*	80*	77*	83*	76*	80*	---	---	---	---	---	---	---	---	1

Model 4D

	T1	T2	T3	T4	T5	T6	M1	M2	M3	M4	M5	M6	G1	G2	V1
T1	1														
T2	99*	1													
T3	64*	76*	1												
T4	70*	74*	61*	1											
T5	57*	58*	62*	84*	1										
T6	78*	79*	68*	93*	98*	1									
M1	0	0	0	0	0	0	1								
M2	0	0	0	0	0	0	71*	1							
M3	0	0	0	0	0	0	72*	71*	1						
M4	0	0	0	0	0	0	72*	73*	79*	1					
M5	0	0	0	0	0	0	66*	65*	79*	77*	1				
M6	0	0	0	0	0	0	72*	72*	78*	67*	65*	1			
V1	28*	25*	14	30*	17	22*	66*	56*	69*	68*	66*	67*	---	---	1

Model 4C

	T1	T2	T3	T4	T5	T6	M1	M2	M3	M4	M5	M6	G1	G2	V1
T1	1														
T2	99*	1													
T3	92*	95*	1												
T4	92*	94*	92*	1											
T5	89*	89*	94*	97*	1										
T6	94*	95*	94*	98*	99*	1									
M1	0	0	0	0	0	0	1								
M2	0	0	0	0	0	0	0	1							
M3	0	0	0	0	0	0	0	0	1						
M4	0	0	0	0	0	0	0	0	0	1					
M5	0	0	0	0	0	0	0	0	0	0	1				
M6	0	0	0	0	0	0	0	0	0	0	0	1			
V1	77*	75*	71*	79*	73*	75*	08	-10	06	12	08	16*	---	---	1

Model 4E

	T1	T2	T3	T4	T5	T6	M1	M2	M3	M4	M5	M6	G1	G2	V1
T1	1														
T2	98*	1													
T3	92*	96*	1												
T4	93*	94*	93*	1											
T5	90*	91*	93*	96*	1										
T6	94*	96*	95*	98*	98*	1									
V1	81*	80*	77*	83*	76*	80*	---	---	---	---	---	---	---	---	1

Model 4B

	T1	T2	T3	T4	T5	T6	M1	M2	M3	M4	M5	M6	G1	G2	V1
T1	1														
T2	106*	1													
T3	64*	80*	1												
T4	66*	84*	77*	1											
T5	45*	59*	101*	100*	1										
T6	71*	82*	81*	110*	110*	1									
G1	0	0	0	0	0	0	---	---	---	---	---	---	1		
V1	27*	26*	20*	35*	26*	29*	---	---	---	---	---	---	---	74*	1

Table 10 continued on next page

Table 10 Continued

Factors in MTMM Models

	T1	T2	T3	T4	T5	T6	M1	M2	M3	M4	M5	M6	G1	G2	V1
Model 4A															
T1	1														
T2	104*	1													
T3	96*	99*	1												
T4	96*	98*	98*	1											
T5	94*	95*	101*	103*	1										
T6	98*	99*	100*	104*	107*	1									
V1	81*	80*	76*	83*	77*	80*	---	---	---	---	---	---	---	---	1
Model 2D															
M1							1								
M2							73*	1							
M3							71*	70*	1						
M4							74*	77*	79*	1					
M5							67*	69*	78*	78*	1				
M6							72*	72*	77*	68*	65*	1			
G2							0	0	0	0	0	0	1		
V1							65*	53*	67*	67*	63*	64*	33*	1	
Model 1D															
M1							1								
M2							77*	1							
M3							75*	75*	1						
M4							78*	79*	82*	1					
M5							73*	73*	81*	80*	1				
M6							77*	76*	81*	72*	70*	1			
V1							73*	61*	75*	74*	71*	72*	72*	1	

Note Only factor correlation from the Psi matrix are presented. Parameter estimates are presented without decimal points. All parameters with values of 1 or 0 are fixed, whereas factor correlations greater than 100 are out-of-range estimates.

* $p < 05$

[a] See Table 4 for a description of the MTMM factors. [b] See Table 2 for a description of the models