

Focus on Psychometrics

More on MTMM: The Role of Confirmatory Factor Analysis

Aurelio J. Figueredo, Sandra L. Ferketich, and Thomas R. Knapp

This article is the second of two on the use of confirmatory factor analysis (CFA) as a method to assess construct validity. The construct validation criteria required by the conventional MTMM approach are satisfied only by certain ideal data sets, such as those in which the method variance of measures is very low. The CFA approach to multitrait-multimethod (MTMM) data is more general, in that violations of those stringent criteria can be managed. Another limitation of the conventional MTMM approach is that only a relatively small number of indicators can be examined by bivariate analysis. The economy of the CFA approach permits the analysis of a much larger number of indicators. In this article, a data set is analyzed using the CFA approach. Results are presented that illustrate the application of this statistical method.

Although Campbell and Fiske (1959) made important contributions to our understanding of the nature of construct validation procedures, their method for analyzing a multitrait-multimethod matrix has some shortcomings. Those shortcomings were discussed in the article, "The multitrait-multimethod approach to construct validity" (Ferketich, Figueredo, & Knapp, 1991). An alternate approach is the use of confirmatory factor analysis (CFA) to analyze the data from a multitrait-multimethod (MTMM) study. The purpose of this article is to describe the findings from a CFA approach to the analysis of data from an MTMM-inspired study.

The selection of this particular data set was prompted by two current trends in nursing. The first trend is the use of multiple indicators of complex phenomena, often with the underlying assumption that no single measure is an adequate index of the phenomena. The second trend is the

incorporation of both self-report and physiological measures of psychosocial phenomena, with the goals of obtaining greater accuracy of measurement and more adequate indexing of responses.

As we discussed in the previous article (Ferketich, Figueredo, & Knapp, 1991), few studies are specifically designed to provide data on different traits from different measurement methods. The data set selected, although not from nursing, provides data on three different psychological states using two measurement methods, thus yielding a particularly rich data base for analysis. These data were obtained from a psychological experiment that was laboratory-based (Smith, Figueredo, & Schwartz, 1991). The experiment was designed to measure states rather than traits. We do not believe, however, the use of states alters the basic logic of the approach (Figueredo & Petrinovich, 1991). Hereafter, for accuracy, the term Multistate-Multimethod (MSMM) will be used

Aurelio J. Figueredo, PhD, is an assistant professor in the Department of Psychology, University of Arizona, and a member of the Evaluation Group for Analysis of Data. Sandra Ferketich, PhD, RN, FAAN, is an associate professor of nursing and the director of the Predoctoral and Postdoctoral Clinical Nursing Research Instrumentation Fellowship Program (NRSA Institutional Grant No. 5T32 NR07029) at the University of Arizona. Thomas R. Knapp, EdD, is a professor in the College of Nursing and the College of Education, Ohio State University.

This is the sixth article in the series, *Focus on Psychometrics*, edited or contributed by Dr. Ferketich.

The authors thank Bradley H. Smith, MS, for permission to use these results as an example of an MTMM analysis.

Requests for reprints can be addressed to Dr. Sandra Ferketich, College of Nursing, University of Arizona, Tucson, AZ 85721.

to refer to this particular application of the MTMM approach. Although the actual experiment reported is not from nursing research, the design and approach could be used for nursing studies by changing the stimulus conditions.

EXPERIMENT

Briefly, the experiment was done as follows. Subjects were shown 12 stimulus cards each designed to evoke different emotions. Two methods of measurement were used to assess the responses of the subject to the content of the card. The first measurement method was self-report, using a list of Likert-scaled items. The nine verbal labels used in the Likert scales were: *happy, sad, angry, fearful, anxious, depressed, ambivalent, surprised, and disgusted*. The second measurement method was physiological, using three indicators of physiological response: two of muscle activity and one of heart rate. Muscle activity associated with smiling was indexed by electromyography (EMG), measuring changes in the zygomatic major muscle which controls retracting the corners of the mouth. Muscle activity associated with frowning was indexed by electromyography (EMG), measuring changes in the corrugator supercilli muscle which controls wrinkling the eyebrows. Heart rate was measured by digital photoplethysmography (PPG).

Physiological recordings were made by EMG and digital PPG for both: (1) 15 s while the card was viewed by the subject with eyes open and (2) 15 s while the subject reexperienced the elicited emotion but kept eyes closed. Since movement artifacts might be associated with either holding or examining the card, both situations were measured. In this experiment, no such artifacts were actually found, as indicated by the similarity between factor loadings for the corresponding conditions. The means and standard deviations of these continuous recordings were used as separate indicators in the analysis, since facial emotional expressions may often be characterized by fleeting motions rather than fixed positions. Therefore, the standard deviation contained information not evident from the mean. After each stimulus presentation, each subject was asked to self-report the emotions experienced by rating each verbal label as describing the emotion(s) experienced, using the Likert-scale to assess each item.

CFA Model

The MSMM-CFA model for this experiment contains two novel features. A few clarifying re-

marks concerning certain unusual features of the model are, therefore, in order. The purpose of the study was to apply the principle of MTMM analysis (Campbell & Fiske, 1959) and the techniques of CFA developed for such data (Widaman, 1985) to the study of transient emotional states rather than permanent personality traits.

The first unusual feature of the model is that *subject residuals*, which are unstandardized deviations from the subject means, were used on all measures in place of the raw scores. This eliminated the main effects of any stable individual differences, or *traits*, and permitted the desired analysis of transient emotional states manifested by intraindividual variation. Because a transient emotional state represents a temporary deviation from an individual's baseline condition, using subject residuals statistically removes from the analysis the baseline personality characteristics of that individual. Second, there was no causal relationship posited between the physiological response and the cognitive response (as in James, 1890, and Cannon, 1927). Instead, this MSMM-CFA approach treats the physiological and cognitive responses as common expressions, or manifest indicators, of a latent transient state, which is not directly observable.

CFA Method

Clearly, using the conventional MTMM approach (Campbell & Fiske, 1959) to examine the bivariate correlations between 21 different measures, 9 self-report and 19 physiological measures, would be unwieldy. Such a matrix would contain a total of 21 different diagonal, or *reliability*, coefficients and 210 different off-diagonal, or various *validity*, coefficients, requiring an unreasonable number of comparisons. Thus, although multiple indicators are generally a useful way of gathering data about a phenomenon, an excess of these may render bivariate analyses impractical.

For the MSMM-CFA approach, the research hypotheses were converted into the programming statements necessary to run the latent variable model program. The statements are specific to the software used, and will not be detailed here. However, the overall approach to the model and the expected relationships will be discussed.

The simple theoretical model was composed of five orthogonal factors. Three factors were latent variables for emotional states and two factors were latent variables for methods. The emotional states were labeled as *Good Emotions, Bad Emotions, and Strong Emotions*. Smith, Figueredo, and

Schwartz (1991) posited that certain self-reports, such as *happy* and *surprised*, and certain physiological recordings, such as zygomatic activity (*smiling*), will be indicators of a single state construct for Good Emotions. Other self-reports, such as *sad* and *anxious*, and other physiological recordings, such as corrugator activity (*frowning*), will be indicators of a second state construct for Bad Emotions. Finally, certain self-reports, such as *surprised* and *anxious*, and certain physiological recordings, such as heart rate (*arousal*), will be indicators of a third state construct for Strong Emotions. In part, Campbell and Fiske's (1959) Conditions 1 and 2 can be assessed by examining the factor loadings of the indicators. The factor loadings of these indicators on the emotion constructs should be significant on the hypothesized factors, such as both *Bad* and *Strong* for anxious self-report, and zero on any others. The correlation of any state factor to any given method factor should be zero.

The positing of the two latent variables for methods is the principal utility of the MSMM-CFA approach. That is, the method variance associated with the different measures is statistically controlled by the construction of latent *method* factors. Explicitly modeling the contribution from these latent constructs (a) disattenuates the correlation between dissimilar measures of the same emotions, enhancing the assessment of convergent

validity across methods, and (b) explains the correlation between similar measures of different emotions, enhancing the assessment of discriminant validity between states. This form of statistical control is not possible using the Campbell and Fiske (1959) MTMM approach in which the researcher must rely on discerning consistent patterns of relationships among the various bivariate correlations in both the heteromethod and monomethod blocks. (See Campbell and Fiske Conditions 3 and 4 in Ferketich, Figueredo, & Knapp, 1991.)

RESULTS

Using Widaman's (1985) CFA approach, the rather simple model described above was tested. The matrix of loadings of indicators on the emotion factors and method factors is shown in Table 1. Surprisingly, this simple model showed an excellent degree of practical goodness-of-fit to the data. Although a number of alternative models might also adequately fit the data, the hypothesized measurement model appears sufficient to generate most of the observed covariation.

One of the shortcomings of the Campbell and Fiske (1959) approach is that the method of analyzing the matrix is highly subjective, especially when the pattern of relationships among the bivariate correlations is not consistent. With the

Table 1. Multistate–Multimethod Model Standardized Factor Loadings

	Good Emotions	Bad Emotions	Strong Emotions	Self-Reports	Physio Records
Happy	.806			-.223	
Sad		.618		-.638	
Angry		.831	.175	.290	
Fearful		.267	.463	-.086	
Anxious		.410	.845	.066	
Depressed		.575		-.460	
Ambivalent			-.038	-.053	
Surprised	.567		.239	.139	
Disgusted		.721		.454	
Corrugator M1		.095			.125
Corrugator S1		.061			.306
Zygomatic M1	.223				.368
Zygomatic S1	.187				.429
Heart rate M1			.071		.022
Heart rate S1			.077		.207
Corrugator M2		.237			.311
Corrugator S2		.072			.364
Zygomatic M2	.283				.467
Zygomatic S2	.053				.122
Heart rate M2			.068		.021
Heart rate S2			.111		.299

Note. M = Mean; S = Standard Deviation; 1 = eyes open; 2 = eyes closed.

CFA approach, several test indices are provided that help the researcher make judgments about the fit of the data to the model. In EQS, a latent variable modeling software (Bentler, 1989), two indices are provided to assess the "practical" goodness-of-fit. These two indices, the Bentler–Bonett Normed (NFI) and Nonnormed Fit (NNFI) Indices, should be higher than .90 to be acceptable, indicating that most of the covariance between measures has been reproduced by the model. These indices were .994 (NFI) and .997 (NNFI), respectively. The chi-square goodness-of-fit index was statistically significant ($\chi^2 = 463.835$ for 270 degrees of freedom, $p < .001$). However, this statistic has been shown (Bentler & Bonett, 1980) to be oversensitive to large sample sizes, rejecting virtually any model tested because even the smallest residuals are almost never exactly equal to zero.

As was hypothesized, the factor for Good Emotions loaded positively on the self-reports of *happy* and *surprised* and both the means and standard deviations of zygomatic EMG activity in both response conditions (see Table 1). Also, the factor for Bad Emotions loaded positively on the self-reports of *sad*, *angry*, *fearful*, *anxious*, *depressed*, and *disgusted*, and both the means and standard deviations of corrugator EMG activity in both response conditions. The factor for Strong Emotions loaded positively on the self-reports of *angry*, *fearful*, *anxious*, and *surprised*, negatively on the self-report of *ambivalent*, and positively on both the means and standard deviations of heart rate PPG activity in both response conditions. This result was predictable by anyone who has experienced strong, or truly "heartfelt," emotions. The factor for self-report methods loads positively on *angry*, *anxious*, *surprised*, and *disgusted*, and negatively on *happy*, *sad*, *fearful*, *depressed*, and *ambivalent*. This indicates that the systematic bias introduced by self-report is not consistent across items, but dependent (i.e., "item-specific") on what question is being asked due to overreporting some emotional labels and underreporting others. The factor for physiological recording methods loads positively on both the means and standard deviations of corrugator EMG activity, zygomatic EMG activity, and heart rate PPG activity, always in both response conditions. This indicates that the systematic bias introduced by physiological recording is consistent in direction, generally tending to overestimate emotional response.

Thus, it can be seen that a major traditional obstacle, failing to control for systematic *method* variance, is surmounted with this approach. If this method variance is not controlled through the appropriate psychometric methods, bivariate correlations will be greatly attenuated.

SUMMARY

How useful to nursing researchers is the use of this complex approach to assess construct validity? We believe MTMM–CFA holds promise because it performs the analysis better than the conventional approach, even though the statistical model is more complex. The two basic tenets of the Campbell and Fiske (1959) approach to construct validity are addressed by the MTMM–CFA approach. In this example, it could be concluded that the measures of Good Emotions are more strongly related no matter which method is used to gather the data. That is, Good Emotions explains the covariance in the self-reports of *happy* and *surprised* and the physiological recordings of zygomatic activity. Thus, if we were looking at Campbell and Fiske's (1959) basic tenets, Tenet One would be met. The indicators of the construct, Good Emotions, correlate among themselves no matter what the method. The same situation is clearly evident with Bad Emotions. The fact that the measures do not "mix" across constructs of Good and Bad Emotions supports the second tenet that tests measuring different constructs, although using the same method, should not be excessively related.

Substantively, the use of physiological recordings and self-report measures to assess psychosocial phenomena holds promise for researchers assessing states across individuals as well as states within individuals. Such designs may hold promise for researchers of children's pain, for example, who might want to measure transient states within the individual. Family researchers may be interested in aggregating family data based on both physiological and self-report measures of each of the family members. Thus, nursing researchers may find challenges and solutions in applying this type of design and analysis to nursing problems.

We hope this set of two articles has stimulated thinking about both alternate approaches to and substantive uses of the MTMM–CFA approach. Its versatility allows the researcher to examine a number of states or traits and a number of methods. Additionally, a model can be designed and analyzed to handle multiple methods and traits, as well as multiple indicators of each.

REFERENCES

- Bentler, P.M. (1989). *EQS: Structural equations program manual*. Los Angeles: BMDP Statistical Software, Inc.
- Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of co-

- variance structures. *Psychological Bulletin*, 88, 588–606.
- Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81–105.
- Cannon, W.B. (1927). The James–Lange theory of emotions: A critical examination and an alternative theory. *American Journal of Psychology*, 39, 106–124.
- Ferketich, S., Figueredo, A.J., & Knapp, T.R. (1991). Focus on psychometrics: The multitrait-multimethod approach to construct validity. *Research in Nursing & Health*, 14, 315–320.
- Figueredo, A.J., & Petrinovich, L. (1991). *Multivariate methods in ethology*. Symposium conducted at the Western Psychological Association Conference, San Francisco, CA. Monograph in preparation.
- James, W. (1890). *Principles of psychology*. New York: Holt.
- Schachter, S., & Singer, J.E. (1962). Cognitive, social, and physiological determinants of emotional states. *Psychological Review*, 69, 379–399.
- Smith, B.H., Figueredo, A.J., & Schwartz, G. (1991). A multistate multimethod model of human emotional expression. Symposium presentation of the Western Psychological Association Conference, San Francisco, CA. A.J. Figueredo & L. Petrinovich (Eds.), *Multivariate methods in ethology*. Monograph in preparation.
- Widaman, K.F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, 9, 1–26.