Emotion
2005, Vol. 5, No. 2, 236–237

1528-3542/05/$12.00 DOI: 10.1037/1528-3542.5.2.236

REPLY TO COMMENTARY

A Discrepancy in Analyses of the MSCEIT—Resolving the Mystery and Understanding Its Implications: A Reply to Gignac (2005)

John D. Mayer
University of New Hampshire

A. T. Panter
University of North Carolina at Chapel Hill

Peter Salovey
Yale University

David R. Caruso
Work-Life Strategies

Gill Sitarenios
Multi-Health Systems

G. E. Gignac (2005) reanalyzed the factor structure of the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT) and found results that differed from those the authors obtained initially. The authors tracked down the surprising sources of those discrepancies. G. E. Gignac's hierarchical model of emotional intelligence appears promising, and the authors anticipate that further investigations of the MSCEIT factor structure may yield additional information.

Keywords: emotional intelligence, MSCEIT, structural equation modeling, factor analysis, goodness of fit

Does the Mayer–Salovey–Caruso Emotional Intelligence Test (MSCEIT) measure one unified emotional intelligence (EI) ability, or two, or four intercorrelated abilities, or something more complex? Such a question has implications both for scoring the test and for the nature of EI generally. Gignac's (2005) "Evaluating the MSCEIT V2.0 via CFA: Comment on Mayer et al. (2003)" raises questions about a factor analysis we used to address such questions.

A Mysterious Discrepancy

We had used structural equation modeling to test one-, two-, and four-factor models of the MSCEIT. Gignac noted that our reports of fit indices (the normed fit index [NFI] and the Tucker–Lewis index [TLI]) for our one- and two-factor models were quite high and that values he obtained using our correlation matrix and AMOS software were lower. He remarks, "It is difficult to discern what error Mayer et al. (2003) made" (Gignac, 2005, p. 234).

When the journal *Emotion* forwarded Gignac's (2005) comment to us, we were able to confirm that we had specified our models in the recommended fashions and had correctly reported them. Yet, when we reran the analyses ourselves, using the identical specifications and data, we too found that the NFI and TLI fit indices were lower than those indicated in the printouts from our earlier work—even though nearly every other value in our printout—including all parameter estimates, exactly matched our old records. What was the explanation of this odd state of affairs?

The Mystery Explained

In between our first analyses and the later analyses by Gignac and ourselves, AMOS Version 4 was upgraded to 4.02. Among other changes, this upgrade altered how the incremental fit indices were computed. Fit indices are calculated by comparing the researcher-specified model to a comparison baseline model. There is, however, some latitude in what such a comparative baseline model might be, and fit indices will change accordingly (Widaman & Thompson, 2003).

AMOS, in fact, uses different comparative baseline models in different analytic situations, which it selects automatically (Arbuckle, 2003, Appendix F, p. 85). The Mayer, Salovey, Caruso, and Sitarenios (2003) results were based on the AMOS 4 program; it compared both the one- and two-factor models to a baseline model that constrained means, intercepts, and covariances to zero. The four-factor model, however, was compared against a baseline-

Editor's Note. Michael Eid served as guest editor.—RJD

John D. Mayer, Department of Psychology, University of New Hampshire; A. T. Panter, Department of Psychology, University of North Carolina at Chapel Hill; Peter Salovey, Department of Psychology, Yale University; David R. Caruso, Work-Life Strategies, New Canaan, Connecticut; Gill Sitarenios, Research and Development, Multi-Health Systems, Toronto, Ontario, Canada.

John D. Mayer, Peter Salovey, and David R. Caruso receive royalties from the Mayer–Salovey–Caruso Emotional Intelligence Test, which is published by Multi-Health Systems (MHS), Toronto, Canada. Gill Sitarenios is Director of Research and Development at MHS.

Correspondence concerning this article should be addressed to John D. Mayer, Department of Psychology, 10 Library Way, University of New Hampshire, Durham, NH 03824. E-mail: jack.mayer@unh.edu

comparison model that relied on modeling independence among covariances only.

It was this switch in baseline comparison models in AMOS 4 that led to the anomaly first noticed by Gignac: The NFI we reported unexpectedly dropped for our four-factor model (relative to the one- and two-factor models). Gignac (2005) noted that "it is impossible for NFI to decrease from model to model" (p. 233) given that the baseline-comparison chi-square remained constant. If the baseline comparison model does change, however, as it did in AMOS 4, it would account for this anomaly.[1] With the introduction of AMOS 4.02, however, the same comparative-baseline model (independence among covariances only) was used across all three models.[2] When that new baseline model is applied to the one- and two-factor solutions, the fit indices are lower because the new baseline model already has a better fit to most data sets, and improvements made by researcher-specified models perform less well by comparison.[3]

Other users of the newer AMOS version have noted this issue as well. For example, in an SPSS Technical Support memo (personal communication, March 14, 2003; re: Resolution no. 31828), a user noted that fit indices were often "quite different" across versions 4.01 and 4.02 of AMOS and asked whether any changes had been made to the program. The technical staff responded that the algorithms in AMOS had been altered such that "many fit measures . . . are worse."

What Is the Factor Structure of the MSCEIT?

Given the revised fit measures provided by AMOS, what is the best factor model for the MSCEIT? Gignac's preferred solution is a hierarchical one in which an overall EI factor has more specific factors nested within it. He concludes that one- and two-factor solutions lacking such nested factors provide a poor fit. In addition, he notes that a three-factor solution that combines Branches 1 and 2 is better than a four-factor solution. Such a model would fit our conceptual model fairly closely, except in its combining of Branches 1 and 2 into a three-factor model. We ourselves have noted the plausibility of a similar three-factor model for EI (Mayer, Caruso, & Salovey, 2000), although such a solution also has its critics (Roberts, Zeidner, & Matthews, 2001, p. 216).

Although his hierarchical model appears promising, Gignac points out—and we agree—that perhaps a factor solution based on only two observed variables for each latent variable is limited. To remedy this, Gignac suggests adding more subscales to the entire instrument. If one is interested in the factor structure of EI ability generally, a broader range of subscales makes sense. If one is interested in the factor structure of the MSCEIT itself, however, a better approach may be to further divide the test by focusing on 16 or 32 parcels of the scale. Item parcels are small groups of items (MacCallum & Austin, 2000). These parcels might be composed more specifically of item testlets—sets of items with common stems that can be analyzed using item-response models or item-level factor analytic approaches (Steinberg & Thissen, 1996).

Until the MSCEIT is analyzed using such a further division of the branches, we believe the ultimate choice between hierarchical and nonhierarchical, three-, four-, or other-level factor models should, perhaps, be postponed. Debates over the dimensionality and hierarchality of human abilities have gone on for decades (e.g., Carroll, 1993). They are not likely to be quickly settled in the area of EI any more than they have been in the case of, say, cognitive intelligence. In the meantime, it is reassuring to note that a hierarchical model supports the idea of a unitary EI, and either a three- or a four-factor model provides some good discrimination among different kinds of EI within that broader approach.

[1] After we discovered the change in results from AMOS 4 to AMOS 4.02, we went back to our original printouts to examine the baseline comparison model used in greater detail. There, we found that the chi-square for the baseline comparison model in AMOS 4 changed dramatically between the two- and four-factor models: 51,874.78 for the one- and two-factor models; 4,163.87 for the four-factor model. Carrying this through, for the one-factor model, the NFI = (51,874.78 − 626.56)/51,874.78 = 0.9879, for the two-factor model, NFI = (51,874.78 – 347.32)/51,874.78 = 0.9933, and for the four-factor model, with variances constrained, NFI = (4,163.87 – 94.28)/4,163.87 = 0.9774, all as reported originally.

[2] According to Arbuckle (personal communication, August 26, 2004), the rationale for the shift in baseline model for the newer version of AMOS in computing incremental fit indices is as follows: It is important to make the baseline or independence model as bad as it could possibly be. In earlier versions of AMOS, it was assumed that most users would be modeling mean and intercept parameters explicitly in their SEMs and that they would use these parameters in hypothesis testing. More commonly, however, modelers use means and intercepts only when they have missing data, and in those cases, they do not typically constrain means and intercepts. Thus, there was no strong rationale for constraining means in the baseline model. The baseline model used in the newer version of AMOS now is consistent with the baseline model in the other major SEM programs and used in calculations of incremental fit indices.

[3] More technically, on a continuum of baseline-comparison models ranging from zero covariance among indicators to a saturated model, the closer the baseline model is to the saturated model, the smaller the discrepancy between the chi-square for the baseline comparison (null) model and the substantive model, resulting in lower proportional improvement.

References

Arbuckle, J. L. (2003). *Amos 5.0 update to the AMOS users guide.* Chicago: SPSS/Smallwaters.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytical studies.* New York: Cambridge University Press.

Gignac, G. E. (2005). Evaluating the MSCEIT V2.0 via CFA: Comment on Mayer et al. (2003). *Emotion, 5,* 233–235.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology, 51,* 201–226.

Mayer, J. D., Caruso, D., & Salovey, P. (2000). Emotional intelligence meets traditional standards for an intelligence. *Intelligence, 27,* 267–298.

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2.0. *Emotion, 2,* 97–105.

Roberts, R. D., Zeidner, M., & Matthews, G. (2001). Does emotional intelligence meet traditional standards for an intelligence? Some new data and conclusions. *Emotion, 1,* 196–231.

Steinberg, L., & Thissen, D. (1996). Uses of item response theory and the testlet concept in the measurement of psychopathology. *Psychological Methods, 1,* 81–97.

Widaman, K. F., & Thompson, J. S. (2003). On specifying the null model for incremental fit indices in structural equation modeling. *Psychological Methods, 8,* 16–37.

Received September 3, 2004
Revision received October 13, 2004
Accepted October 14, 2004 ■