
LETTER TO THE EDITOR

The role of factor analysis in construct validity: Is it a myth?

STEPHEN C. BOWDEN

Department of Psychology, The University of Melbourne, Parkville, Victoria, Australia

In their recent article, Delis et al. (2003) criticized the use of factor analysis for evaluating construct validity. Focusing on a key component of their argument, they reported a high correlation between two memory test scores in a community sample but a low correlation between the same scores in a sample of people with Alzheimer's disease. As a consequence, they argued that the presence of a "dissociation" between the two variables in the Alzheimer's sample contradicted the single-factor result derived from studies of community samples and other clinical groups: "Two variables that share a high degree of variance in normal participants . . . and thus appear to measure a unitary cognitive construct, can dissociate into two distinct functions, but only in certain homogeneous patient populations" (p. 940).

However, the evidence Delis and colleagues provide does not necessitate their conclusion about different factor structures. Nor does it lessen the value of the factor-analytic approach. Instead, their data highlight the importance of examining factor structures in more detail than is usually the case.

The observation of different correlations between two variables, in different samples, may tell us more about the impact of sampling strategies than about different trait composition of the scores in different groups (Ree et al., 1994) and is quite compatible with the assumption that the same factor structure applies in both of the groups. Since the patients with Alzheimer's disease in Delis and colleagues' study were selected on the basis of poor memory performance, perhaps including poor delayed memory scores, their study displays a design feature sometimes termed "criterion contamination" (Sackett et al., 2000). As a consequence their sample may not be representative of population scores on the memory variables. This dilemma is a common, and perhaps unavoidable, confound in much clinical

research but requires special caution with regard to inferences about variables that were used for patient selection.

Methods exist to test directly the hypothesis of differences in the factor structure of cognitive function across groups and, although of direct relevance to evaluation of many clinical hypotheses, these methods are not well known in neuropsychology (see Horn & McArdle, 1992; Vandenberg & Lance, 2000; Widaman & Reise, 1997). To examine the invariance, or equality across groups of a factor structure or latent variable measurement model it is necessary to examine aspects of the factor analytic model not available from principle components analysis (PCA).

Even the more detailed confirmatory factor analytic model usually is only reported in terms of the familiar factor loading matrix, the residual (or error) variance-covariance matrix and the matrix of variances and covariances between the latent variables (or factors). A full measurement model also includes a vector of observed score intercepts and a vector of latent variable means (Meredith, 1993). In single group analyses the elements in these latter two vectors are usually set to zero and not reported but are important in multiple group analysis.

The factor loading matrix, together with the vector of observed score intercepts, provides information on the regression relationships between each of the observed scores and the respective latent variables. These matrices should not be confused with the latent variable variance-covariance matrix that, in standardized form, provides information on the standard deviations and correlations between the latent variables. Since these matrices represent separate components of a measurement model, it is possible for the correlations between latent variables to change across groups even though the factor structure does not change (Widaman & Reise, 1997).

Correlations between latent variables may change across groups for a variety of reasons, for example, because of unrepresentative sampling, changes in the reliability of scores, or changes in the variability of scores due to increasing disability on the one hand, or floor effects on the other. Correlations

Reprint requests to: Stephen Bowden, Ph.D., School of Behavioural Sciences, The University of Melbourne, Parkville, Victoria, 3010 Australia. E-mail: sbowden@unimelb.edu.au

between scores also may change because of different trait composition of the observed scores in different groups. It is only this latter result that requires the inference of different factor structures in different groups (Meredith, 1993).

Without exploring all components of the measurement model across groups, namely examining measurement invariance, it is difficult to make sense of differences in a correlation between two variables measured in separate groups, or to interpret seemingly different PCA results. A recent examination of measurement invariance in a large, diagnostically homogenous group of alcohol dependent patients, many of whom displayed neurological signs of Wernicke-Korsakoff syndrome, supported the generality of the factor structure derived from healthy community samples and concluded that a distinction between immediate and delayed memory was not necessary (Bowden et al., 2001).

To advance our understanding of construct validity we need to make the best use of a variety of approaches, including detailed examination of factor analytic or latent variable models. However, robust latent variable analysis requires large samples. This often poses logistic difficulties in clinical research. Collaborative research between centers may facilitate acquisition of the samples required for such analysis.

REFERENCES

- Bowden, S.C., Ritter, A.J., Carstairs, J.R., Shores, E.A., Pead, J., Greeley, J.D., Whelan, G., Long, C.M., & Clifford, C.C. (2001). Factorial invariance for combined WAIS-R and WMS-R scores in a sample of patients with alcohol dependency. *Clinical Neuropsychologist*, *15*, 69–80.
- Delis, D.C., Jacobson, M., Bondi, M.W., Hamilton, J.M., & Salmon, D.P. (2003). The myth of testing construct validity using factor analysis or correlations with normal or mixed clinical populations: Lessons from memory assessment. *Journal of the International Neuropsychological Society*, *9*, 936–946.
- Horn, J.L. & McArdle, J.J. (1992). A practical and theoretical guide to measurement invariance in aging research. *Experimental Aging Research*, *18*, 117–144.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, *58*, 525–543.
- Ree, M.J., Carretta, T.R., Earles, J.A., & Albert, W. (1994). Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formulas. *Journal of Applied Psychology*, *79*, 298–301.
- Sackett, D.L., Straus, S.E., Richardson, W.S., Rosenberg, W., & Haynes, R.B. (2000). *Evidence-based medicine: How to practice and teach EBM* (2nd ed.). Edinburgh: Churchill Livingstone.
- Vandenberg, R.J. & Lance, C.E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods*, *3*, 4–69.
- Widaman, K.F. & Reise, S.P. (1997). Exploring the measurement invariance of psychological instruments: Applications in the substance abuse domain. In K.J. Bryant & M. Windle (Eds.), *The science of prevention: Methodological advance from alcohol and substance abuse research*, (pp. 281–324). Washington, DC: American Psychological Association.