Use Progress in Psychometrics to Advance Nursing Science
Revisiting Factor Analysis

Listening to patients has always been valued in nursing practice, and patient-reported outcomes are fundamental to understanding the processes of health and outcomes of interventions across all areas of nursing research. Self-report instruments used to measure biobehavioral, emotional, and social aspects of health (e.g., pain, anxiety, physical functioning, self-efficacy, resilience) are a systematic way of listening to people describe their health. Measurement based on self-report relies on psychometric theory and methods, so progress in psychometrics is important to nursing research.

Advances in psychometrics over the past 50 years have been remarkable in depth and scope, but it has been noticed that this progress is not fully reflected in contemporary behavioral research. In fact, Borsboom’s (2006) impression, stated in a fiercely titled article, was that research practices have not changed much from those used by earlier generations of scientists and that they bear “an uncanny resemblance to the psychometric state of the art as it existed in the 1950s.” His subsequent thoughtful assessment of the situation and recommendations for rectifying it are worth considering because a similar situation exists in nursing science, where suboptimal or outmoded psychometric practices also persist. Factor analysis practice is a case in point.

Factor analysis in some form has long been used to investigate and explain the interdependence of item responses by effects of a common latent (unmeasured) variable or variables. In fact, the 100th “birthday” of factor analysis was celebrated in 2004 (Cudeck & MacCallum, 2007). For a quick overview of the impressive intellectual and technical developments that advanced factor analysis over the first 100 years after publication of the seminal article of Spearman (1904), visit the celebration Web page at www.fa100.info (an easy-to-type URL), select “materials for download,” and peruse the “factor analysis genealogy” and “factor analysis timeline.” Two issues are especially important for everyday research practice involving self-report data. The first involves the distinction between principal components analysis (PCA) and factor analysis. The second involves selection of descriptive, historical methods or statistically based methods of factor analysis for use in research.

**PRINCIPAL COMPONENTS**

Scan the current contents of virtually any journal reporting results based on self-report data and you’ll find, as Borsboom (2006, pp. 426–427) noted, that truncated PCA is sometimes substituted for factor analysis. But PCA is not factor analysis at all. Factor analysis is focused on accounting for the covariances (or correlations) among responses to items, whereas PCA is a variance-focused technique that uses linear recombinations of observed variables to account for the observed variances in a set of variables. PCA is a data reduction technique that focuses on the reproduction of diagonal elements of a matrix. PCA does not analyze the off-diagonal elements that summarize interdependence relationships (Skrondal & Rabe-Hesketh, 2004, pp. 70–71). Because PCA does not contribute to understanding the structure of a test, its use for that purpose should be discontinued.

**FACTOR ANALYSIS**

Common factor analysis using principal axis factoring and other early approaches were ingenious solutions to solve the factor analytic problem, suitable for days when statistical theory for factor analysis was undeveloped and computing resources were scarce (e.g., Cliff, 1987, pp. 114–115). But, this has not been the case for over 40 years (Jöreskog, 1969; Lawley & Maxwell, 1963). It is time to stop using historical methods of factor analysis. Statistically based methods of factor analysis should replace them.
Nonstatistical methods of factor analysis are exploratory processes that involve many ad hoc procedures such as looking at the scree plot of eigenvalues associated with each factor, searching for “salient factor loadings,” crossing off factor loadings that appear small, and naming factors based on the pattern of factor loadings and content of items that survived the salience search. Although these activities support close engagement of the investigator with the data, this qualitative assessment encourages post hoc explanation of results. Sometimes, investigators try to justify use of historical methods of factor analysis by saying that little theory or prior research results are available to inform a hypothesized structure for responses to the item set. This argument is disingenuous when presented alongside rich reports of the careful work that preceded the factor analysis; thorough concept clarification, thoughtful item pool development, and assessment of content validity clearly show extensive prior theorizing that fully represents prior results.

When dimensionality is nevertheless uncertain, unrestricted statistical factor analysis (using the ML estimator, for example) can be used to identify the number of factors underlying relationships among the item responses (McDonald, 1985, p. 51). All factor loadings and unique variance estimates are estimated for a series of models with 0, 1, or more factors. (The proposition that the number of factors is 0 is equivalent to hypothesizing that the variables are uncorrelated.) The $\chi^2$ test statistic arises naturally from the estimation method. If more exploration is desired, the solution may be rotated. Rather than search for salient loadings, standard errors are used to identify factor loadings that are significant. Cautious use of standard error estimates for this purpose is needed (Cudeck & O’Dell, 1994) because multiple tests are involved (one for each factor loading, so type 1 error rate should be controlled), and standard errors for loadings of equal value may be substantially different (implying naturally that size of the loading is not an indicator of salience after all).

In general, confirmatory factor analysis is to be encouraged because it allows the ideas of the investigator about the interdependence of responses to the item set to be put to the test. The number of underlying dimensions, patterns of factor loadings, correlations or covariances among factors (depending on how the factors are scaled), and the unique variances can be evaluated within groups (e.g., Jöreskog, 1969; McDonald, 1985) and across groups (using covariance matrices; Jöreskog, 1971; also see Cudeck, 1989, for a general discussion of scaling issues) with a wide range of estimators suitable for use in a variety of distribution situations (Browne, 1984) and for items with a variety of response formats (e.g., dichotomous, polytomous; Muthén, 1984). Either hypothesis testing or model comparison approaches may be used (Jöreskog, 1993). The general confirmatory factor analysis model includes special cases that permit assessment of specific forms associated with critical, recurring measurement issues (Cudeck & MacCallum, 2007; Hoyle, 2012). Some of these include the model for unidimensionality and the bifactor model, which may be used to evaluate assumptions on item response theory models (Gibbons & Hedeker, 1992) and models for measurement equivalence in multiple populations (Millsap, 2011; see Sousa, West, Moser, Harris, & Cook, 2012, for an example). As with any complex statistical model, there are challenges in the use of factor analytic models. But because factor analytic models are critical in answering many measurement-related questions in nursing science, we should know the models and the logical and technical issues involved in estimating and interpreting them.

**Closing Comment**

Measurement is a fundamental scientific activity. Consistent use of optimal measurement practices is necessary to obtain dependable and interpretable results from scientific studies used as the basis for evidence-based practice. To ensure that voices of participants expressed via self-report instruments are heard, use advances in psychometric theory and methods.

**Susan J. Henly, PhD, RN**

Editor

henly003@umn.edu

The author has no funding or conflicts of interest to disclose.

DOI: 10.1097/NNR.0b013e318294b509

**References**


