Reliability Concepts and Methods
Special Focus Section

Susan J. Henly, PhD, RN, FAAN

O matter the area of science, fundamental concepts are inexorably linked with instrumentation used to obtain measurements of them. In nursing science (Ventura, Hinshaw, & Atwood, 1981) as in all sciences, measurements should be faithful representations of characteristics of interest that gauge quantity with minimal error and enable examination of relationships and testing of hypotheses arising from theory and past research findings. Thus, accurate, precise measurement is the bedrock of all scientific endeavors, and advances in understanding are often preceded by and result from advances in measuring techniques.

The basics of measurement science are the same across the disciplines, with variations and extensions reflecting the nature of the subject matter. Many critical constructs in nursing science are in the biobehavioral and psychosocial realms. Measurements are often obtained using self (patient)-report. Proxy reports (about child health status, obtained from parents; about patient status, obtained from nurses) and direct observation of health-related behaviors (from carefully trained researchers) are also used. All these practices rely on psychometric principles, concepts, and methods—especially those related to reliability (repeatability) of measurements.

Use of self-report on a single occasion is pervasive in nursing research (and research in related health sciences), and Cronbach’s alpha (Cronbach, 1951) is often regarded as “the” way to report reliability. Rules of thumb are invoked to determine whether reliability is “adequate.” These habits point out gaps in design of measurement studies and limited consideration of advances in psychometrics that, if used, would enhance quality of instrumentation—and minimize errors in measurement that is an impediment to advancing nursing science.

In his Presidential Address to the Psychometric Society, Prof. Klaas Sijtsma (2012) considered the future of psychometrics and asked “what can psychometrics do for psychology?” He extended his commitment to linking advances in psychometric theory and methods with substantive scientific areas when he suggested a series of papers on reliability for Nursing Research. The resulting “Reliability Concepts and Methods” Special Focus Section in this issue of Nursing Research is a moderated dialogue about reliability of measurements of biobehavioral and psychosocial constructs in nursing research. Papers were contributed by four invited groups of authors composed variously of psychometricians, nursing scientists, and statisticians. The process was “open,” with authorship of the target paper (Sijtsma & van der Ark, 2015a) and comments (Barbaranelli, Lee, Vellone, & Riegel, 2015; Gajewski, Price, & Bott, 2015; Yang & Green, 2015) known to everyone involved. The target paper was circulated to others, and their comments were likewise submitted to Drs. Sijtsma and van der Ark for the purpose of writing a wrap-up rejoinder (Sijtsma & van der Ark, 2015b). My roles as editor were, first, to invite participation and moderate the process and, now, to introduce the Special Focus Section to you.

The Special Focus Section opens with Sijtsma and van der Ark’s (2015a) presentation of fundamental definitions and methods used in reliability estimation arising from classical test theory, factor analysis, and generalizability theory; they used a simulation example to illustrate differences among the methods. In reply, Gajewski, Price, and Bott (2015) extended the discussion to include information about variability in point estimates of reliability and conditional standard errors of measurement. Barbaranelli et al. (2015) summarized their scientific work with the Self-Care of Heart Failure Index and addressed applied aspects of reliability and dimensionality assessment. They argued that dimensionality questions should be addressed prior to estimating classical test theory-based reliability estimates and discussed implications for reliability estimation when graded response options are used for measurement. Yang and Green (2015) synthesized many issues, emphasizing the bifactor model when considering the interplay of validity (in the dimensionality sense) and reliability; they urged investigators to immerse themselves in their data to come to a complex understanding of the consistency of scores in their research. In their rejoinder, Sijtsma and van der Ark (2015b) noted the depth and vitality of research and advances in reliability concepts and methods reflected in the commentaries and described a new approach to nearly unbiased estimation of reliability irrespective of dimensionality (the divisive latent class reliability coefficient).

The papers have more equations than is typical in Nursing Research. This feature is necessary for an in-depth dialogue about technical issues, however. Here are some tips for reading this information. First, technical text is a type of shorthand whose terms and operators allow clear and concise presentation of the flow of ideas. Authors have defined all the terms; take notes as you read, and watch for “English names” of terms. For example, recognize that Sijtsma and van der Ark (2015a) use the symbol $X_j$ to stand for sumscore and $X_j$ to mean the sumscore for a specific set of $j$ items. Know the...
Greek letters so the terms can be named. Both scalar and matrix representations are used in equations; remember that in everyday scalar arithmetic, $\sum$ means to “add up” the terms that follow and that in matrix arithmetic, bolding is used to indicate matrices and vectors. Second, arguments are made in parallel in the text and equations; try to follow both. Last, practice! Like any skill, the ability to read technical material improves with repetition. The payoff is greater understanding and performance—in this case, in the planning, conduct, and critique of our scientific endeavors.

“Thank you” to all the authors of papers in the Special Focus Section who have so generously shared their expertise! They have provided a treasure trove of information about reliability concepts and methods for nursing and health sciences research. Study and apply the information they provided—move beyond reflexive use of what Sijtsma (2012, p. 8) aptly called “the common lore of dos and don’ts” about psychometric methods, including reliability estimation. Be informed and thoughtful in merging substantive issues and methodological knowledge to resolve measurement problems. Explain your perspectives and create reasoned arguments to support your approach to reliability assessment.

References


Susan J. Henly, PhD, RN, FAAN, is Editor of Nursing Research. The Editor has no conflicts of interest to disclose.