

March 2004: Scales of Measurement (Rule 1.13)

Rules of the month are numbered in accordance with the numbering in the book. Thus, Rule 1.1 refers to the first rule in Chapter 1. And so on. These comments do not repeat the material in the book but highlights and amplifies it. A rule is stated as found in the book and then discussed.

“Be flexible about scale of measurement determining analysis.” (Rule 1.13)

Further Comments on the Rule

The issue of scale of measurement continues to be discussed by researchers, especially in the social sciences. In this note I want to make two comments.

A recent book by Cliff and Keats (2003) was inspired by the authors' careful analysis of scales of measurement in psychometrics. The authors voice strong disagreement with the current work in item response theory that does not follow Rasch modeling; basically, that all items in a test have the same difficulty. They argue that Likert-like scales (ordinal scales) should only be analyzed using ranking methods. They argue strongly that to assume an interval scale is invalid. For example, going from 1 to 2 may not be the same as going from 4 to 5 on a Likert scale with 5 cut points. They then “put their money where their mouth is” and present alternative ways of analyzing such data. As could be expected they use rank statistics adjusted for ties. In an earlier book Cliff (1996) in the same vein Cliff writes, “For at least two generations, statistical methods in the behavioral sciences have been dominated by normal-based linear models, analysis of variance, and regression, and their relatives...I have become increasingly concerned that they often are not the best possible ones for our purposes. They overestimate the quality of our data; they answer questions that are not really the ones we want to ask; they are overly sensitive to the failures of assumptions that are frequently unrealistic.” (Cliff, 1996), page ix). It is unlikely that the authors will convince the majority of psychometric researchers but at least they need to be heard.

A second comment about scales is from a paper by Fischer (2003). He points out that differences on an interval scale form a ratio scale. An interval scale has no true origin; a ratio scale has. An example of an interval scale is temperature, an example of a ratio scale is body mass. Fischer gives the example that a change of 10 degrees is twice as much as a change of 5 degrees—no matter what the scale. Thus, the somewhat surprising result that one can go from one scale to another without additional assumptions.

It is not clear if there are general rules for going validly from one scale to another. For example, is there a situation where a function of an ordinal scale produces an interval scale?

These two points illustrate again that the typology of scales of measurement is not yet based on a “closed dogma.”

References

Cliff, N. (1996). *Ordinal Methods for Behavioral Data Analysis*. Lawrence Erlbaum Associates. Mahwah, N.J.

Cliff, N. and Keats, J.A. (2002). *Ordinal Measurement in the Behavioral Sciences*. Lawrence Erlbaum Associates, Mahwah, NJ.

Fischer, G.H. (2003). The precision of gain scores under an item response theory perspective: a comparison of asymptotic and exact conditional inference about change. *Applied Psychological Measurement*. **27**: 3-26.