

---

**Book Review**

**Reasons for reconsidering quantitative research based on the use of Likert scale and other social science data sets**

---

Michell, J. (1999). *Measurement in psychology: Critical history of a methodological concept*. United Kingdom: Cambridge University Press. Price: \$99.00 (hardback)

A sense of common cause largely motivates my review of this book. Like the author, Joel Michell, I am increasingly inclined to the view that there is "an error in scientific method fundamental to quantitative psychology" (p. xi), one that generalises across the social sciences (although I hadn't glimpsed the full extent prior to looking through this book).

In my current capacity as statistical adviser to a Faculty of Education, I encounter an ocean of research that washes between the shores of qualitative (e.g., case studies) and quantitative (e.g., attitude) research. The latter is largely focused on the use of survey instruments based on the Likert scale. The "rule of thumb" seems to be that a continuum of response categories such as "strongly disagree, disagree, undecided, agree, strongly agree", can be recoded numerically as 1, 2, 3, 4, 5. Usually, the resulting set of numbers is analysed as if possessing ratio or at least equal interval properties. That is, participant responses signifying, for example, agreement vs. disagreement typically are analysed by computing central tendency and range (descriptive statistics) and then examining significant differences between conditions by means of quantitative statistical procedures (inferential statistics). To my mind, this approach to the analysis of human data is built on a foundation of sand, the sand being the inability of such numerical scales to amount to more than a numerical representation of an ordinal, at best, set of response categories.

At this point, the book under review gets a mention. The author's major assertion is that the unfounded assumptions I've noted in relation to the use of the Likert scale in educational research are present throughout the social sciences.

Michell begins by defining quantitative science as seeking to discover Euclidean relationships between changes in the intensity of attributes (e.g., length) that are additive and can be expressed in terms of a ratio (i.e., M2 is twice as long as M1; length x breadth = area). In these terms, the goal of measurement can be restated as the process of discovering ratios. He contrasts the approach of psychological measurement, including scales based on item

response theory, as making the error of actively ignoring the difficulties inherent in demonstrating that attributes such as ability, self-esteem, etc., are additively related or can be expressed in terms of ratios.

Michell provides a lucid discussion of how this error became entrenched in psychology by way of Newtonian physics, Helmholtz, Fechner, Stevens, behaviourism, and positivist philosophy. The error is depicted as stemming from attempts by experimental psychologists to bridge the gap between psychophysical and physical attributes, attempts that culminated in Stevens' (1946) redefinition of measurement as the "assignment of numerals to objects and events according to rules" (p. 677). Stevens' redefinition amounted to a claim to be able to measure psychological attributes. This redefinition allowed Stevens to ask subjects to judge when one sound was twice as loud as another and then to treat these direct judgements as a sufficient reason for proposing that judgements of loudness could be represented in terms of a ratio scale. On this matter, Michell has quoted Bertrand Russell (1919) as saying, "The method of 'postulating' what we want has many advantages; they are the same as the advantages of theft over honest toil" (p. 71).

Michell illustrates the paradigmatic quality of this definition of measurement in present-day psychology by documenting psychology's indifference to potentially error-correcting developments. Michell has cited Luce and Tukey's (1964) theory of conjoint measurement as a case in point. This theory provides a basis for comparing objects in terms of the ratios between attributes (e.g., the mass and density of objects made up of differing materials). He comments that its application to psychological attributes (e.g., the ability and motivation of persons when performing differing tasks) has the potential to test the extent to which such attributes are quantitative and measurable (additive). Yet conjoint theory does not rate a mention in mainstream commentaries or textbooks on test theory. The author concludes that to say that the relevant theoretical attributes of psychology are quantitative is to propose an empirically testable hypothesis, one that has yet to be properly tested.

The starting point for this review was a sense of common cause concerning the author's misgivings about the application of measurement theory to human data, and, more specifically, its application to Likert scale data. In answer to the unspoken question, what does one do instead, this final paragraph presents an approach to Likert scale data sets based on a decision to treat them as embodying counts or frequencies rather than as possessing interval properties. A logical starting point is to simplify the raw responses to particular items by collapsing items into dichotomous categories (e.g., the Yes/No and Agree/Disagree categories identified by Babbie, 1995). The act of dichotomising items renders them linear (i.e., creates the dummy variables outlined by Tabachnick & Fidell, 1995). This act then facilitates contextual (e.g., What response is this item eliciting?) and graphical (e.g., sorted by percentage) exploration. A further step might be to undertake formal analyses based on item or summary scores. For example, one might examine the effect of demographic or other variables per item by way of contingency analysis (chi-square or continuity corrected chi-square). Again, a summary score

might be computed: With scores per item equal to 0 or 1, Likert scale items can be added. The resulting distribution of discrete scores might be examined graphically and the effect of "independent" variables analysed by using nonparametric or parametric procedures, depending upon the normalcy of the distribution. This essentially conservative approach to Likert scale analysis offers a treatment of human data that takes Michell's critique into account insofar as it avoids implausible assumptions about the properties of the data set and thus increases the likelihood of obtaining reliable and meaningful (valid) outcomes.

### **Address for Correspondence**

Dr Peter Grimbeek  
Faculty of Education  
Griffith University  
Nathan Queensland 4111  
Phone: (07) 3875 5629  
E-mail: p.grimbeek@mailbox.gu.edu.au

### **References**

- Babbie, E. (1995). *The practice of social research* (7th ed.). USA: Wadsworth.
- Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement, *Journal of Mathematical Psychology*, 1, 1-27.
- Russell, B. (1919). *Introduction to mathematical philosophy*. London: Routledge.
- Stevens, S. S. (1946). On the theory of scales of measurement. *Science*, 103, 667-80.
- Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). USA: Harper Collins College Publishers.