



On measurements and their quality. Paper 4: Verbal anchors and the number of response options in rating scales



Jason W. Beckstead*

University of South Florida College of Nursing, USA

ARTICLE INFO

Article history:

Received 12 September 2013

Accepted 13 September 2013

Keywords:

Measurement theory

Psychological scaling

Likert scales

Intervention research

ABSTRACT

This is the last in a short series of papers on measurement theory and practice with particular relevance to intervention research in nursing, midwifery, and healthcare. Understanding how it is that people respond to the questions posed by researchers is fundamental to progress in the social and health sciences. For decades methodologists in psychology, marketing, education, and survey research have studied this issue. In this paper I review this diverse empirical literature to synthesize basic principles for creating rating scales which can reduce measurement error and increase the quality of resulting data. After introducing a theoretical framework known as the cognitive aspects of survey methods (CASM), I review the fundamentals of psychological scaling theory and discuss how it has been used to study the meanings of verbal response options and provide an illustration of how the quality of measurements may be influenced by our choice of the verbal phrases we present as response options. Next, I review the research on the optimal number of response options to use in various measurement situations and how verbal and numeric anchors can combine to influence data quality. Finally, I summarize the issues covered and present recommendations for best practice when creating and using rating scales in research.

© 2013 Elsevier Ltd. All rights reserved.

1. Models of the question answering process

As researchers we are often interested in the relationships between what, speaking broadly, may be called external or environmental conditions and the internal or psychological sensations (e.g., pleasure, confidence, satisfaction, regret, discomfort, or anger) that they may produce. Dependent upon our specialization, the environmental conditions that interest us may include assorted clinical therapies and prescribed treatment regimens, benefits and hardships associated with living under various governments, or the degree of difficulty inherent in various decision-making situations. Regardless of our

specialty, when we rely on individuals to quantify their own sensations for us to study, the veracity of our conclusions is limited by the accuracy of their judgments. It is therefore in our best interests to make the respondent's task of providing us with information as unambiguous and as easy to complete as possible.

Since the 1980s, there has been considerable interest among social scientists in how people respond to the questions posed by researchers. This line of inquiry is referred to in the literature as the cognitive aspects of survey methods (CASM). Investigators in the CASM tradition have made considerable progress in illuminating the cognitive and communicative processes underlying survey responding. Drawing on psychological theories of language comprehension, memory and judgment, CASM researchers have formulated models of the question answering process and tested them in laboratory experiments and in surveys employing random assignment of respondents to different question formats (Schwarz, 2007).

* Correspondence to: University of South Florida College of Nursing, 12901 Bruce B. Downs Boulevard, MDC 22, Tampa, FL 33612, USA.
Tel.: +1 813 974 7667; fax: +1 813 974 5418.

E-mail address: jbeckste@health.usf.edu.

Most experts in this field agree that there are benefits to modeling the act of answering questions into sequential stages based on the specific cognitive processes involved. Generally speaking, there is an initial acquisition or question comprehension stage, followed by a judgment formation stage, and finally a translation stage in which the respondent's internal judgment is translated or mapped onto the response options provided by the researcher. More complex models have been proposed to address specific aspects within this general framework (see [Chessa and Holleman, 2007](#); [Shulruf et al., 2008](#); [Tourangeau, Rips & Rasinski, 2000](#)). An excellent review of the developments in CASM is provided by [Lietz \(2010\)](#).

In this article, the last in the series on measurements and their quality, we are concerned with the final stage of the question answering process, namely, the mapping of internal judgments onto available response options. Even when the wording of our survey questions or attitude statements accurately represents what we intend to measure, the quality of the resulting data may be compromised due to translation errors introduced by the set of response options we provide to the participant or respondent. The remainder of this article is organized as follows. First I introduce psychological scaling theory and discuss how its methods have been employed to construct scales that quantify the meaning of words and phrases commonly used for expressing the degree of abstract concepts such as frequency, intensity, or probability. Second, I work through an illustration of how the quality of measurements may be improved (or hampered) by our choice of verbal phrases presented as response options. Third, I review the research relevant to deciding on the number of response options to use in various measurement situations. Fourth, I discuss how verbal and numeric anchors can combine to influence data quality. Finally, I summarize the issues covered and present recommendations for best practice when creating and using rating scales in research.

2. Determining the meaning of verbal response options: applying psychological scaling theory

Psychological scaling methods are procedures for constructing scales for the quantification of psychological attributes. Their origins can be traced to the field of psychophysics in the 19th century. Psychophysics requires both the measurement of physical attributes of stimuli (e.g., luminance, weight, or volume) and the quantification of sensations (e.g., brightness, heaviness, and loudness, respectively) in order to study the relationships between physical and psychological magnitudes. In the 1920s Louis Leon Thurstone pointed out how many of the scaling methods of psychophysics could be used to accurately measure the psychological attributes of stimuli that have no relevant measurable physical correlates, such as, for example, attitude statements or nationality preferences ([Thurstone, 1927a,b, 1928, 1929](#)).

[Torgerson \(1958\)](#) succinctly summarizes the basic tenets of Thurstone's theory of scaling as follows. First, each stimulus when presented to an observer gives rise to a reaction or *discriminal process* which has some value on

the psychological continuum of interest. Second, because of momentary fluctuations in the observer, a given stimulus does not always activate the same discriminial process, but may activate one with a higher or lower value on the continuum. If we present any stimulus to an observer a large number of times we can imagine a frequency distribution along the psychological continuum of discriminial processes associated with that stimulus. The frequencies with which discriminial processes are associated with any given stimulus are assumed to form a normal distribution on the continuum. Third, the discriminial process (reaction) most often associated with a given stimulus is defined as the *modal discriminial process*; the scale value of the stimulus on the psychological continuum is then taken as the value of the modal discriminial process. The standard deviation of the reactions associated with a given stimulus is referred to as its *discriminal dispersion*.

Unfortunately, the observer cannot report the values of the discriminial processes activated and so these cannot be used directly to define scale values of stimuli. Thurstone's original contribution was to show how an observer's *judgments of the differences between stimuli* could, however, be used indirectly to scale stimuli. This result has become known as Thurstone's Law of Comparative Judgment. Various experimental procedures have been developed to deduce the scale values and dispersions of stimuli from judgments of the differences among stimuli presented. These include the method of paired-comparisons, the method of successive intervals, the method of successive categories, and the method of magnitude estimation (see [Torgerson, 1958](#) for review).

Of particular interest, these methods have been employed to construct scales that quantify the meaning of words and phrases commonly used for expressing the degree of abstract concepts such as frequency, intensity, or probability. For example, if we wished to construct a scale of phrases for patients to use when expressing how frequently they are bothered by various symptoms, we could start by assembling a list of words and phrases that we believe express varying degrees of frequency; our list might include the phrases: *seldom, on occasion, sometimes, once in a while, and often*. Next, we could form all ($5 \times 4 / 2 = 10$) pairs of phrases and then ask several individuals to select the phrase in each pair that indicates the greater degree of frequency. Finally, the proportion of times each phrase is chosen over another is calculated. These proportions serve as the basis for defining the position, precision and overlap of the phrases on a psychological scale of frequency.

Sporadically over many years, investigators have used scaling methods to obtain psychological scale values for words and phrases that express degrees of frequency and amount ([Bass et al., 1974](#); [Lilly, 1968](#); [Pohl, 1981](#); [Schriesheim and Schriesheim, 1978](#); [Simpson, 1944, 1963](#); [Stone and Johnson, 1959](#)), intensity ([Bashaw and Anderson, 1968](#); [Cliff, 1959](#); [Dudek, 1959](#); [Mosier, 1941](#)), evaluation ([Jones and Thurstone, 1955](#); [Lodge et al., 1976](#); [Spector, 1976](#)), and probability ([Howe, 1963, 1969](#); [Lichtenstein and Newman, 1967](#); [Reagan et al., 1989](#)). Some of the common findings from these studies provide

Download English Version:

<https://daneshyari.com/en/article/7515990>

Download Persian Version:

<https://daneshyari.com/article/7515990>

[Daneshyari.com](https://daneshyari.com)