Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: do they matter?

Ellen Greenberger*, Chuansheng Chen, Julia Dmitrieva, Susan P. Farruggia

*Department of Psychology and Social Behavior, School of Social Ecology, University of California,
Irvine, CA 92697-7085, USA*

Abstract

Some researchers contend that the Rosenberg Self-Esteem Scale taps two dimensions of self-image, whereas others argue that the two dimensions (positive and negative) are merely an artifact of item wording. To directly test these competing views, we had 741 ethnically diverse university undergraduates take one of three versions of the 10-item Rosenberg Scale: the original version comprised of five positively worded and five negatively worded items, or one of two alternative versions comprised of 10 positively worded or 10 negatively worded items. Analyses indicated that the original version fit a two-factor model, whereas the reworded versions generally fit a one-factor model. All three versions had high validity for different ethnic groups, but the revised-positive version had less overlap with a measure of depression, and both revised versions had less overlap with a measure of self-deception.
© 2003 Elsevier Ltd. All rights reserved.

Keywords: Self-esteem; Dimensionality; Item-wording; Construct validity; Confirmatory factor analysis; Rosenberg Scale

The widely used Rosenberg Self-Esteem Scale was conceptualized by its author as a single-factor scale with scores ranging along a continuum of low self-esteem to high self-esteem (Rosenberg, 1965). The "actual" or empirical factor structure of this scale has been the target of debate, however, for over 30 years (Owens, 1994). Several researchers who conducted factor analyses of the 10-item scale have suggested that the scale reflects a two-dimensional construct, comprised of positive and negative images of the self (Bachman & O'Malley, 1986; Goldsmith, 1986; Kaplan & Pokorny, 1969; Owens, 1993). In these and other studies, the five positively-worded items (e.g. "I feel that I have a number of good qualities") loaded on one factor, referred

---

to variously as "positive self-esteem," "positive self-worth," and "positive self-image," whereas the five negatively worded items (e.g. "At times I feel that I am no good at all") loaded on a separate factor, referred to variously as "negative self-esteem," "self-derogation," "self-deprecation," and "negative self-image."

In support of their contention that Rosenberg's Scale taps two distinct dimensions of self-esteem, some researchers have argued that the positive and negative dimensions of this measure lead to different outcomes and are influenced by different experiences. For example, Owens (1994), based on a longitudinal study using an eight-item version of Rosenberg's scale (with slight re-wording and a slightly different response-continuum), examined the reciprocal effects between three scores derived from the scale [total or Global Self-Esteem, "Self-Deprecation" (i.e. the subscale comprised of negatively worded items), and "Positive Self-Worth" (i.e. the subscale comprised of positively worded items)] and measures of grades, depression, and self-reported delinquency. Among the results of the study, self-deprecation scores were substantially more related to depression in both causal directions, i.e. self-deprecation led to increases in depression over time, and depression also led to increases in self-deprecation, than were scores on either the positive self-worth subscale or total (global) self-esteem scores. In a related study, using samples of British college students and adults, Sheasby, Barlow, Cullen, and Wright (2000) found evidence favoring a two-factor structure of the Rosenberg Scale, with negatively worded items all loading on one factor. However, one positively worded item emerged on the "negative items" factor for both samples, and a second positively worded item loaded on the factor for the adults in this study. In a recent, cross-cultural study (Farruggia, Chen, Greenberger, Dmitrieva, Tally, & Macek, unpublished manuscript), a single-factor solution for the Rosenberg Self-Esteem Scale did not fit the data for adolescent samples in any of the four countries (US, Czech Republic, China, and Korea), but a modified, nine-item, two-factor model fit the data adequately. Although the Owens study, in particular, appears to make a good case for the nuances of self-esteem that may be uncovered when both positive and negative dimensions of self-esteem are considered, the earlier mentioned studies did not test the possibility that the apparent bi-factorial structure of the Rosenberg scale may be an artifact of item-wording.

The possibility that the presumed two-factor structure of Rosenberg's Self-Esteem Scale is an artifact of item-wording has been proposed by several researchers (e.g. Carmines & Zeller, 1979; Hensley & Roberts, 1976; Marsh, 1996; Tomas & Oliver, 1999). For example, the two factors that emerge may merely reflect issues of response set or response bias. Thus, positive and negative self-esteem factors might result from respondents' tendency to agree with positively worded statements about the self and disagree with negatively worded statements about the self, or from more general tendencies to engage in "yea-saying" or "nay-saying" (Couch & Keniston, 1960). Marsh (1996) took a confirmatory analysis approach to examining the dimensionality of a seven-item version of the Rosenberg scale (four positively worded items and three negatively worded items) and compared the validity coefficients of Global (total) Self-Esteem, Positive Self-Esteem and Negative Self-Esteem) with 13 measures relevant to the validity of these three measures. Marsh's analyses of Rosenberg's scale reflected a single, substantively meaningful factor and the existence of a method effect, primarily due to the negative items, that complicates interpretation of the total or global Self-Esteem score. Marsh argued that agreeing with negative items such as "I feel I do not have much to be proud of," but especially *disagreeing with* or negating negatively worded items, adds a degree of cognitive complexity to the task of responding to a questionnaire. In