ALBERT KOZMA AND M. J. STONES

SOCIAL DESIRABILITY IN MEASURES OF SUBJECTIVE WELL-BEING: AGE COMPARISONS*

(Received 18 January, 1987)

ABSTRACT. The present investigation evaluated the relationship between measures of psychological well-being and social desirability in three age groups: 21—40, 41—60, and 61—82 year-old samples. Data on 330 people, consisting of community and clinical groups, yielded high correlations between three measures of well-being (the MUNSH, the LSI-Z, and the PGC) and the Edwards Social Desirability Scale for all age groups, but only moderate ones between well-being scales and the Marlow-Crown Social Desirability Scale. Partial correlations between well-being measures and an external criterion of happiness, controlling for social desirability, failed to improve on the 0-order criterion/well-being relationship. Controlling for social desirability, therefore, does not enhance the construct validity of well-being scales in adult populations at any age. These results, combined with those on the factor structure of scale totals and on the discriminant validity of the well-being measures, suggest that the high 0-order correlations between measures of well-being and the Edwards scale are more readily attributed to content similarity between the Edwards scale and measures of well-being than to a social desirability response bias in well-being measures.

An important consideration for any one employing self-report questionnaires is the extent to which such measures are free from response bias. Until recently, this problem had escaped the attention of researchers working in the area of psychological well-being. However, a study by Carstensen and Cone (1983) has already had enough impact on the area to bring about a major change in this state of affairs (Kozma and Stones, 1985). Carstensen and Cone found that two of the most frequently used measures of psychological well-being, the Philadelphia Geriatric Center Morale Scale (PGC) (Lawton, 1972) and the Life Satisfaction Index-B (LSI-B) (Neugarten et al., 1961) correlated with the Edwards Scale of Social Desirability (ESDS) (Edwards, 1957) 0.70 and 0.58, respectively. For the PGC this value was higher than that representing its relationship to the LSI-B ($r = 0.64$). Not surprisingly, Carstensen and Cone suggested that well-being scales need extensive refinements to free them from the influence of a social desirability response bias.

At first glance, these results do appear to pose insurmountable
problems for the construct validity of well-being measures. However, before large scale revisions are undertaken of otherwise serviceable instruments, an alternative interpretation of the Carstensen and Cone data needs to be explored. Such a strategy is particularly advisable if one considers the experiences that researchers have had with the response bias problem in other areas of psychology. As early as 1965, Block argued that item content rather than response bias was the basic determiner of self-report inventories. Evidence in support of Block's position has persisted up to the present time (Rorer, 1965; Wiggins 1973; McCrae and Costa, 1983; to list only a few studies).

Our basic position is that measures of well-being are primarily measures of content and not measures of response set or style. We are attributing the high correlations between the ESDS and the LSI-B and PGC, in large part, to content overlap among the three scales. The Edwards scale contains too many items measuring a well-being/pathology dimension (Crowne and Marlowe, 1960; Kozma and Stones, 1985; McCrae and Costa, 1985) to serve as a good measure of response bias. If the ESDS is administered to populations relatively free of pathology, such as the Carstensen and Cone (1983) sample, failure to endorse negative items, and the tendency to endorse positive ones, is more likely to reflect an absence of psychopathology than an attempt at "faking good". In psychometric terms, subjects respond more to content than to response set or style.

Kozma and Stones (1987) discuss three general strategies for determining the relative merits of the two positions. The first one involves the inclusion of a social desirability measure, such as the Marlow-Crown Social Desirability Scale (MCSDS), in which content overlap with well-being scales is less extensive (Kozma and Stones, 1987). Substantially lower correlations between measures of well-being and the MCSDS than between well-being scales and the ESDS would imply that content overlap is responsible for the large ESDS/well-being coefficient.

The second approach involves a statistical control for a social desirability bias. If the "parcelling out" of such an effect from well-being measures (i.e., the removal of variance due to social desirability) led to an increase in the correlation coefficient between well-being measures