

Subjective Well-Being: The Convergence and Stability of Self-Report and Non-Self-Report Measures

Ed Sandvik, Ed Diener, and Larry Seidlitz

Abstract The validity of self-report measures of subjective well-being (SWB) was examined and compared with non-self-report measures using a sample of 136 college students studied over the course of a semester. A principal axis factor analysis of self- and non-self-report SWB measures revealed a single unitary construct underlying the measures. Conventional single-item and multi-item self-report measures correlated highly with alternative measures, with theoretical correlates of SWB, and with a principal axis factor underlying five non-self-report measures of well-being. Comparisons of family versus friend informant reports demonstrated the considerable cross-situational consistency and temporal stability of SWB. Evidence of the discriminant validity of the measures was provided by low correlations of the various SWB measures with constructs theoretically unrelated to well-being. It was concluded that conventional self-report instruments validly measure the SWB construct, and that alternative, non-self-report measures are useful for providing a comprehensive theoretical account of happiness and life satisfaction.

The present article provides evidence concerning the construct validity of various self-report and non-self-report measures of subjective well-being (SWB). The analyses serve several major purposes: (a) to assess the convergence of conventional single- and multi-item self-report measures of SWB with each other and with alternative measures of the construct, (b) to provide evidence of the unitary nature of the well-being construct, (c) to assess the convergence of conventional and alternative SWB measures with a number of theoretical correlates of well-being, (d) to assess the consistency of individual differences in SWB over time and across situations, and (e) to provide evidence of the discriminant validity of conventional and alternative well-being measures. These analyses are necessary to clarify what traditional well-being scales measure as well as to elucidate the nature of individual differences in SWB.

E. Diener (✉)

Department of Psychology, University of Illinois, Urbana-Champaign, Champaign, IL 61820, USA
e-mail: eddiener@uiuc.edu

In recent years interest in SWB has increased both in the frequency with which well-being is assessed in research and survey projects, and in the diversity of research interests which have found it to be a useful construct. These include sociology (Phillips, 1967), geriatric research (McNeil, Stones, & Kozma, 1986), clinical psychology (Cameron, Titus, J. Kostin, & M. Kostin, 1973), personality trait (Diener, 1984), and cognitive affect (Schwarz & Strack, 1991). Although few concepts have more intuitive reality and relevance to life than personal happiness, it is a construct which presents unique challenges to its assessment.

The single most common form of well-being assessment has been the collection of self-reports. A variety of such well-being measures have commonly been used, such as the Delighted-Terrible (D-T) Scale (Andrews and Withey, 1976) and the Fordyce Happiness Measure (1977). These two scales are comprised of single-item, face-valid inquiries into personal well-being: "How do you feel about how happy you are?" (response possibilities ranging from "delighted" to "terrible"), and "In general, how happy or unhappy do you usually feel?" (response possibilities ranging from "extremely happy" to "neutral" to "extremely unhappy"). Regardless of the simple face-validity of such inquiries, there are reasons for not automatically accepting responses to them as valid.

There has always existed considerable skepticism in the social sciences concerning the validity and interpretation of self-report data in general (e.g., Nisbett & Wilson, 1977), and there are reasons for more rather than less skepticism in the case of reports of personal well-being. It is possible that society imposes strong norms concerning the social desirability of happiness that are less of a problem for other constructs such as extraversion. To claim to be happy may be the ultimate assertion of success in our society, and to admit unhappiness could be the single greatest summary of failure in life that an individual could concede. The potential for social desirability artifacts provides good grounds for initial skepticism in the interpretation of self-reports of well-being. Another possibility for artifactual interpretation of self-reports, the influence of current mood, has been addressed elsewhere (Diener, Sandvik, Pavot, & Gallagher, 1991).

Demonstrations that self-reports of well-being are not simply interpretable in various artifactual terms, although important, do not establish the construct validity of such measures. If SWB research findings are to be interpretable, it is also necessary to determine the extent to which well-being shows itself to be a unitary construct which converges across a variety of measures. Because no one type of SWB measure by itself can be considered the touchstone of the well-being construct, the nomological network in which well-being exists must be elaborated in terms of the construct's converging and discriminating relations with affiliated constructs (Cronbach & Meehl, 1955).

Most previous research on construct validity in well-being assessment has focused on the relation of self-reports of well-being to other self-reports of well-being or related self-reports of personality traits. For example, in his recent review of his Happiness Measure, Fordyce (1988) reports on the correlations between his well-being measure and a host of similar measures, such as the D-T Scale (Andrews and Withey, 1976), the Affectometer 2 (Kammann & Flett, 1983), as well

as measures such as the Beck Depression Inventory (Beck, 1978), the Mood Survey (Underwood & Froming, 1980), the Institute for Personality and Ability Testing (IPAT) Depression Scale (Krug & Laughlin, 1970), and omnibus inventories such as the Minnesota Multiphasic Personality Inventory (Hathaway & McKinley, 1951), the Personality Research Form (Jackson, 1967), and the California Personality Inventory (Gough, 1957). Similarly, in their classic study, Andrews and Withey (1976) report on the intercorrelations of a host of self-report scales of well-being and personality. They find a high concordance among many different phrasings of self-report well-being scales, and based on such comparisons, provide estimates on the true well-being variance accounted for in measures such as their D-T happiness measure. Although such information is important, the cautious researcher may hesitate to incorporate happiness measures in his or her research or to interpret findings without greater evidence that scores on self-report happiness scales correspond with something other than similar self-reports. Cogent arguments have been voiced against naive interpretations of self-reports as such (Nisbett & Ross, 1980; Nisbett & Wilson, 1977). Although such concerns have been considered in abstract (Dulany, 1968; Ericsson & Simon, 1980), it is clearly desirable for well-being researchers to demonstrate convergence between the commonly used, traditional one-time self-report well-being measures and other diverse sources of information which would not be subject to the same limitations or biases as traditional self-report inventories.

In their recent review, Hogan & Nicholson (1988) note that, although the idea of construct validity is familiar, it has not been practiced as thoroughly as it has been preached. The result, they claim, has been a rehash of what they term “1960s-vintage criticisms” in assessment. The concern for construct validity arose out of a need to develop scientifically rigorous justification and procedures for dealing with the measurement of inferred states, and a skepticism about a naive interpretation of self-reports.

Unfortunately, there are unique difficulties involved in attempting to collect diverse forms of evidence in the case of subjective well-being. Andrews and Withey (1976) stated the situation as follows:

Evaluating measures of people's feelings about their lives presents major problems because there seem to be no clear and directly observable phenomena that can serve as criteria. People's behavior and the conditions under which they live, while related to their perceptions of well-being, are influenced by many factors in addition to their feelings and hence are not appropriate criterion variables. It is the feelings themselves that we seek to measure. Although very real and important to the people who hold them, they are internal and inherently unobservable phenomena. (p. 215)

Thus, special care must be used in selecting objective indices of SWB, and caution must be exercised in their interpretation.

Several studies, however, have attempted to go beyond self-report information on well-being. One such attempt was by Andrews and Withey (1976). In addition to collecting self-report information, they gathered reports from two or three individuals who knew the subject. They found low correlations between these ratings by knowledgeable others and the D-T Scale and concluded that future collection of such information is unlikely to be useful. Recent methodological reviews have

suggested, however, that because informed others usually see individuals only in a limited number of settings, a greater number of reports than that collected by Andrews and Withey is necessary to show convergence (Moskowitz, 1986). Other studies gathering peer reports have in fact had greater success. Carroll (1952), for example, reported a 0.39 correlation between self- and other reports of well-being. Similarly, Hartmann (1934) reported a 0.34 correlation between self-reports and those of four associates, and Goldings (1954) reported a 0.38 correlation between self-reports and those of five experimenters who knew the subject moderately well.

Although each of the studies mentioned above assessed the convergence of a self-report measure of well-being with a non-self-report measure, none have attempted a broad study of the relation between commonly used traditional self-report measures of SWB and a broad spectrum of non-self-report measures. We have examined standard single-item and multiple-item self-reports, as well as various alternative strategies, including informant reports, a forced-choice measure which controls for social desirability, a memory sampling measure, expert ratings, and a longitudinal study of daily mood over 42 days. The convergence among these diverse well-being measures should provide important information on the construct of SWB, on how traditional well-being scales should be interpreted, and on how much is gained by using non-self-report strategies in addition to traditional self-report measures.

Although well-being is inherently internal and experiential, there are reasons to believe that methods other than self-report might be valid indicators of this experience. Furthermore, it does not seem that simple self-report should be automatically afforded privileged status in the assessment of private experience. For example, experts should be able to estimate participants' SWB based on information provided in a written interview. Although based on self-report, the written interview can be much more detailed, ideographic, and complex than the traditional self-report of well-being. Such an interview can include a subject's hopes and concerns, goals and interests, and pleasures and pains. It is harder for respondents to avoid unpleasant aspects of their lives when formulating their responses because information about these domains is directly requested. The expert rater then judges this complex material and converts it into a numeric response. Thus, the method avoids problems associated with traditional self-report measures such as differences between subjects in the strategies used to retrieve and integrate relevant information, and differences in numerically scaling their level of happiness.

A measure of daily moods over an extended period is another alternative that avoids other problems associated with traditional self-reports of SWB. The advantage to a daily score is that subjects' memories are likely to be much more accurate for a single day than for a prolonged period. Furthermore, although subjects might be loathe to admit that they are very unhappy, they would be much more likely to admit that they had a bad day. Therefore, the daily score is likely to give a good indication of subjects' average moods.

The reports of informants are likely to summarize emotional information expressed by subjects over time—information conveyed in both verbal and nonverbal channels. Because some nonverbal channels (e.g., vocal parameters) are less controllable than self-reports (DePaulo & Rosenthal, 1979), knowledgeable others may

gain an accurate idea of a subject's emotional life. The reports of informants have a strength which is complementary to self-report: Respondents are less likely to be able to hide their true feelings from knowledgeable others.

A memory measure of SWB is based on the idea that important emotional experiences will be encoded in subjects' memories. Indeed, a primary source of self-reports of well-being must be the recall of such remembered material. A number of studies have shown that happy persons recall more positive versus negative life events than unhappy persons (e.g., Diener, Sandvik, Pavot, & Gallagher, 1991; Pavot, Diener, Colvin, & Sandvik, 1991). Seidlitz and Diener (1993) demonstrated that happy individuals can recall more positive events and fewer negative events because they initially encode the events in this manner. Differences between happy and unhappy persons in the encoding of life events may be due to both differences in the incidence of positive versus negative objective events and to differences in the interpretation of events. Other research has indicated that mood, both at the time of encoding (Bower, Gilligan, & Monteiro, 1981) and retrieval (Clark & Teasdale, 1985; Forgas, Bower, & Krantz, 1984), has effects on the valence of information recalled (for a review, see Blaney, 1986). Thus, there are empirical and theoretical reasons to believe that counts of memories for positive and negative events will tap SWB. The idea that a simple count of such memories will be a valid index is based on the finding that happy persons frequently experience positive feelings and infrequently experience negative feelings (Diener, Sandvik, & Pavot, 1991), and therefore should have more positive events available for easy recall. Thus, memory for happy versus unhappy events is likely to tap SWB, and will be complementary to self-report in that it is less susceptible to response artifacts such as differences in using numbers to rate SWB or social desirability.

Finally, a forced-choice measure was specifically designed to control social desirability. In assessing the frequency with which various positive versus negative emotions are experienced, this measure requires subjects to choose between emotions and personality traits that have been balanced for the social desirability of the items. Scores on this measure are based on the positive and negative emotion terms chosen by subjects over trait adjectives having approximately equal social desirability. Thus, the measure complements traditional self-report measures in that it taps the SWB construct uncontaminated by social desirability.

Therefore, in this study a number of nontraditional measures of SWB which appear to have promise as alternative ways to assess subjects' private experiences were employed. Although some of these alternative measures rely on self-report to some degree, they all guard against at least some important threat to traditional self-report measures, and thus as a group represent an index of SWB which is relatively free from traditional self-report biases.

We compared the construct validity of various well-being measures by examining their correlations with measures of various constructs theoretically related to SWB such as extraversion, neuroticism, optimism, and symptoms of ill-being. Extraversion and neuroticism have been found to be related to well-being in a number of studies (e.g., Costa & McCrae, 1980; Emmons & Diener, 1985), but the exact reasons for these relations are not yet firmly established (Pavot, Diener, & Fujita,

1990). Emmons and Diener (1986) found that the extraversion-SWB relation resulted from a relation of the sociability component of extraversion with positive affect, and speculated that the reinforcement and satisfaction of needs provided by others lead to an increase in positive affect. Emmons and Diener (1986) also found that neuroticism was related to negative but not positive affect, and that impulsivity may have contributed in small degree to the neuroticism–negative affect relation.

Optimism has been theoretically and empirically related to SWB in previous studies through its association with successful coping with stress and physical ill-being. Scheier, Weintraub, and Carver (1986) found that optimists were more likely than pessimists to use problem-focused coping strategies, seek social support, and emphasize positive aspects of stressful situations. Carver and Gaines (1987) found that optimism moderated the tendency toward depression after a stressful life change, the birth of a child. In a study with coronary bypass surgery patients, Scheier, Matthews, and Owens (1989) found optimism to be associated with problem-focused coping, faster recovery, and higher quality of life 6 months after surgery. Thus, optimism was expected to be associated with SWB because of its association with successful coping strategies.

Symptoms of illness are predicted to correlate inversely with well-being because of causal paths in both directions. Evidence from a number of studies suggests that illness makes people less happy (Okun, Stock, Haring, & Witter, 1984); however, other studies suggest that the relation between self-reported health and SWB results primarily from unhappy people perceiving more symptoms of illness (Costa, McCrae, & Norris, 1981; Okun & George, 1984). Similarly, a high grade point average might correlate with well-being both because it is a valued resource, and also because depressed individuals may function less well academically.

For empirical and theoretical reasons, both family income and divorce or death of parents are expected to be associated with SWB. E. Diener, Sandvik, Seidlitz, and M. Diener (1992) have shown income to be associated with SWB in both a cross-country study and a longitudinal study within the United States. Their evidence suggested that income is a resource that not only satisfies basic needs but also contributes to SWB after basic needs have been met. Divorce and death within a family are among the most stressful life events and have been shown to be associated with lower SWB (Diener, 1984).

To test the discriminant validity of SWB instruments, we have also examined the correlations of the various measures with several constructs which are a priori theoretically independent of SWB: stimulus screening (Mehrabian, 1976), intelligence (Otis, 1939), and affect intensity (Larsen, 1984). Larsen's (1984) Affect Intensity Measure, for a variety of theoretical reasons, was predicted not to correlate substantially with measures of well-being. First, this measure assesses the intensity of *both* a person's pleasant and unpleasant emotions, and thus should not be correlated with happiness. Second, Diener, Sandvik, et al. (1991) have empirically demonstrated that happiness is more related to emotional frequency than to its intensity. Larsen and Diener (1987) review evidence showing that individual differences in emotional intensity are unrelated to individual differences in happiness. Finally, Diener, Colvin, Pavot, and Allman (1991) have shown that intense positive emotions

can entail costs. Thus, based on theoretical reasons as well as past empirical work, affect intensity was predicted to exhibit low, discriminant correlations with measures of SWB.

Stimulus screening is the individual difference characteristic of the automatic screening out of irrelevant stimuli, and of the habituation to distracting, irrelevant stimuli. The characteristic is closely related to arousability, such that screeners impose a structure on complex situational stimuli, resulting in lower arousal. Unlike extraversion, however, this concept does not emphasize sociability, which is associated with well-being. Thus, screeners have lower arousal and less distraction due to irrelevant stimuli than nonscreeners, but there is no theoretical relation between these characteristics and frequencies of positive and negative affect or cognitive judgments of overall life satisfaction, on which SWB is based.

The third discriminant variable, intelligence, consistently has been found to be unrelated to SWB in past studies (Diener, 1984). It has been conjectured that the rewards of being intelligent are offset by higher expectations and by an ability to understand problems and counterfactual alternatives more clearly. Thus, based on previous research using a variety of measures of mental ability and well-being, intelligence was hypothesized to be unrelated to SWB.

Method

Participants

Participants included 136 University of Illinois students enrolled in a course that involved extended participation in research in SWB during the fall of 1986. Before any statistical analyses were performed, the number of participants was reduced to 130,46 men and 84 women, due either to suspect or largely incomplete data by 6 respondents. The number of participants included in particular analyses varied due to a number of cases with partially missing data.

Happiness Measures

Standard self-report. Participants completed a number of conventional and non-conventional SWB measures during the semester. Conventional measures included the Fordyce Happiness Measure (Fordyce, 1977), the D-T Scale (Andrews and Withey, 1976), the Bradburn Affect Balance Scale (1969), and the Affectometer 2 (Kammann & Flett, 1983).

Alternative Measures

An alternative SWB measure included a written interview that participants were given to complete at home during the first week. The written interview contained

open-ended questions pertaining to the participants' happiness and satisfaction with life. The detailed five-page, 19-item questionnaire included diverse questions on topics such as subjects' typical moods, suicidal ideation, and the happiest and unhappiest times in their life. Examples of items include the following: "Describe your life right now; the good points and bad." "When, how intensely, and how often do you feel excited and enthusiastic about your life?" "Describe your average moods, your ups and downs, your pleasant and unpleasant feelings and when these occur." These questionnaires were later given global ratings by the senior investigator (Diener) and two research assistants for evidence of SWB on a 7-point scale on which 0 represented "very depressed in general" and 6 represented "elated most of the time." All ratings were made blind to the identity of subjects.

Subjects also completed daily affect reports on 42 occasions over a period of 6 weeks during the latter part of the semester, after the traditional self-report measures had been taken. Participants were asked to recall their daily mood each evening by rating the degree to which they experienced each of four positive emotions and four negative emotions on a 0–6 scale ranging from "not at all" to "extremely much." The ratings of the negative adjectives were reverse-scored and added to the ratings of the positive adjectives to yield a total daily affect score. The average score for each subject for the 42 days was then computed.

As a non-self-report measure of happiness, a minimum of seven informants rated each participant's happiness using the Fordyce scale. Informants included at least three family members and at least three friends of the subject. Comparisons of reports from family members with those of friends would provide information concerning the situational consistency of subjective well-being. These informant reports were confidential and sent directly to the experimenters.

Participants were given a 3-min, timed test to recall as many positive life events as they could that had occurred during their lifetimes, and a separate 3-min test to recall as many negative life events as they could. Participants also were given separate 2-min, timed tests to recall as many positive and negative life events as they could that had occurred during the past year. Subjects generally recall different types of events under the different instruction sets, with a tendency for major life events to be recalled under the lifetime set, and minor hassles and uplifts to be recalled under the past year set. Because the life and year memory tasks are included in the same testing period, subjects assume that nonredundant information is being requested and thus list different events. The 2- and 3-min tests were combined and an event memory index of SWB was derived by subtracting the total number of negative events recalled from the total number of positive events recalled.

To obtain a social desirability-free assessment of SWB, a measure was constructed based on Edwards's forced-choice approach. Participants were asked to make a choice between 54 pairs of affect versus personality trait adjectives that were previously matched for social desirability by raters (across the entire list, not item by item). Subjects were asked to choose the one alternative for each pair that was most self-descriptive. Examples of item pairs were "happy-kind" and "unhappy-dishonest." The number of times participants chose the mood adjective of the pair was summed to yield separate positive and negative affect scores. An

overall forced-choice affect balance well-being score was computed by subtracting the negative affect score from the total number of negative affect adjectives, and adding this reversed negative affect score to the positive affect score. This score represented subjects' affect balance, uninfluenced by need for approval or social desirability. When considered on an item-by-item basis, words were paired such that always answering in the socially desirable direction would yield a score of only one higher than would be obtained if a subject answered randomly. Because the range of the scale was 54 points and the standard deviation of scores was 10, social desirability could have only a trivial effect on scores. Thus the scores were likely to be highly saturated by affect information uncontaminated by social desirability.

Theoretical Correlates and Noncorrelates of Well-Being

Measures of a number of theoretical correlates of SWB were also collected during the semester to examine the construct validity of the various measures. These included the Brief Symptom Inventory (BSI; Derogatis, 1975), the Life Orientation Test (LOT; Scheier & Carver, 1985), and the extraversion and neuroticism scales of the NEO Personality Inventory (NEO-PI; Costa & McCrae, 1985). The positive affect subscale of the NEO-PI extraversion scale was omitted so that correlations with the SWB measures would not be artifactually inflated.

Other SWB correlates included self-reported grade point average, a trouble index (composed of the sum of unit weighting of below-median self-reported parental income, parental divorce, and death of father or mother), and a daily ill-feeling index (composed from the sum of two standardized items included on a questionnaire completed daily over a period of 42 days—the items inquired how many aspirin or other pain relievers had been taken and whether the participant had felt under the weather that day).

Several personality measures of constructs theoretically independent of SWB were also completed. These included the Affect Intensity Measure (AIM; Larsen, 1984), the Otis Quick Scoring Mental Abilities Test (Otis, 1939), and the Stimulus Screening Measure (Mehrabian, 1976).

Results

The means, standard deviations, and reliabilities of the various measures are presented in Table 1. The numbers of participants completing the measures varied because of partially missing data for several of the participants. The reliabilities of the well-being instruments were obtained using several methods. Although most reliabilities shown in Table 1 are alphas, in some cases traditional alpha coefficients were unavailable. The coefficient for the single-item Fordyce is based on the 2-week test-retest, whereas the reliability for the D-T Scale was provided by an estimate from Andrews and Withey (1976). The alpha for the written interview uses three raters as items, and the alpha for the informant reports uses seven informants as

Table 1 Means, standard deviations, and reliabilities of measures

Measures	<i>N</i>	Mean	<i>SD</i>	Reliability ^a
Standard self-report measures				
Fordyce	126	7.32	1.16	0.72 ^b
D-T Scale	128	2.76	1.20	0.70 ^c
Affectometer 2	128	111.60	23.27	0.96
Bradburn	128	6.48	1.81	0.53
Alternative measures				
Written interview	119	3.56	0.92	0.68
Daily affect	129	31.15	4.70	0.88
Informant reports	126	7.44	0.89	0.72
Event memory	126	3.98	6.44	0.69
Forced-choice	130	82.67	9.73	0.89
Theoretical correlates				
Trouble index	130	0.46	0.70	
Daily ill-feeling index	129	0.00	1.76	
Grade point average (5-point scale)	127	4.07	0.61	
Brief Symptom Inventory	120	11.00	3.62	
Optimism (LOT) ^d	126	19.46	5.27	
NEO extraversion	117	98.70	13.37	
NEO neuroticism	117	91.02	23.39	
Discriminant variables				
Stimulus Screening	128	174.27	67.71	
Otis	127	64.55	7.13	
Affect Intensity Measure	112	152.29	20.32	

^a Reliabilities are alpha coefficients unless otherwise noted. Alphas for the written interview and informant reports are based on agreement between raters or informants, instead of items.

^b Two-week test-retest reliability.

^c Provided as estimate for D-T Scale (Andrews and Withey, 1976).

^d Life Orientation Test.

items. The alpha for the event memory uses the two tests (past year and lifetime) as items. In addition to the alphas shown in the table, the daily affect measure's test-retest reliability (first 21 daily reports vs. second 21 reports) may be of interest—it was 0.94. All remaining reliability coefficients shown in the table are traditional alphas.

Intercorrelations of SWB Instruments

The intercorrelations of the various SWB measures are presented in Table 2. The conventional scales are listed first, followed by the alternative SWB measures. The intercorrelations show satisfactory convergent validity for all of the measures. Of the alternative measures, the written interview, daily affect reports, and informant reports correlated most highly with the conventional measures, most ranging between 0.50 and 0.70. The event memory measure correlated moderately with most of the measures, generally in the 0.30 to 0.50 range, its indirect methodology probably contributing to the lower correlations. The forced-choice measure also correlated in the 0.30 to 0.50 range with the other SWB measures.

Table 2 Intercorrelations of subjective well-being measures

	1	2	3	4	5	6	7	8	9
1. Fordyce	–								
2. D-T Scale	0.62	–							
3. Affectometer 2	0.57	0.72	–						
4. Bradburn	0.46	0.54	0.69	–					
5. Written interview	0.68	0.71	0.76	0.60	–				
6. Daily affect	0.54	0.66	0.70	0.56	0.59	–			
7. Informant reports	0.58	0.58	0.54	0.34	0.60	0.52	–		
8. Event memory	0.33	0.41	0.41	0.20	0.47	0.45	0.34	–	
9. Forced-choice	0.27	0.50	0.50	0.35	0.47	0.37	0.36	0.18	–

Note. $N = 110$ to 128 . For $r > 0.17$, $p < 0.05$. For $r > 0.22$, $p < 0.01$.

Principal Axis Factor Analyses

Table 3 shows the factor loadings of the SWB measures on the first factor obtained in an exploratory principal axis factor analysis of the self-report instruments and the five alternative measures. The purpose of the analysis was to explore the number of factors needed to adequately describe the data. The first factor extracted had an eigenvalue of 4.82 and accounted for 53.6% of the variance. It was the only factor with an eigenvalue above one, and the scree plot distinctly indicated a single factor. Both the self-report and non-self-report measures loaded highly on the first factor, indicating convergence on a single underlying construct. Although the various measures have acceptably high levels of reliability, an estimate of what the factor loadings would be if there were no measurement error was obtained by disattenuating the correlation matrix for unreliability using a standard formula (Ghiselli, Campbell, & Zedeck, 1981), and then factoring the disattenuated matrix. As seen in the right-hand column, the disattenuated loadings are quite high. The first factor underlying the disattenuated correlation matrix accounted for 61.1% of the variance.

Because the first factor analysis included a preponderance of self-report measures, we performed a second exploratory factor analysis including only the five alternative measures to determine if a single factor would emerge. The first factor in the principal axis factor analysis had an eigenvalue of 2.31 and explained 46.2%

Table 3 Factor loadings of subjective well-being measures

Measure	Factor loading	Disattenuated loading
Fordyce	0.74	0.78
D-T Scale	0.85	0.92
Affectometer 2	0.89	0.85
Bradburn	0.66	0.81
Written interview	0.85	0.92
Daily affect	0.77	0.79
Informant reports	0.68	0.74
Event memory	0.53	0.58
Forced-choice	0.52	0.57

Table 4 Factor loadings of alternative well-being measures

Measure	Factor loading	Disattenuated loading
Written interview	0.84	1.00
Daily affect	0.73	0.79
Informant reports	0.71	0.82
Event memory	0.55	0.64
Forced-choice	0.51	0.56

of the variance in the measures. It was the only factor with an eigenvalue above one, and again the scree plot of eigenvalues clearly suggested a one-factor solution. Table 4 shows the loadings of the five alternative measures on the first factor. As indicated in the table, the factor analysis shows strong convergence of the nonconventional measures. The factor loadings average 0.67, and the loadings based on the correlation matrix disattenuated for unreliability average 0.76. The single factor based on the correlation matrix of disattenuated measures accounted for 60.2% of the variance. Given the extremely different methodologies used in the alternative measures, this finding is very encouraging in terms of pointing to a unified underlying construct of SWB.

Neither factor analysis indicates that variance due simply to the self-report method is a major component of the SWB measures. Several of the highest-loading variables are self-report. It should be noted, however, that the one measure which is most clearly not a self-report assessment, informant reports, loads highly in both factor analyses, and that the self-report forced-choice method fares more poorly. If one compares the average correlation between the four conventional self-report measures with each other (0.60) to the average correlation of these measures with the informant reports (0.46), an estimate of 15% of the variance in self-report measures due to the self-report methodology would be obtained (0.60^2 minus 0.46^2).

Correlations of SWB Measures with Theoretical Correlates

Table 5 shows the correlations of the traditional self-report and alternative SWB measures with the theoretically related variables and with the principal axis factor underlying the five alternative measures. The conventional measures again are listed first, followed by the alternative measures. The conventional self-report measures compared favorably to the nonconventional measures in predicting several of the theoretical correlates, including the trouble index, optimism, extraversion, and neuroticism. Even the single-item measures performed at a level comparable to the more elaborate alternative measures. Nevertheless, it is encouraging that the nonconventional measures often correlated significantly with the theoretical correlates, indicating that the correlation of these variables is not simply due to self-report artifacts.

The correlations of the self-report measures with the SWB factor underlying the alternative measures were very respectable. Indeed, the average correlation of

Table 5 Correlations of subjective well-being measures with theoretical correlates

Measures	TI	DIFI	GPA	BSI	LOT	EXT	NEUR	FAC
Fordyce	-0.36	-0.01	0.16	-0.22	-0.59	0.40	-0.41	0.71
D-T Scale	-0.18	-0.14	0.14	-0.12	-0.59	0.30	-0.57	0.79
Affectometer 2	-0.11	-0.21	0.17	-0.17	-0.72	0.49	-0.73	0.82
Bradburn	-0.06	-0.22	0.04	-0.06	-0.50	0.36	-0.57	0.61
Written interview	-0.15	-0.21	0.20	-0.20	-0.67	0.42	-0.60	-
Daily affect	-0.21	-0.23	0.21	-0.17	-0.50	0.46	-0.55	-
Informant reports	-0.27	-0.07	0.29	-0.14	-0.51	0.38	-0.34	-
Event memory	-0.16	-0.33	0.25	-0.21	-0.36	0.22	-0.29	-
Forced-choice	-0.06	-0.25	0.02	-0.09	-0.38	0.13	-0.39	-

Note. $N = 99$ to 128 . For $r > 0.17$, $p < 0.05$. For $r > 0.22$, $p < 0.01$. TI = trouble index, composed of the sum of unit weightings of parental death, parental divorce, and below-median parental income (less than \$43,000); DIFI = daily ill-feeling index, composed of two standardized items: daily self-reports of feeling under the weather (yes or no) and number of aspirin taken; GPA = grade point average; BSI = Brief Symptom Inventory; LOT = Life Orientation Test; EXT = NEO extraversion, excluding positive affect subscale; NEUR = NEO neuroticism; FAC = principal axis factor underlying the five alternative subjective well-being measures.

the self-report measures with the non-self-report factor (0.73) slightly outstripped the average loading of the alternative measures (0.67) from which the factor was derived. These high correlations are especially encouraging because the common factor underlying the alternative method measures reflects the SWB construct untainted by many of the problems associated with self-report methodology.

Although it is encouraging that the SWB measures correlate with the theoretically related variables, their correlations with the LOT and NEO-PI neuroticism scales are as high as the intercorrelations among the well-being scales themselves. An analysis of the neuroticism items of the NEO-PI indicates that they essentially measure negative affect. Indeed, Watson and Clark (1984) suggest that neuroticism is the propensity to experience negative affect. In confirmatory factor-analytic work, Fujita, Diener, and Pavot (1993) find that neuroticism and negative affect are indistinguishable. Similarly, generalized expectancies about the future as measured by the LOT might be based on feelings of positive or negative affect. Thus, several personality constructs may have a strong affective underpinning and not be discriminable from the construct of SWB. Further work is needed in order to fully assess the discriminant validity of neuroticism and optimism from SWB—work with large samples and with multiple measures of each construct.

Comparisons Among Self-Report Instruments

The correlations of the four self-report instruments with the alternative SWB measures and with the theoretically related constructs were compared to determine which self-report instruments had greater convergent validity. Direct comparisons of the self-report instruments may be of interest to survey researchers and others who, because of practical constraints, are forced to choose between available conventional instruments. Four types of comparisons were made: (a) between the two single-item

Table 6 Z scores for differences in correlations for pairs of alternative subjective well-being measures

	D-T Scale vs. Fordyce	Affectometer vs. Bradburn	Affectometer vs. D-T	Affectometer vs. Fordyce
Fordyce	–	<i>ns</i>	<i>ns</i>	–
D-T Scale	–	3.40	–	<i>ns</i>
Affectometer 2	2.59	–	–	–
Bradburn	<i>ns</i>	–	3.02	3.60
Written interview	<i>ns</i>	2.96	<i>ns</i>	<i>ns</i>
Daily affect	2.03	2.76	<i>ns</i>	2.70
Informant reports	<i>ns</i>	3.17	<i>ns</i>	<i>ns</i>
Event memory	<i>ns</i>	3.12	<i>ns</i>	<i>ns</i>
Forced-choice	3.25	2.52	<i>ns</i>	3.13
Trouble index	–2.43	<i>ns</i>	<i>ns</i>	–3.28
Daily ill-feeling	<i>ns</i>	<i>ns</i>	<i>ns</i>	2.52
Grade point average	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>
Brief Symptom Inventory	<i>ns</i>	<i>ns</i>	<i>ns</i>	<i>ns</i>
Optimism (LOT) ^a	<i>ns</i>	3.97	2.66	2.66
NEO extraversion	<i>ns</i>	1.96	3.01	<i>ns</i>
NEO neuroticism	2.29	2.90	3.04	4.53

^a LOT = Life Orientation Test.

Note. Positive value indicates the first instrument of the pair had the higher correlation with the measure.

measures, (b) between the two multi-item measures, (c) between the multi-item Affectometer 2 and the Fordyce measure, and (d) between the Affectometer 2 and the single-item D-T Scale.

Table 6 shows the Z scores for the differences in correlations between the two instruments in each comparison with the various well-being measures and related constructs listed in the left column. For example, the first entry, 2.59, represents the Z score for the difference between (a) the correlation of the D-T Scale with the Affectometer 2, and (b) the correlation of the Fordyce measure with the Affectometer 2. The positive value of the Z score indicates that the first instrument of the pair listed in the heading, the D-T Scale, had the higher correlation with the Affectometer 2. In comparisons between the D-T Scale and the Fordyce measure, the D-T Scale also correlated more highly with two alternative well-being measures and with neuroticism, whereas the Fordyce measure better-predicted the trouble index. The Affectometer 2 was substantially superior to the Bradburn Affect Balance Scale in correlations with the other well-being instruments and with several related variables, not surprising given the relatively low reliability of the Bradburn scale. Shortcomings of the Bradburn scale have been discussed elsewhere (Diener, 1984). The D-T Scale's correlations were generally comparable to those of the Affectometer 2, but it had lower correlations with the Bradburn scale and several of the theoretically related variables. The Fordyce measure compared less favorably to the Affectometer 2; it had lower correlations with the Bradburn, daily affect, and forced-choice measures, as well as with several of the theoretical correlates.

Cross-Situational Consistency

To examine more systematically the long-term cross-situational consistency of SWB, we separately analyzed the informant reports completed by family members and those completed by friends. Family members observed the student participants in very different situations from those in which the informant friends encountered them. The participants, mostly juniors and seniors, had lived away from home for several years. Family members primarily interacted with the participants several years before in a home setting. In contrast, friends primarily saw the participants in a large university setting, living in an apartment with other students, at parties, etc. Thus, convergence of these two sets of informant reports would provide evidence of the long-term cross-situational consistency of SWB.

Table 7 shows the separate intercorrelations of the two sets of informant reports with self-report measures, the alternative well-being measures, and the theoretically related variables. As shown in Table 7, the two sets of informant reports correlated 0.44 with each other, and correlated highly and similarly with the alternative well-being measures. They also correlated similarly and at a moderate level with several variables theoretically related to well-being. The nonsignificant correlations of the informant reports with the daily ill-feeling index and the BSI probably resulted because the minor health problems assessed by these measures were frequently unobservable. Overall, the results support the cross-situational consistency of SWB.

If one disattenuates the correlation between friends and family for the reliability of these measures (0.85 and 0.74), one obtains an estimated correlation of 0.70. The coefficient of determination based on this number suggests that about half of the

Table 7 Informant reports separated by family versus friends

	Family reports	Friends' reports
Friends' reports	0.44	
Fordyce (Self)	0.48	0.50
D-T Scale	0.52	0.49
Affectometer 2	0.46	0.46
Bradburn	0.34	0.26
Written interview	0.54	0.52
Daily affect	0.50	0.42
Event memory	0.37	0.27
Forced-choice	0.34	0.28
Trouble index	-0.30	-0.18
Daily ill-feeling index	-0.13	-0.04
Grade point average	0.33	0.23
Brief Symptom Inventory	-0.11	-0.12
Optimism (LOT) ^a	-0.40	-0.47
NEO extraversion	0.31	0.33
NEO neuroticism	-0.33	-0.27

^a LOT = Life Orientation Test.

Note. $N = 113$ to 126 . For $r > 0.17$, $p < 0.05$. For $r > 0.23$, $p < 0.01$.

variance in SWB can be attributed to long-term and cross-situational factors such as personality and stable environmental conditions. The average intercorrelation of individual friends with each other was 0.49, of family members with each other was 0.65, and of family with friends, 0.32. This pattern suggests some stability across life periods, but greater levels of stability of SWB within a life period because there is greater agreement within groups than between groups. It could be, however, that the greater agreement within groups arises because there is more discussion of the individual within the rated groups than between them. Thus, the greater consensus within groups might be a function of discussion rather than behavioral consistency per se.

Discriminant Validity

Although the previous results show a unitary construct underlying the various SWB measures and consistency across situations, the discriminant validity of the well-being measures remains to be demonstrated. If the various scales and measures correlated substantially with constructs which were theoretically independent of SWB, the precise nature of the underlying construct would be in doubt. Table 8 presents the correlations of the various SWB measures with several constructs theoretically distinct from well-being: stimulus screening (Mehrabian, 1976), intelligence (Otis, 1939), and affect intensity (Larsen, 1984). Only 2 of the 27 correlations reached significance: the correlation of the daily affect with the Otis Quick Scoring Mental Abilities Test, and the correlation of the Affectometer 2 with the Stimulus Screening Measure. There were no particular SWB measures that showed substantially higher or lower correlations with the discriminant constructs than the other measures.

Table 8 Discriminant validity

	Stimulus Screening Measure	Intelligence ^a	Affect Intensity Measure
Fordyce	0.17	-0.06	0.11
D-T Scale	0.10	0.07	0.02
Affectometer 2	0.19*	0.02	0.01
Bradburn	0.17	0.05	0.03
Written interview	0.16	0.12	0.02
Daily affect	0.15	-0.22*	0.09
Informant reports	-0.03	-0.03	0.18
Event memory	0.15	0.02	0.04
Forced-choice	0.12	0.14	-0.15

^a Measured by the Otis Quick Scoring Mental Abilities Test.

* $p < 0.05$.

Note. $N = 104$ to 128.

Discussion

The results are encouraging for the use of self-reports in the assessment of well-being. For instance, it could be concluded that standard self-report well-being scales are adequate for many or most research purposes. The traditional self-report measures of SWB demonstrated high convergent validity by their agreement with alternative SWB measures and their relations with theoretically related constructs. In this respect they appear as or more valid and useful than most personality measures in use today. The results thus indicate a unitary core of experience for well-being, which self-reports reflect to a great extent. This core of experience has some temporal and situational stability. Thus, researchers using standard well-being scales can generally expect they are obtaining meaningful, interpretable information from these scales under ordinary conditions.

Situational influences, however, should still be considered. The correlations obtained between our diverse measures and standard self-report scales are high, but certainly not perfect. There is room in these relations for situational effects and momentary mood to have significant influences on the outcome of individual studies. Schwarz and his colleagues (Schwarz & Clore, 1983; Schwarz & Strack, 1991; Schwarz, Strack, Muller, & Chassein, 1988) have demonstrated that situational effects such as the weather or question ordering can influence self-reports of well-being. Momentary effects and simple error may also account for a part of this variance. Such effects may be minimized by increasing reliability through use of multi-item measures and by repeated measurements. Other problems associated with self-reports of SWB, such as repression or denial of negative emotions, and cultural differences in emotional experience and labeling, are more difficult to correct (Diener, 1994).

Several practical implications can be drawn from the present findings. First, the results imply several recommendations concerning the self-report instruments. The Affectometer 2 was shown to be a superior instrument to Bradburn's Affect Balance Scale. In other work (Diener, Emmons, Larsen, & Griffin, 1985; Pavot & Diener, 1993; Pavot et al., 1991), we have demonstrated that the Satisfaction with Life Scale is also a reliable and valid SWB instrument. The multi-item Affectometer 2 showed greater reliability and validity than the single-item measures, and thus its use is preferable when practical considerations permit. Nevertheless, the two single-item measures loaded highly on the factor underlying the various well-being measures and showed respectable convergent and discriminant validity. Thus, their use in survey research, when practical considerations preclude more elaborate measures, is defensible. Although both single-item measures demonstrated significant convergent and discriminant validity, comparisons between the two scales suggested the D-T Scale to be slightly preferable.

Second, although self-report well-being scales may be adequate for many purposes, they do not tell the whole story or necessarily contain all the information a given researcher might want or need. When possible, a broader base of measures is

desirable to investigate the experiential, communicative, behavioral, and physiological components of well-being, and their interconnections. For example, if groups differed in informant report versus self-report assessments, this would point to interesting hypotheses about the processes underlying the two types of assessment. Various assessment strategies are also useful for generating research hypotheses concerning the internal and external determinants of well-being. It may be, for example, that self-reports are adequate when comparing persons within a homogenous culture, but that additional measures are desirable when comparing persons' SWB across different cultures. In addition, alternative measurement methods such as informant reports and event memory seem desirable when time permits because they help rule out various artifactual explanations of results.

Finally, what do the results indicate about the stability of SWB? The answer seems to be that there are both long-term consistencies in average mood and a fair amount of fluctuation in short-term moods (Diener & Larsen, 1984). The present study clearly indicates that there are long-term consistencies in average mood, and that a variety of self-report and non-self-report methods can tap into this long-term average. Thus, SWB or happiness is a scientifically defensible area of study for personologists.

References

- Andrews, F. M., & Withey, S. B. (1976). *Social indicators of well-being: America's perception of life quality*. New York: Plenum Press.
- Beck, A. (1978). *Beck depression inventory*. Philadelphia: Center for Cognitive Therapy.
- Blaney, P. H. (1986). Affect and memory: A review. *Psychological Bulletin*, *99*, 229–246.
- Bower, G. H., Gilligan, S. G., & Monteiro, K. P. (1981). Selectivity of learning caused by affective states. *Journal of Experimental Psychology: General*, *110*, 451–473.
- Bradburn, N. M. (1969). *The structure of psychological well-being*. Chicago: Aldine.
- Cameron, P., Titus, D. G., Kostin, J., & Kostin, M. (1973). The life satisfaction of nonnormal persons. *Journal of Consulting and Clinical Psychology*, *41*, 207–214.
- Carroll, J. B. (1952). Ratings on traits measured by a factored personality inventory. *Journal of Abnormal Social Psychology*, *47*, 626–632.
- Carver, C. S., & Gaines, J. G. (1987). Optimism, pessimism, and postpartum depression. *Cognitive Therapy and Research*, *11*, 449–462.
- Clark, D. M., & Teasdale, J. D. (1985). Constraints on the effects of mood on memory. *Journal of Personality and Social Psychology*, *48*, 1595–1608.
- Costa, P. T., Jr., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: Happy and unhappy people. *Journal of Personality and Social Psychology*, *38*, 668–678.
- Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., Jr., McCrae, R. R., & Norris, A. H. (1981). Personal adjustment to aging: Longitudinal prediction from neuroticism and extraversion. *Journal of Gerontology*, *36*, 78–85.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.
- DePaulo, B. M., & Rosenthal, R. (1979). Ambivalence, discrepancy, and deception in nonverbal communication. In R. Rosenthal (Ed.), *Skill in nonverbal communication* (pp. 204–248). Cambridge, MA: Oelgeschlager, Gunn, & Hain.

- Derogatis, L. R. (1975). *Brief symptom inventory*. Baltimore: Clinical Psychometric Research.
- Diener, E. (1984). Subjective well-being. *Psychological Bulletin*, *95*, 542–575.
- Diener, E. (1994). Assessing subjective well-being: Progress and opportunities. *Social Indicators Research*, *31*, 103–157. (2005 reprinted in *Citation classics from Social Indicators Research*)
- Diener, E., Colvin, C. R., Pavot, W., & Allman, A. (1991). The psychic costs of intense positive affect. *Journal of Personality and Social Psychology*, *61*, 492–503.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction with Life Scale. *Journal of Personality Assessment*, *49*, 71–75.
- Diener, E., & Larsen, R. J. (1984). Temporal stability and cross-situational consistency of affective, behavioral, and cognitive responses. *Journal of Personality and Social Psychology*, *47*, 871–883.
- Diener, E., Sandvik, E., & Pavot, W. (1991). Happiness is the frequency, not the intensity, of positive versus negative affect. In F. Strack, M. Argyle, & N. Schwarz, (Eds.), *Subjective well-being: An interdisciplinary perspective* (pp. 119–139). London: Pergamon.
- Diener, E., Sandvik, E., Pavot, W., & Gallagher, D. (1991). Response artifacts in the measurement of subjective well-being. *Social Indicators Research*, *24*, 35–56.
- Diener, E., Sandvik, E., Seidlitz, L., & Diener, M. (1992). The relationship between income and subjective well-being: Relative or absolute? *Social Indicators Research*, *28*, 253–281.
- Dulany, D. E. (1968). Awareness, rules, and propositional control: A confrontation with S-R behaviorism. In T. Dixon & D. Horton (Eds.), *Verbal behavior and general behavior theory* (pp. 340–387). Englewood Cliffs, NJ: Prentice-Hall.
- Emmons, R. A., & Diener, E. (1985). Personality correlates of subjective well-being. *Personality and Social Psychology Bulletin*, *11*, 89–97.
- Emmons, R. A., & Diener, E. (1986). Influence of impulsivity and sociability on subjective well-being. *Journal of Personality and Social Psychology*, *50*, 1211–1215.
- Ericsson, K. A., & Simon, H. A. (1980). Verbal reports as data. *Psychological Review*, *87*, 215–251.
- Fordyce, M. W. (1977). *The happiness measures: A sixty-second index of emotional well-being and mental health*. Unpublished manuscript, Edison Community College, Ft. Myers, FL.
- Fordyce, M. W. (1988). A review of research on the Happiness Measures: A sixty-second index of happiness and mental health. *Social Indicators Research*, *20*, 355–381.
- Forgas, J. P., Bower, G. H., & Krantz, S. E. (1984). The influence of mood on perception of social interactions. *Journal of Experimental Social Psychology*, *20*, 497–513.
- Fujita, F., Diener, E., & Pavot, W. (1993). *An investigation of the relations between extraversion, neuroticism, positive affect, and negative affect*. Manuscript submitted for publication.
- Ghiselli, E. E., Campbell, J. P., & Zedeck, S. (1981). *Measurement theory for the behavioral sciences*. San Francisco: Freeman.
- Goldings, H. J. (1954). On the avowal and projection of happiness. *Journal of Personality*, *23*, 30–47.
- Gough, H. (1957). *California personality inventory*. Palo Alto, CA: Consulting Psychologists Press.
- Hartmann, G. W. (1934). Personality traits associated with variations in happiness. *Journal of Abnormal and Social Psychology*, *29*, 202–212.
- Hathaway, S. R., & McKinley, J. C. (1951). *Minnesota multiphasic personality inventory*. New York: Psychological Corporation.
- Hogan, R., & Nicholson, R. A. (1988). The meaning of personality test scores. *American Psychologist*, *43*, 621–626.
- Jackson, D. (1967). *Personality research form*. Port Huron, MI: Research Psychologists Press.
- Kammann, R., & Flett, R. (1983). *Sourcebook for measuring well-being with Affectometer 2*. Dunedin, New Zealand: Why Not? Foundation.
- Krug, S., & Laughlin, J. (1970). *IPAT depression checklist*. Champaign, IL: Institute for Personality and Ability Testing.
- Larsen, R. J. (1984). Theory and measurement of affect intensity as an individual difference characteristic. *Dissertation Abstracts International*, *85*, 2297B. (University Microfilms No. 84-22112).

- Larsen, R. J., & Diener, E. (1987). Affect intensity as an individual difference variable: A review. *Journal of Research in Personality, 21*, 1–39.
- McNeil, J. K., Stones, M. J., & Kozma, A. (1986). Subjective well-being in later life: Issues concerning measurement and prediction. *Social Indicators Research, 18*, 35–70.
- Mehrabian, A. (1976). *Manual for the questionnaire measure of Stimulus Screening and Arousability*. University of California, Los Angeles.
- Moskowitz, D. S. (1986). Comparison of self-reports, reports by knowledgeable informants, and behavioral observation data. *Journal of Personality, 54*, 294–317.
- Nisbett, R. E., & Ross, L. (1980). *Human inference: Strategies and shortcomings of social judgment*. Englewood Cliffs, NJ: Prentice-Hall.
- Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*, 231–259.
- Okun, M. A., & George, L. K. (1984). Physician- and self-ratings of health, neuroticism and subjective well-being among men and women. *Personality and Individual Differences, 5*, 533–539.
- Okun, M. A., Stock, W. A., Haring, M. J., & Writter, R. A. (1984). Health and subjective well-being: A meta-analysis. *Journal of Aging and Human Development, 19*, 111–132.
- Otis, A. S. (1939). *Otis Quick Scoring Mental Abilities Test*. New York: Harcourt, Brace & World.
- Pavot, W., & Diener, E. (1993). Review of the satisfaction with life scale. *Psychological Assessment, 5*, 164–172.
- Pavot, W., Diener, E., Colvin, C. R., & Sandvik, E. (1991). Further validation of the Satisfaction with Life Scale: Evidence for the cross-method convergence of well-being measures. *Journal of Personality Assessment, 57*, 149–161.
- Pavot, W., Diener, E., & Fujita, F. (1990). Extraversion and well-being. *Personality and Individual Differences, 11*, 1299–1306.
- Phillips, D. L. (1967). Social participation and happiness. *American Journal of Sociology, 72*, 479–488.
- Scheier, M. F., & Carver, C. S. (1985). Optimism, coping, and health: Assessment and implications of generalized outcome expectancies. *Health Psychology, 4*, 219–247.
- Scheier, M. F., Matthews, K. A., & Owens, J. F. (1989). Dispositional optimism and recovery from coronary artery bypass surgery: The beneficial effects on physical and psychological well-being. *Journal of Personality and Social Psychology, 57*, 1024–1040.
- Scheier, M. F., Weintraub, J. K., & Carver, C. S. (1986). Coping with stress: Divergent strategies of optimists and pessimists. *Journal of Personality and Social Psychology, 51*, 1257–1264.
- Schwarz, N., & Clore, G. L. (1983). Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states. *Journal of Personality and Social Psychology, 45*, 513–523.
- Schwarz, N., & Strack, F. (1991). Evaluating one's life: A judgment model of subjective well-being. In F. Strack, M. Argyle, & N. Schwarz (Eds.), *Subjective well-being: An interdisciplinary perspective* (pp. 27–48). London: Pergamon.
- Schwarz, N., Strack, F., Muller, G., & Chassein, B. (1988). The range of response alternatives may determine the meaning of the question: Further evidence on informative functions of response alternatives. *Social Cognition, 6*, 107–117.
- Seidlitz, L., & Diener, E. (1993). Memory for positive versus negative life events: Theories for the differences between happy and unhappy persons. *Journal of Personality and Social Psychology, 64*, 654–664.
- Underwood, B., & Froming, W. (1980). The mood survey: A personality measure of happy and sad moods. *Journal of Personality Assessment, 44*, 404–414.
- Watson, D., & Clark, L. A. (1984). Negative affectivity: The disposition to experience aversive emotional states. *Psychological Bulletin, 96*, 465–490.