RANDY J. LARSEN, ED DIENER, AND ROBERT A. EMMONS

AN EVALUATION OF SUBJECTIVE WELL-BEING MEASURES

(Received 10 November, 1984)

ABSTRACT. A number of subjective well-being scales were compared and evaluated. The Satisfaction With Life Scale emerged as a good measure of general life satisfaction and the Affective Intensity Measure appeared to adequately assess the characteristic level of emotional intensity. Most other scales seemed to reflect both life satisfaction and duration of positive versus negative affect. Of the single item measures, those created by Fordyce were the strongest, whereas for the multi-item scales, several performed at adequate levels. The widely used Bradburn scales showed several undesirable psychometric properties and alternative scales are suggested.

Research on subjective well-being has increased dramatically in recent years. It is not uncommon to see articles in popular magazines and newspapers reporting on the "state of the nation" in terms of how happy we are, how satisfied we are with our lives, and how we stand on other indicators of subjective well-being. The topic of subjective well-being is also gaining popularity as an area of psychological research. In *Psychological Abstracts* the headings "happiness" and "satisfaction" did not become regular features until 1973 and the number of entries under these titles has risen rapidly.

As this research area has grown, so too has there been a proliferation of measures of subjective well-being. The development of such measures has helped provide clearer conceptualizations of the construct of subjective well-being. Also, the existence of readily available and easy to administer measures has led to a variety of empirical findings on the antecedents, consequences, and correlates of subjective well-being. Although investigators have produced many interesting substantive findings, basic research on the adequacy of the various scales is still insufficient. A number of evaluations of existing scales are available (Conte and Salamon, 1982; George and Bearon, 1980; Larson, 1978; Lohmann, 1977), but these reviews have focused on geriatric subjective well-being scales. Thus, there is a need for an evaluation and comparison of the commonly used measures that are appropriate for general adult samples.

In the present study several popular measures of subjective well-being and happiness were evaluated and compared in diverse samples of subjects. Only

Social Indicators Research **17** (1985) 1–17. 0303–8300/85.10

scales that are appropriate for general adult populations were included. Thus, the geriatric scales which are inappropriate for younger adults were not included in this paper. In addition to temporal reliability and internal consistency, we also evaluated the measures in terms of four general types of validity: (1) Criterion validity – the degree to which the scales predicted daily moods over time as well as peer reports of happiness and satisfaction, (2) Convergent validity – the degree to which the scales correlated with each other, (3) Content validity – the degree to which the scales reflect various components of well-being (e.g. duration of positive affect, intensity of affect, or life satisfaction), and (4) Construct validity – the degree to which the scales correlated with several personality dimensions as would be theoretically predicted (e.g. neuroticism and self-esteem).

METHOD

Subjects

Five diverse subject samples were used to evaluate the well-being measures:

Sample 1 consisted of 163 introductory psychology students at the University of Illinois at Urbana-Champaign. These subjects completed several subjective well-being scales as well as a battery of personality and temperament questionnaires. Two months later 76 of these students were rescheduled and took the subjective well-being measures again.

Sample 2 consisted of 63 students at the College of the Virgin Islands enrolled in undergraduate psychology classes. This mostly black Caribbean sample completed a battery of subjective well-being measures.

Sample 3 was 130 University of Illinois students who were enrolled in a semester long research project on daily mood (see Diener and Emmons, 1984, for a complete description of the methodology). These subjects completed a detailed mood report every day for a period ranging from 6 to 10 weeks. A subsample of 62 subjects also provided "peer" reports from at least five sources (e.g. parents and roommates).

Sample 4 consisted of 176 introductory psychology students at the University of Illinois. These subjects took a battery of subjective well-being measures.

Sample 5 included 34 elderly persons who were recruited through various church and volunteer organizations in the Urbana-Champaign community.

The age range was from 48 to 85, with the median age being 65. These subjects completed a battery of well-being scales twice (with a 30 day test-retest interval).

Well-being measures

Several of the subjective well-being measures are single item scales. We evaluated five of these:

(1) Andrews and Withey (1976) developed a question which gives seven explicit response alternatives ranging from "Delighted" at one end to "Terrible" at the other, and hence the name, "The D-T scale". The questions asks, "How do you feel about how happy you are?"

(2) Cantril (1965) created the "self-anchoring ladder". This measure shows a picture of a nine step ladder anchored at one end with the phrase "Best possible life for you", and at the other with, "Worst possible life for you". The intervening responses (rungs on the ladder) are not labelled. The subject responds to the question, "Where on the ladder do you stand at the present time?" Thus, the wording suggests an ipsatized index which is specific to the subject.

(3) Fordyce's (1978) questions is, "In general, how happy or unhappy do you usually feel?" There are 11 response choices that are each graphically anchored with a series of mood adjectives. For example, the highest choice is anchored with "Feeling extremely happy, ecstatic, joyous, and fantastic". This measure will be referred to as Fordyce 1.

(4) Fordyce (1978) developed another question that asks, "What percent of the time do you feel happy, what percent of the time do you feel unhappy, and what percent of the time do you feel neutral?" The subject is required to estimate these percentages such that they add up to 100 percent. We will refer to the "percent happy" score as Fordyce 2.

(5) Gurin, Veroff and Feld's (1960) popular item, is "Taking all things together, how would you say things are these days?" There are three response choices: "Very happy", "Pretty happy", and "Not too happy".

There are also several multi-item measures of subjective well-being that we compared:

(1) Bradburn and Caplovitz (1965) developed a ten item true/false inventory that yields a positive affect score (PAS) and a separate negative affect score (NAS). The difference between these two scales is often used to create a

third score called "affect balance" (ABS). It is noteworthy that Bradburn found positive and negative affect tended to vary rather independently across persons and we have replicated this result with other methodologies (Diener and Emmons, in press).

(2) Campbell, Converse, and Rodger's (1976) eight item scale which consists of semantic differential type items. The subject rates his or her life in general along such dimensions as interesting-boring and worthwhile-useless.

(3) Tellegen (1979) has developed the Differential Personality Questionnaire (DPQ) which contains a 24 item true/false Well-Being subscale.

(4) Underwood and Froming (1980) created the Mood Survey (MS), a 15 item inventory that yields two scores: one that reflects average level of positive affect or hedonic tone (the "Level" subscale), and another that reflects emotional reactivity and lability (the "Reactivity" subscale).

(4) Larsen (1983) has developed the 40 item Affect Intensity Measure (AIM) which assesses the typical intensity or strength with which individuals characteristically experience their emotions. The items are balanced in terms of positive and negative emotions.

(6) Diener, Emmons, Larsen and Griffin (in press) have developed a 5 item Satisfaction With Life Scale (SWLS) that is focused explicitly and exclusively on life satisfaction as a cognitive judgmental evaluation of one's life as a whole.

RESULTS

The studies were conducted over a several year period and some scales were not used in the earliest studies. This accounts for the missing values in several of the tables. In Sample 1 ($n = 163$) we examined sex differences. Using *t*-tests we found no significant differences between males and females on any of the well-being measures, a finding which is consistent with other recent findings (Diener, 1984).

Reliability

Table I shows the test-retest reliabilities for two different samples. In general the single item scales performed poorly, with reliabilities in the 0.30 to 0.50 range. The exceptions were the single item Fordyce scales whose reliabilities were stronger. The reliabilities of the Gurin and Cantril scales were quite low.

TABLE I

Test-retest and alpha coefficient reliabilities of subjective well-being scales

Scale	Elderly sample (*T-R*) $N = 34$	Student sample (*T-R*) $N = 76$	Student sample (Alpha) $N = 176$
AIM	*	*	0.87
Andrews D-T	0.66	0.45	*
Bradburn			
PAS	0.41	0.50	0.66
NAS	0.52	0.63	0.48
ABS	0.57	0.64	*
Campbell	0.35	0.51	0.91
Cantril	0.41	0.32	*
DPQ	*	*	0.89
Fordyce 1	0.59	0.59	*
Fordyce 2	0.81	0.60	*
Gurin	0.37	0.29	*
Mood survey			
Level	*	*	0.88
Reactivity	*	*	0.74
SWLS	*	0.83	0.93

Note. *T-R* = test-retest reliability. The interval was one month for the elderly sample and two months for the student samples. Asterisks represent missing values or the statistic does not apply.

The multi-item scales tended to show modest reliabilities. However, the SWLS, which is a cognitive life satisfaction measure, showed high temporal reliability. It is possible that the life satisfaction judgment tapped by this scale is more stable over time than is affect, which the other scales reflect to a greater degree.

The fact that the test-retest reliabilities are only of modest magnitude suggests that the components of subjective well-being possess both trait and state aspects (Kirkaldy, 1984). The high reliability for life satisfaction found with the SWLS is consistent with the finding of Diener and Larsen (1984) that life satisfaction was the most stable and cross-situationally consistent of any response they measured. Since the measure and daily report methodology used by Diener and Larsen were quite different from the SWLS, this provides convergent evidence that life satisfaction is relatively stable and consistent. The Gurin and Cantril measures show disappointing reliabilities, perhaps partly because they contain only one item. In addition, the Gurin scale has only 3 response options, and on the Cantril scale the response options are unlabelled. Because the Cantril scale is phrased in an ipsatized way, it may be

expected that it would be less reliable because it may be more sensitive to change within a respondent's life.

Internal consistency

Coefficient alpha was computed for each of the multi-item scales based on Sample 4 data. The values are also presented in Table I. The Bradburn scales had relatively low alphas, whereas the internal consistency estimates were higher for the other scales, ranging from 0.74 to 0.93. Once again the SWLS showed a very high value.

Convergent validity

In Tables II and III are the inter-scale correlation matrices for four of our samples. The scales tended to correlate with each other at moderate to high levels. The convergent validities tended to be higher for the multi-item scales (DPQ, Level, and Campbell). However, the single item Fordyce scales also showed higher convergent validities and the Bradburn scales tended to yield somewhat lower values. A factor analysis was conducted on the correlations

TABLE II

Interscale correlations, sample 4

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. AIM													
2. Andrews	14												
3. PAS	30	51											
4. NAS	09	−38	00										
5. ABS	08	62	66	−76									
6. Campbell	22	64	54	−38	64								
7. Cantril	13	52	39	−39	55	61							
8. DPQ	17	59	56	−33	62	77	56						
9. Fordyce 1	25	58	45	−28	52	65	50	66					
10. Fordyce 2	22	56	37	−21	38	62	47	71	70				
11. Gurin	14	42	31	−31	42	63	53	60	55	54			
12. MS-level	28	64	51	−34	59	77	49	77	74	70	56		
13. MS-react.	21	−36	−12	41	−40	−41	−33	−35	−36	−33	−35	−44	
14. SWLS	09	68	50	−37	62	75	62	66	58	58	59	67	−33

Note. Decimals are omitted.
$N = 176$
MS = Mood Survey.

TABLE III

Inter-scale correlations, three separate samples

Scale	Sample	1	2	3	4	5	6	7	8	9	10
1 Andrews D-T	A										
	B										
	C										
2. PAS	A	46									
	B	49									
	C	49									
3. NAS	A	−28	−11								
	B	−05	−08								
	C	−33	−22								
4. ABS	A	49	72	−77							
	B	29	42	−87							
	C	52	−87	−78							
5. Campbell	A	45	30	−30	46						
	B	37	34	−25	41						
	C	70	62	−36	59						
6. Cantril	A	46	50	−39	59	54					
	B	36	39	−27	44	61					
	C	45	50	−19	44	69					
7. Fordyce 1	A	53	53	−33	54	49	58				
	B	52	38	−35	41	56	60				
	C	*	*	*	*	*	*				
8. Fordyce 2	A	54	53	−33	49	49	51	68			
	B	50	38	−35	63	54	60	64			
	C	*	*	*	*	*	*	*			
9. Gurin	A	71	42	−35	51	54	48	54	53		
	B	71	46	−35	55	55	52	62	52		
	C	73	48	−41	59	71	45	*	*		
10. SWLS	A	58	52	−31	55	65	64	60	63	49	
	B	*	*	*	*	*	*	*	*	*	
	C	*	*	*	*	*	*	*	*	*	

Note. A = Student sample, N = 163. B = Elderly sample, N = 34. C = Virgin island sample, N = 63. Decimals are omitted.

for our largest sample on the subjective well-being measures, as well as Larsen's AIM and the Emotionality and Activity subscales of the EASI-III (Buss and Plomin, 1975). Principle axis factor analysis with iterations was used, followed by varimax rotation. The criterion for retaining factors for rotation was an eigenvalue greater than unity. A clear two-factor solution emerged which accounted for 60% of the common variance. The first factor consisted of all the standard subjective well-being scales and the second factor was composed of Affect Intensity, Reactivity, Emotionality, and Activity. The varimax rotated factor loadings are presented in Table IV. As can be seen, multi-item

TABLE IV

Factor loadings of well-being and temperament scales

	Factor 1	Factor II
Andrews D-T	0.79	
PAS	0.59	
NAS	–0.40	
Campbell	0.89	
Cantril	0.67	
DPQ	0.86	
Gurin	0.68	
Fordyce 1	0.79	
Fordyce 2	0.77	
MS-level	0.87	
SWLS	0.84	
MS-reactivity	–0.48	0.48
Activity		0.31
AIM		0.74
Emotionality		0.57

Note: $N = 176$. Only loadings greater than 0.30 are shown.

scales (with the exception of Bradburn's) show slightly higher loadings than the single item scales on Factor 1.

Since most of the well-being scales loaded on a single major factor, we created a global measure of subjective well-being consisting of the average of the measures loading on Factor 1. In effect we considered each well-being scale to be a separate item on a global subjective well-being measure. The coefficient alpha for this global scale was 0.66. This finding, along with the factor analysis, strongly suggests that the individual well-being scales share substantial common variance. The correlations between each scale and the sum of all the others (in effect, the corrected item-total correlations) are given in Table V. The Bradburn positive minus negative "affect balance" score is not included in Tables V or VI because it is a linear combination of the other two subscales.

Criterion Validity

Bradburn and Caplovitz (1969) have defined subjective well-being as the experience of more positive affect than negative affect over a given period of time. Taking this as one operational definition of subjective well-being, we measured positive and negative affect on a daily basis for periods ranging

TABLE V

Analysis of scales as items on a global well-being measure, sample 4, students

	Item-total correlation
Andrews D-T	0.66
PAS	0.48
NAS	−0.81
Campbell	0.75
Cantril	0.59
DPQ	0.81
Fordyce 1	0.76
Fordyce 2	0.72
Gurin	0.64
MS-level	0.81
SWLS	0.76

Note. N = 176.

from 40 to 70 days for Sample 3. For each subject we computed the percentage of days in which he or she reported feeling more positive than negative affect. We also computed the mean positive affect on days when positive exceeded negative affect. We refer to this as the intensity of positive affect because it represents for each subject how positive, on the average, their good days were. We also computed the converse; negative affect intensity refers to the amount of negative affect they reported feeling on the average for those days when negative affect predominated over positive affect. Unlike simple averages of daily affect, these daily intensity scores are not influenced by how frequently a person feels an emotion, but only by how intense the emotion is on the average when it is experienced (Diener, Larsen, Levine and Emmons, in press).

Duration of positive affect and positive and negative affect intensity served as criteria for evaluating the well-being scales. Table 6 shows the correlations between each of the well-being scales and these three criteria for Sample 3. It is clear from these correlations that the well-being scales tap the duration of positive affect much more than they reflect its intensity. In other words, these scales give an indication of the percentage of time a person experiences more positive than negative affect, but little indication as to the intensity or strength with which they experience their emotions. The correlations between duration and the intensity scores were not significantly different from zero, suggesting that how frequently one is happy or unhappy is not related to how intensely one

TABLE VI

Daily affect criterion correlations, sample 3

	Duration	Positive intensity	Negative intensity
AIM	−0.21	0.58	0.55
Andrews D-T	0.44	0.18	0.05
PAS	0.37	0.11	−0.07
NAS	−0.24	−0.01	0.02
ABS	0.39	0.05	−0.09
Campbell	0.37	0.28	0.08
Cantril	0.51	0.04	−0.03
DPQ	0.52	0.14	−0.02
Fordyce 1	0.57	0.20	−0.08
Fordyce 2	0.49	0.13	0.00
Gurin	0.41	0.11	−0.07
SWLS	0.37	0.16	0.05

Note. $N = 130$, except AIM = 62.

feels happiness or unhappiness. These findings complement the factor analytic finding discussed above showing that intensity and frequency of positive affect are distinct components of subjective well-being. Note that the AIM loaded on a separate factor from the other well-being scales and correlates much more highly with daily affect intensity. However, the AIM does not correlate substantially with the duration of positive affect. The criterion validity correlations between the well-being scales and duration of positive affect averaged over several weeks are only minimally different for the single versus the multiple item scales. However, once again Bradburn scales show lower validity coefficients than most other measures.

A subsample of 62 of the daily report subjects each provided between five and ten peer reports. "Peers" (including friends, relatives and roommates) rated the subject for global positive affect, for the percentage of time they thought the subject was happy, and for how intensely they thought the subject experiences his or her positive emotions. The correlations between these reports and each of the self-report well-being scales are presented in Table VII. The peer reports of global positive affect and the peer reports of percent of time happy correlated moderately with the well-being scales, in general ranging from 0.40 to 0.60. There were only minimal differences between the single and multiple item scales in terms of their correlations with

TABLE VII

Peer rating correlations, sample 3, peer report measures

	Global positive affect	Percent of time happy	Intensity of happiness
Andrews D-T	0.56	0.36	0.16
PAS	0.44	0.45	0.17
NAS	−0.10	−0.35	−0.02
ABS	0.52	0.46	0.14
Campbell	0.41	0.34	0.18
Cantril	0.53	0.23	0.28
DPQ	0.34	0.34	0.23
Fordyce 1	0.51	0.43	0.42
Fordyce 2	0.49	0.54	0.29
Gurin	0.45	0.47	0.13
SWLS	0.25	0.30	0.20

Note. $N = 130$.

peer reports. There was a tendency, however, for the SWLS to show lower correlations with peer reports of affect. Peer reports of the intensity of positive emotions correlated only weakly with the standard well-being scales. This suggests once again that the intensity component of subjective well-being is not being assessed by these scales.

Construct Validity

Subjective well-being should relate in a predictable way to personality and temperament variables. In Table VIII we present the correlations between the well-being measures and several personality and temperament scales for Sample 1. The well-being scales covaried strongly with the Rosenberg Self-Esteem Scale, and moderately with temperament measures of sociability and activity. There also were strong positive correlations between the well-being scales and ratings of being satisfied with various life domains such as being satisfied with friendships, financial situations, love life, housing, grades, etcetera. There were moderate to strong negative correlations between the well-being scales and neuroticism, self-reported symptoms, and the temperament of emotionality.

TABLE VIII

Correlations between SWB scales and personality measures

	S-E	Sym	Neu	Dom	Emo	Act	Soc
Andrews D-T	0.44	−0.34	−0.40	0.40	−0.31	0.16	0.23
PAS	0.36	−0.31	−0.31	0.42	−0.18	0.19	0.27
NAS	−0.39	0.45	0.45	−0.24	0.28	−0.06	0.00
ABS	0.50	−0.51	−0.52	0.44	−0.32	0.17	0.17
Campbell	0.54	−0.36	−0.36	0.46	−0.38	0.18	0.38
Cantril	0.49	−0.44	−0.35	0.51	−0.30	0.07	0.22
Fordyce 1	0.48	−0.43	−0.41	0.50	−0.30	0.17	0.34
Fordyce 2	0.44	−0.27	−0.39	0.38	−0.34	0.21	0.39
Gurin	0.47	−0.40	−0.42	0.41	−0.30	0.16	0.25
SWLS	0.65	−0.43	−0.51	0.57	−0.37	0.14	0.25

Note. S-E = Self-esteem. Sym = Symptom checklist. Neu = Neuroticism. Dom = Domain satisfaction. Emo = Emotionality. Act = Activity. Soc = Sociability.
$N = 163$.
Correlations of 0.14 are significant at the 0.05 level and 0.18 at the 0.01 level.

Social Desirability

We examined social desirability response bias by correlating each of the well-being scales with the Marlowe-Crowne scale (Crowne and Marlowe, 1964). Only two scales showed significant ($p < 0.05$, one tailed) correlations with social desirability: The Gurin scale correlated 0.24 and the SWLS correlated 0.22 with the Marlowe-Crowne scale. These correlations, however, represent only meager portions of variance due to response bias. Considering the nature of the construct of subjective well-being, it is surprising that the correlations with social desirability were not higher because people with a positive outlook tend to respond in a positive way on the Marlowe-Crowne scale. One valuable direction for future research would be to correlate the scales with other social desirability measures such as the SD scale of Edwards (1957) which are not highly correlated with the Marlowe-Crowne scale.

DISCUSSION

Are the present results affected by the rather narrow samples used in these studies? It would appear that the results are generalizable for several reasons. First, each of the samples gave very similar results, suggesting that the correlations we report are generalizable to other populations. Second, although the

samples we used were rather narrow, several were quite different from each other and yet yielded similar results. Finally, the homogeneity of the samples, if anything, would lower the correlations compared to a more heterogeneous national sample, and therefore the correlations we present can be taken as conservative estimates. It is an unfortunate fact that the extensive amount of data collected on each subject in these studies would be difficult or impossible to obtain from a national sample.

The factor analytic and convergence results presented in this article are dependent on the particular questionnaires we used. In other words, the sample of tests used will influence which load most highly on a factor analysis. In these studies we attempted to include the major scales in current use and to include scales that varied in length and response format. Thus, we feel confident that high factor loadings are not due to methodological components, but are due to high saturation with the content that these scales measure. Thus, a scale such as the DPQ that shows high convergence with the other scales must reflect to a high degree the content that is measured by most subjective well-being scales.

The test-retest reliabilities of the scales suggest one answer to the question about the extent to which the measurement of subjective well-being simply reflects the respondent's mood at the time. Schwarz and Clore (1983) have demonstrated that current mood does indeed influence subjective well-being responses. However, the temporal reliabilities found here imply that these measures are not simply a reflection of mood at the time of responding. There was a substantial portion of common variance between measures taken one or two months apart, especially for life satisfaction. Hence, the scales must also reflect longer-term trends in mood. Diener and Larsen (1984) have also found temporal stability as well as cross-situational consistency in the various components of subjective well-being.

In choosing which scales to include in one's research there are several considerations, one of which is the length of the scale. There are several single item measures of subjective well-being. However, these scales are generally less reliable over time than multi-item scales, are probably more susceptible to asquiescene response bias, are more likely to be affected by the particular wording of the item, may not be entirely suitable for parametric analysis since responses tend to be highly skewed, and do not provide an assessment of the separate components of subjective well-being. On the other hand, these scales are quick to administer and several of them have been in

use for a decade or more. As a consequence, much is known about their properties, norms on national samples have been compiled, and results over time and across studies can be compared. Thus, the choice of which measures to use is a cost/benefit one which must be made in terms of the goals of a particular study.

What other general conclusions can be drawn in evaluating the scales? In general, most scales performed at adequate levels, a finding that is consonant with data on the geriatric scales (George and Bearon, 1980). The SWLS shows high temporal reliability, a high alpha for internal consistency, and good convergence with other measures. Findings suggest that life satisfaction may be more stable than affective aspects of well-being, although the relationship between life satisfaction and positive affect is not yet well understood (Diener, 1984).

The existence of multi-item scales raises the question as to the multidimensional nature of subjective well-being. The present study, along with the work of Andrews and Withey (1976), suggests that there are several components of subjective well-being. These might include, for example, duration of positive affect, the intensity of affect, and life satisfaction. Life satisfaction shows an empirical relationship to duration of positive affect. Nevertheless, life satisfaction is at least conceptually distinct from duration of positive affect and should play a role in the theory and measurement of subjective well-being. It should be noted that the SWLS correlated substantially with other scales but correlated at a lower level with the strictly affective measures used in the daily study. The lower correlations of the SWLS with affect stands in contrast to the high correlations found for the SWLS in other areas (e.g. with domain satisfactions). This suggests that the SWLS measures life satisfaction as a cognitive evaluation of one's life rather than duration of positive affect per se. The pattern of correlations also indicates that the majority of subjective well-being scales are influenced by both positive affect and life satisfaction, since the other scales correlate substantially with both the SWLS and with daily affect.

Certain measures, notably the single item Gurin and Cantril scales seem to have weaker psychometric properties, for example low temporal reliability. The widely used Bradburn scales appear to have only modest alphas and test-retest reliabilities, and tend to show lower convergence with other scales. The Bradburn scales were also less highly correlated with daily affect. The "affect balance" (Positive minus Negative) score did perform somewhat

better than either the positive or negative subscales, supporting Bradburn's contention that happiness is a combination of these two independent factors. The weakness of these scales is disappointing since they are the only ones we assessed which measure positive and negative affect separately. Fortunately there are two newer scales which are designed to measure positive and negative affect and which show initially promising results (Kammann and Flett, 1983a and b; Kozma and Stones, 1980). Although Kammann and Flett present very promising data on their Affectometer 2, this scale was not known to us until after we completed the present studies. Of the measures we assessed, the SWLS appears to be the most promising instrument for measuring life satisfaction per se. The Fordyce measures (especially Fordyce 1), although short, performed adequately in most ways and therefore deserves wider use. Several multi-item scales performed well on most criteria – the DPQ, Campbell, and the Level scales.

In terms of the construct validity correlations, it is noteworthy that the personality dimensions of self-esteem and neuroticism show substantial relationships with subjective well-being. Although the causal direction of this influence is unknown, these findings do indicate the importance of the study of personality variables in relation to happiness.

Intensity of affect emerged as a component of subjective well-being that is unrelated to both duration of positive affect and to life satisfaction and refers to individual differences in how strongly people characteristically experience their emotions. The Affect Intensity Measure (AIM) appears to be the measure of choice for assessing the intensity dimension. Although Underwood and Froming's scale measures a related construct, reactivity, it appears that their scale is also influenced by the duration of negative affect. Although affective intensity is not related to the duration of happiness, it does influence the qualitative experience of happiness for an individual. When duration and intensity measures are used in combination, they can provide a finely differentiated view of individual differences in affective well-being. For example, someone who is:

(1) High in duration of positive affect and high in intensity will experience subjective well-being as an exuberant, enthusiastic, and actively cheerful sort of happiness.

(2) High in duration of positive affect and low in intensity will experience subjective well-being as a calm, untroubled, and dispassionate contentment.

(3) Low in duration of positive affect and low in intensity will experience

a lack of subjective well-being as a chronic low level unhappiness, a mild but persistent melancholy.

(4) Low in duration of positive affect but high in intensity will experience a lack of subjective well-being as a more acute sort of depression accompanied by a variety of strongly felt negative emotions such as embarrassment, guilt, grief, and shame.

Subjective well-being is a rich and multi-faceted construct. Researchers should be aware that there exist several measures that tap global well-being as well as several measures that tap the more specific components of subjective well-being. Ultimately the choice of which measures to use will depend on the format of the study and on the questions the experimenter wants to answer. Since research on subjective well-being has increased dramatically, questions about the validity of measures are bound to arise. In particular, questions about problems related to the self-report nature of these measures are important. Diener (1984) discusses these issues and the results presented in this paper support his argument that these scales do contain substantial portions of valid variance. This conclusion was reaffirmed by Weinstein (1982) who found that a self-report measure of happiness correlated substantially with observational measures of smiling and laughing. Nonetheless, our results suggest that some measures are stronger than others. In addition, results from the present study help clarify the relationship between the various components of subjective well-being such as positive affect, life satisfaction, and affect intensity.

BIBLIOGRAPHY

Andrews, F. M., and S. Withey: 1965, Social Indicators of Well-Being (Plenum, New York).

Bradburn, N. M., and D. Caplovitz: 1965, Reports on Happiness (Aldine, Chicago).

Buss, A. H., and R. Plomin: 1975, A Temperament Theory of Personality Development (Wiley, New York).

Campbell, A., P. E. Converse, and W. L. Rodgers: 1976, The Quality of American Life (Russell Sage, New York).

Cantril, H.: 1965, The Patterns of Human Concern (Rutgers University Press, New Brunswick, N.J.).

Conte, V. C. and M. J. Salamon: 1982, 'An objective approach to the measurement and use of life satisfaction with older persons', Measurement and Evaluation in Guidance 15, pp. 194–200.

Crowne, D., and D. Marlowe: 1964, The Approval Motive (Wiley, New York).

Derogatis, L. R., R. S. Lipman, J. Rickels, E. H. Uhlenhuth, and L. Covi: 1974, The Hopkins symptom checklist: A measure of primary symptom dimensions, in P. Pichot (Ed.), Psychological Measurements in Psychopharmacology: Modern Problems in Pharmopsychiatry (S. Krager, Basel).

Diener, E.: 1984, 'Subjective well-being', Psychological Bulletin 95, pp. 542–575.

Diener, E. and R. A. Emmons:1984, 'The independence of positive and negative affect', Journal of Personality and Social Psychology 47, pp. 1105–1117.
Diener, E. and R. J. Larsen: 1984, Temporal stability and cross-situational consistency of affective, behavioural, and cognitive responses', Journal of Personality and Social Psychology 47, pp. 871–883.
Diener, E., R. J. Larsen, S. Levine and R. A. Emmons: 'Frequency and intensity: The underlying dimensions of positive and negative affect', Journal of Personality and Social Psychology, in press.
Diener, E., R. A. Emmons, R. J. Larsen, and S. Griffin: 'The satisfaction with life scale: A measure of life satisfaction', Journal of Personality Assessment, in press.
Eysenck, H. J., and S. B. G. Eysenck: 1968, Manual for the Eysenck Personality Inventory (Educational and Industrial Testing Service, San Diego).
Fordyce, M. W.: 1978, Prospectus: The Self-Description Inventory (Unpublished manuscript, Edison Community College, Fort Myers, Florida).
George, L. K., and L. B. Bearon: 1980, Quality of Life in Older Persons: Meaning and Measurement (Human Sciences Press, New York).
Gurin, G., J. Veroff, and S. Feld: 1960, American's View of Their Mental Health (Basic, New York).
Kammann, R. and R. Flett: 1963, 'Affectometer 2: A scale to measure current level of general happiness', Australian Journal of Psychology 35, pp. 257–263. (a)
Kammann, R., and R. Flett: 1983, Sourcebook for Measuring Well-Being with Affectometer 2 (Why Not? Foundation, Dunedin, New Zealand).
Kirkaldy, B. D.: 1984: 'The interrelationship between state and trait variables', Personality and Individual Differences 5, pp. 141–149.
Kozma, A. and M. J. Stones: 1980, 'The measurement of happiness: Development of the Memorial University of Newfoundland Scale of Happiness (MUNSCH)', Journal of Gerontology 35, pp. 906–912.
Larsen, R. J.: 1983, Manual for the Affect Intensity Measure (Unpublished manuscript, Purdue University).
Larsen, R. J.: 1984, Theory and Measurement of Affect Intensity as an Individual Difference Characteristic (Unpublished doctoral dissertation, University of Illinois at Urbana-Champaign).
Larson, R.: 1978, 'Thirty years of research on the subjective well-being of older Americans', Journal of Gerontology 33, pp. 109–125.
Lohmann, N. L. P.: 1977, 'Comparison on life satisfaction, morale and adjustment scales on an elderly population', (Doctoral dissertation, Brandeis University, 1977), Dissertation Abstracts International 38, p. 418B. (University Microfilms No. 77–15, 272).
Rosenberg, M.: 1965, Society and the Adolescent Self-Image (Princeton University Press, Princeton, N.J.).
Schwarz, N. and G. L. Clore: 1983, 'Mood, misattribution, and judgments of well-being: Informative and directive functions of affective states', Journal of Personality and Social Psychology 45, pp. 513–523.
Tellegen, A.: 1979, Differential Personality Questionnaire (Unpublished materials, University of Minnesota).
Underwood, B. and W. J. Froming: 1980, 'The mood survey: A personality measure of happy and sad moods', Journal of Personality Assessment 44, pp. 404–414.
Weinstein, L.: 1982, 'Positive contrast as due to happiness', Bulletin of the Psychonomic Society 19, pp. 97–98.

Dept. of Psychology, Purdue University,
West Lafayette, IN 47907,
U.S.A.