ROBERT A. CUMMINS

18. ON THE TRAIL OF THE GOLD STANDARD FOR SUBJECTIVE WELL-BEING

(Accepted 7 February 1995)

ABSTRACT. The absence of a 'gold standard' for subjective well-being has severely hampered the interpretation of data from empirical studies. This paper demonstrates a remarkable consistency among the results of 16 studies that have investigated 'life satisfaction' among large samples drawn from the general population. It is concluded that a population standard for 'life satisfaction' can be expressed as 75.0 ± 2.5 percent of the measurement scale maximum score.

INTRODUCTION

The most common statement in the quality of life literature is a lament at the lack of a 'gold standard', or some agreed statistic which could form the basis of comparison between empirical studies. In the absence of such a reference, studies are constrained by the need to rely either on internal comparisons or comparisons with others that have used an identical measurement procedure. Unfortunately, studies rarely use comparable measurement procedures and this lack of an empirical frame of reference has severely limited the interpretation of data relating to subjective well-being (SWB).

Despite these difficulties, some general characteristics of such data have been identified. Most notable is that they are not normally distributed but negatively skewed. This has been found almost irrespective of the measuring instrument, population sample, or nationality, as data presented later in this paper will detail.

In order to explain this phenomenon various hypotheses have been proposed. In 1954 Goldings found that college students rated themselves significantly above the scale mid-point on happiness. From these data he proposed a 'behavioral norm' for happiness which was hypothesised to lie within the band of 6.0 to 7.5 on a 9-point scale. The explanation offered for this distribution was that,

Alex C. Michalos (ed.), Citation Classics from Social Indicators Research, 537–558.

since happiness is positively valued in our culture, people select a happiness rating for themselves that is socially acceptable.

Subsequent research, however, has extended the form of this data distribution to include other personal attributes including perceived energy (Gill, 1984) and a wide variety of life satisfaction measures. Clearly, therefore, explanations for this phenomenon need to be broadly based.

Boucher and Osgood (1969) provided a different and more general perspective. They noted that people have a tendency to prefer the use of positive rather than negative concepts when evaluating words, and coined the 'Polyanna Hypothesis' to explain people's preference for the positive end of life satisfaction scales. Headey and Wearing (1988, 1992), on the other hand, have suggested the phenomenon is based on a need to maintain self-esteem through downward social comparisons and describe the negative skew in self-perceptions as providing a 'Sense of Relative Superiority'. So, does the positive distribution represent some actual tendency to 'look on the bright side of life' instead of a response to social expectations as suggested by Goldings?

Evidence in support of an internal, adjustable mechanism is presented by Muthny *et al* (1990) who compared life satisfaction between three different cancer groups and healthy controls. They found that, in terms of general life satisfaction, the group with the worst prognosis had the highest representation (33%) in the top category of 'very satisfied'; almost three times the percentage (13%) of healthy people who used this category. The authors comment, "This clearly points to the issue of comparison with internal standards and ideals; a patient who has faced death might more easily be satisfied with [objectively] bad conditions than a healthy man with high expectations towards quality of life" (p. 156).

This idea of adjustable internal standards is derivative from Helson's (1964) Adaptation Level Theory and has been used by a number of researchers, most notably Brickman *et al.* (1978), to explain the observation that sharp falls or rises in life satisfaction occasioned by a major life event are usually followed by a gradual return to original levels of subjective well-being. But what causes the adaptation level to be set within the positive side of the well-being continuum?

A variety of mechanisms have been suggested which could underpin and maintain a positive perception of self. Taylor *et al.* (1983) have proposed such devices as making social comparisons with less fortunate others, a selective focus on attributes that make the self appear advantaged, creating hypothetical worse-worlds, construing benefit from adverse events, and manufacturing unrealistic normative standards. To this list Headey and Wearing (1992) have added such processes as differential weightings of sub-roles in personal assessments of overall role performance and the use of restricted reference groups.

There is clearly no shortage of psychological devices to nominate for the role of negative-skew maintenance. Indeed, it seems likely that many act in concert to produce the end result. If this is true, then it is likely that the phenomenon represents a true reflection of self-perception, rather than a response to social acquiescence. The multiple nature of supporting devices may also reflect the psychological importance of maintaining positive self-regard.

The conclusions that may be drawn are threefold. Firstly that the existence of the negative skew in subjective quality of life data is ubiquitous. Secondly that a variety of psychological mechanisms have been identified as likely candidates to explain the production and maintenance of such high personal self-regard. Thirdly that the consistency of this phenomenon across widely differing studies would be evidence for the operation of a psychological set-point for feelings of personal well-being.

If these conclusions are valid, then one outcome might be a high level of consistency, from one population to another, in their absolute level of perceived life satisfaction. Should this prove to be the case, a single statistic could provide a reference point, or a 'gold-standard', for all empirical studies into life quality. Such an investigation is the purpose of this paper.

METHOD

In order to provide a large and relatively homogenous data base for this research, the 'normative' studies have been restricted to population samples from Western countries. Data from other countries are presented in comparison to these 'normative' statistics later in the paper.

Normative Studies

More than 1 000 articles and books pertaining to the broad literature on life quality were scanned for normative data that met the following criteria:

1. The scale used must measure perceived 'satisfaction' with either some statement of global life assessment, or with a number of life domains from which a mean score can be derived. Scales measuring happiness have been excluded.
2. Likert scales must be symmetric, ranging from strongly positive to strongly negative, around a neutral mean point.
3. The number of points on the scale must be known and the mean value stated.
4. Each sample must conform to the following specifications.
   (a) It must be drawn from countries with comparable cultures and socio-economic status. The actual samples were drawn from Australia, Canada, England, Norway, Sweden, and U.S.A.
   (b) It must not be selected on medical, psychiatric, or socio-economic criteria.
   (c) It must not be restricted to rural communities, but may be restricted to urban populations.
   (d) It must not include children or adolescents less than 17 years of age and must include adults less than 65 years of age.
   (e) The sample size must exceed 200.

Despite meeting these criteria, a few studies were excluded for the following reasons:

Bortner and Hultsch (1970): The instructions to respondents combined the concepts of 'happiness' and 'satisfaction'.

Hall (1973): Reports the same data as Abrams (1973) with slightly differing domain names and excluding the item 'being a housewife'.

Judge and Watanabe (1993): Combined responses to bipolar adjectives with a measure of global life satisfaction.

White (1992): Combined measures of domain satisfaction with a measure of general happiness.

Comparative Studies

The same collection of articles and books used to identify the normative studies, was also searched for data on population sub-samples. The criteria for their inclusion were the same as (1) and (2) of the normative studies. In addition, a minimum of three studies were required that presented data relevant to a common population sub-group.

RESULTS

Normative Studies

A total of 17 data sets were located that fulfilled all of the normative criteria. The samples and scale details are presented in Table I.

On the right-hand column of this table all scale scores have been converted to the common statistic "Percentage of Scale Maximum" (%SM), calculated through the formula (score -1) $\times$ 100/(number of scale points -1). There are two variations on the application of this formula. In situations where the scale starts from zero, instead of from one, the formula becomes % SM = score $\times$ 100/(number of scale points -1). If the scale starts from a number greater than one, due to the aggregation of scores from multiple items, then the aggregate score is initially divided by the number of items to restore a scale range that commences with one.

Observations on the Normative Studies

Non-probability samples

Three studies are included that restricted their samples to population sub-groups. In each case, the nature of the sample seems unlikely to have influenced the mean score significantly away from population norms.

TABLE I

Studies used for the normative comparisons

Study	Country	Sample	Scale (see below)	N	Mean score	% scale max
Abrams (1973)	England	Quota	A1	593	5.59[1]	76.5
Andrews & Withey (1976)	USA	Probability	A2	4,920	5.42	73.7
Andrews (1991)	USA	Probability				
		– 1972	A2	1,297	5.50	75.0
		– 1988	A2	616	5.80[2]	80.0
Atkinson (1982)	Canada	Probability	A3	2,129	8.69[3]	76.9
Brief *et al.* (1993)	USA	Probability	A4	443	7.19	79.9
Campbell *et al.* (1976)	USA	Probability	A5	2,134	5.5	75.0
Harris & Assoc. (1975)	USA	Probability	A6	1,457	13.35	74.2
Headey & Wearing (1994)	Australia	Probability	A7	619	6.97[2]	74.6
Hoyert & Seltzer (1992)	USA	Female	A8	5,900	4.0	75.0
Gove *et al.* (1983)	USA	Probability	A9	2,174	2.24	74.7
Mastekaasa (1992)	Norway	Never married[4]	A10	6,214	5.6	76.1
Mookherjee (1992)	USA	Probability	A11	2,529	17.97	78.4
Moum (1994)	Norway	Probability	A12	71,896	5.37	72.8
Oppong *et al.* (1988)	Canada	Probability	A13	421	3.87[5]	71.7
Palmore & Luikart (1972)	USA	Probability	A4	502	7.0	70.0
Roy Morgan Research (1993)	Australia	Probability	A14	1,217	6.67	70.9

[1] Average of two measures.

[2] Average of two responses to the question asked about 20 minutes apart.

[3] Average of 1977 and 1979 data.

[4] The sample was actually designated 'unmarried' but their youth (20–39y) would suggest that the vast majorty would never have been married.

[5] Average of 1984 and 1985 data

The scale descriptions are as follows:

A1 – (a) "All things considered, how satisified or dissatisfied are you overall with [domain]. Which number comes closest to how satisfied or dissatisfied you are?". The score was a composite from 12 life domains, each rated on a 7-point scale from 'completely satisfied' to 'completely dissatisfied'. (b) Towards the end of the interview, each respondent was asked to take into account all of the aspects of life that had been discussed and to use the scale to indicate satisfaction or dissatisfication with life as a whole.

A2 – "... we want to find out how you feel about various parts of your life, and life in this country as you see it. Please tell me the feelings you have now – taking into account what has happened in the last year and what you expect in the near future" (p. 19). "How do you feel about your life as a whole?", using a 7-point Delighted-Terrible scale.

A3 – "All things considered, how satisfied or dissatisfied are you with your life as a whole? Which number comes the closest to how you feel?, from 1 = 'completely dissatisfied' to 11 = 'completely satisfied'.

A4 – Global life satisfaction was measured using the Cantril Ladder (Cantril, 1965), a single-item, 10 point scale ranging from 0 ('worst possible life') to 9 ('best possible life').

A5 – "We have talked about various parts of your life, now I want to ask you about your life as a whole. How satisfied are you with your life as a whole these days?" (p. 33), using a 7-point Likert scale from 'completely dissatisfied' to 'completely satisfied'.

A6 – Life satisfaction Index (Neugarten *et al.*, 1961). The scale consists of 18 statements with which respondents agree or disagree. Scores can range from 0 to 18.

A7 – "How satisfied are you with your life as a whole ", scored on a 9-point Delighted-Terrible scale.

A8 – "On the whole, I am satisfied with myself" scored from "1-strongly disagree to 5-strongly agree" (p. 77).

A9 – "How satisfying do you find your life, very satisfying, pretty satisfying, not too satisfying, or not at all satisfying?" (p. 63).

A10 – "Thinking about how you feel about your life these days, are you generally satisfied, or are you generally dissatisfied?" The seven response categories were labelled from "extremely satisfied" to "extremely dissatisfied" (Mastekaasa, 1994).

11 – Composite seven item scale comprising: How happy the respondent was in general, and how satisfied he/she felt with their residence, non-work activities, family life, friendship, health and physical condition, and financial situation. Scored on a 3-point scale 'Not satisfied at all, somewhat satisfied, well satisfied'.

A12 – "When you think about the way your life is going at present, would you say that you are by and large satisfied with life or are you mostly dissatisfied" (p. 8), scored on a 7-point scale from 'extremely dissatified' to 'extremely satisfied'.

A13 – "How satisfied are you with your ____?" (p. 618), from 1-'very dissatisfied' to 5 (or 7) 'very satisfied'. The aggregate score was derived from seven domains as: house, family life, health and physical condition, amount of time for leisure and hobbies, friendships, standard of living, job satisfaction.

A14 – "How pleased are you with your life as a whole", using a 9-point Delighted-Terrible scale.

Hoyert and Seltzer (1992) selected an exclusively female sample. However, probability samples have tended to show a gender difference in subjective well-being that is less than 2.5%, and which favor either males (Mastekaasa, 1992) or females (Branholm and Degerman, 1992; Shmotkin, 1990). One Australian study has reported identical gender values for subjective well-being (Roy Morgan Research, 1993). Table II will present data confirming the absence of a gender difference in life satisfaction.

Two other studies have included large samples that were selected from people who have never married (Hughes and Gove, 1981; Mastekaasa, 1992). These studies have been included on the basis that the SWB scores of such persons tend to fall slightly below people who are married and slightly above people who were once married but now divorced or widowed (Hughes and Gove, 1981; Ying, 1992). Thus it might reasonably be expected that the scores from these studies would differ little from the adult population as a whole.

Scales

Perhaps the most striking feature of the studies listed in Table I is the diversity of their scales. Among the 16 studies, 14 different measures of life satisfaction have been employed.

Apart from Cantril's Ladder, used by two studies, and Neugarten's Life Satisfaction Index, used by one, most others have employed a variation on the Andrews and Withey (1976) global question "How do you feel about your life as a whole?" scored on a 7-point Delightful-Terrible scale. However, the extent of the variations on this theme have been very considerable. Scales have ranged from a minimum of three response points (Mookherjee, 1992) to a maximum of 11 (Atkinson, 1982). Some authors have reported a mean response from two identical questions asked some 20 minutes or so apart (Andrews, 1991; Headey and Wearing, 1994) while most others ask the question only once.

A number of authors reported using quite elaborate 'framing' procedure for the question (e.g. Andrews and Withey, 1976; Moum, 1994). Framing procedures are used to define the context within which the person is expected to respond, and again these differ between studies to the point that the respondents' points of reference might be considered to have been very different. For example,

TABLE II

Sub-population comparisons

Sub-population and study	Country	Sample	Scale (see below)	N	Mean score	% Scale max.
Females						
Branholm & Degerman (1992)	Sweden	Married	B1	47	4.78	75.6[1]
			B2	47	5.05	81.0
Mastekaasa (1992)	Norway	Unmarried 20–39y	A10	2,089	5.50	75.0
Mookherjee (1992)	USA	Probability	A11	?[4]	17.97	78.4
Roy Morgan Research (1993)	Australia	Probability	A14	621	6.67	70.9
MEAN ± SD						75.7 ± 3.05
Aged >65y						
Adams (1969)	USA	Probability	A6	508	12.50	62.5
Andrews (1991)	USA	Probability	A2	25	5.8	80.0
Bowling (1990)	England	Low SES >85y	A6	662	13.30	66.5[1]
			B3	662	4.92	65.3
Bowling *et al.* (1991)	England	Low SES, 65–85y				
		Male	A6	186	12.86	64.3[3]
		Female	A6	279	12.38	61.9
		Low SES, >85y				
		Male	A6	166	11.68	58.4
		Female	A6	496	11.19	56.0
		Middle SES, 65–85y				
		Male	A6	95	13.80	69.0
		Female	A6	193	13.10	65.5
Harris & Assoc (1975)	USA	Probability	A6	2,797	12.20	67.8
Mookherjee (1992)	USA	Probability	A11	?[4]	18.01	78.6
MEAN ± SD						69.55 ± 7.15
Physical disability						
Krause (1992)	USA	Spinal injury >2y	B4	286	3.74	68.5
Schulz & Decker (1985)	USA	Spinal injury >20y	A6	100	11.76	59.8
Stensman (1985)	Sweden	Wheelchair users	B5	36	8.00	80.0
MEAN ± SD						69.4 ± 8.3
Chronic medical						
Baker & Intagliata (1982)	USA	Various disorders	B9	118	5.30	71.6
Ferrans & Powers (1992)	USA	Haemodialysis	B7	349	20.65	68.8

TABLE II Continued

Sub-population and study	Country	Sample	Scale (see below)	N	Mean score	% Scale max.
Huxley & Warner (1992)	USA	Low-income patients	B8	15	5.07	67.8
Kober *et al.* (1990)	USA	Liver disease	B6	12	4.00	50.0
	Germany	Liver disease	B6	9	4.50	58.3
MEAN ± SD						63.3 ± 8.01
Post-organ transplant						
Hicks *et al.* (1992)	USA	Liver >2yr	B7	17	21.60	72.0
Kober *et al.* (1990)	USA	Liver >2yr	B6	12	5.60	75.9
	Germany	Liver >2yr	B6	9	6.10	85.0
Koch & Muthny (1990)	Germany	Kidney	B10	761	3.71	67.8
Lough *et al.* (1985)	USA	Heart >6mo	B11	75	4.58	71.6
MEAN ± SD						74.5 ± 5.86
Intellectual disability						
Bramston (1994)	Australia	Mild/Moderate	B13	59	13.70	71.4
Cummins *et al.* (1994)	Australia	Mild/Moderate	B13	39	27.90	74.6
Schalock *et al.* (1989)	USA	Independent living	B12	25	2.54	76.8
MEAN ± SD						74.3 ± 2.19
Low income						
Andrews (1991)	USA	Probability	A2	67	5.40	73.3
Harris & Associates (1975)	USA	Probability	A6	?[5]	24.98	66.6
Mookherjee (1992)	USA	Probability	A11	?[4]	16.86	70.4
MEAN ± SD						70.1 ± 2.76
High income						
Andrews (1991)	USA	Probability	A2	71	6.00	83.3
Harris & Associates (1975)	USA	Probability	A6	?[5]	30.81	82.8
Mookherjee (1992)	USA	Probability	A11	?[4]	18.96	85.4
MEAN ± SD						83.8 ± 1.15
Whites (USA)						
Andrews (1991)	USA	Probability	A2	491	5.8	80.0
Harris & Associates (1975)	USA	Probability	A6	?[5]	29.80	80.0
Mookherjee (1992)	USA	Probability	A11	?[4]	18.10	79.3

TABLE II Continued

Sub-population and study	Country	Sample	Scale (see below)	N	Mean score	% Scale max.
MEAN ± SD						79.9 ± –
Blacks (USA)						
Andrews (1991)	USA	Probability	A2	53	5.6	76.7
Harris & Associates (1975)	USA	Probability	A6	?[5]	29.09	69.7
Mookherjee (1992)	USA	Probability	All	?[4]	17.14	72.4
MEAN ± SD						72.9 ± 2.91

[1] Where two different scales have been used on the same sample, a mean score for the scales in combination has been used in the calculation of the sub-population score.
[2] The sample contained 206 males and 241 females. The break-down for age-groupings is not stated.
[3] In calculating the sub-group mean for aged >65y, all six individual values from this study have been combined to yield a single score of 62.5% SM.
[4] The total probability sample for this study was 2,529 but the sub-population N was not stated.
[5] The total probability sample for this study was 1,457 but the sub-population N was not stated.

Scales A1 to A13 have been described below Table I. The other scale descriptions are as follows:
B1 – A 'Life Satisfaction Questionnaire' containing 8 items each scored on a 6-point scale from 'very dissatisfying' to 'very satisfying'.
B2 – Cantril Ladder based on 11 rungs.
B3 – Faces scale using 7 points on 5 items.
B4 – Life Situation Questionnaire containing eleven items scored on a 5-point satisifaction scale.
B5 – Single question on 'satisfaction with overall life quality' on an 11-point scale.
B6 – European Organization for Research and Treatment of Cancer Quality of Life Questionnaire; satisfaction with overall quality of life on a 7-point scale (Aaronson *et al.*, 1987).
B7 – Quality of Life Index, which used a formula of Importance × Satisfaction on an overall quality of life, 5-point scale, that ranged from 'very dissatisfied' to 'very satisfied'.
B8 – Quality of Life Profile (Oliver, 1988) containing 9 items on 7-point scales.
B9 Quality of Life Scale (Flanagan, 1978) containing 15 items on 7-point faces scales. "Which face comes closest to expressing how you feel about . . . ?"
B10 – An 8-item scale with each rated from 'very dissatisfied' to 'very satisfied' on 5-points.
B11 – "How satisfied are you with your current quality of life?" on a 6-point scale from 'Not at all satisfied' to 'very satisfied'.
B12 – Quality of Life Scale (Keith *et al.*, 1986) containing 28 items on a 3-point scale.
B13 – Comprehensive Quality of Life Scale – Intellectual Disability (Cummins, 1993) contains 7 domains scored using Importance × Satisfaction on a 5-point scale.

Andrews (1991) sets a complex frame which incorporates satisfaction with the person's own life, life for the national community, in the past, present and future (see A2, Table I). Others are more simple, concentrating on satisfaction with the person's life in the present (see A12, Table I) or even restricted to the person themselves (see A8, Table I). One (Roy Morgan Research, 1993) has used the context of how 'pleased' the person feels about their life. Still others are more general, stating frames such as 'all thing considered' (Atkinson, 1982).

Finally, there are differences in the dimension of 'satisfaction' to which the person responds. Most commonly people respond to a scale defined by 'completely dissatisfied' to 'completely satisfied'. However, the extended form of this scale has also been employed, using the 'Delighted-Terrible' anchors (e.g. Andrews and Withey, 1976), as has also the truncated form which ranges from 'very satisfying' to 'not at all satisfying' (Hughes and Gove, 1981). Others are as diverse as Cantril's Ladder which ranges from the 'worst possible life' to the 'best possible life', and the Life Satisfaction Index which contains statements to which the respondent agrees or disagrees.

All in all, there is remarkably little commonality in the methodology employed by these studies. Their single linking feature is that, in one way or another, they all enquire about the level of life satisfaction.

Calculation of a normative statistic

Table I presents the conversion of each data-set into the common statistic 'Percentage of Scale Maximum' (%SM). The data in this table have been generated through very diverse sample sizes, using widely differing techniques, as has been described in the previous section. Thus, there seems no compelling rationale to employ any weighting system to their combination. The arithmetic mean and standard deviation of the %SM values is 75.02 ± 2.74.

Comparative Sub-population Studies

Table II presents data organized under ten headings which represent population sub-groups.

As in the normative studies, there is little commonality among the measuring instruments other than the fact that they all measure life satisfaction by one means or another. Points to note in relation

to each of the sub-groups are as follows:

Females:	The data indicate a mean and standard deviation which is very similar to the normative value.
Aged > 65 y:	The mean value of 69.55 lies within two standard deviations of the normative values but the degree of variation is considerable.
Physical disability:	The group data may be unduly elevated by the Stensman (1985) data indicating 80.0% satisfaction. The normal control group for this study scored 83%SM, but the precise nature of the question asked to generate these data is not stated. Nevertheless, even with the inclusion of this value, the mean score for the group lies around two standard deviations below the normative mean.
Chronic medical:	This very heterogeneous group display a high degree of variability and a mean that lies four standard deviations below the normative sample.
Post-organ transplant:	With one exception these values lie within, or in one case above, the normal range.
Intellectual disability:	All values lie within the normal range.
Low income:	All values lie towards the lower side of the normal range.
High income:	This is the only sub-group where all values lie above the normal range.
Whites (USA):	All values are firmly placed at the upper margin of the normal range.
Blacks (USA):	The values tend to lie below the normative mean.

TABLE III
Other countries

Study	Country	Sample	Scale	N	Mean score	% SM
Nathawat & Mathur (1993)	India	Housewives	A2	200	4.46	57.7
		Working women	A2	200	4.73	62.2
Shmotkin (1991)	Isreal	Males 21–59y	B2	?[1]	7.90	69.0
			A6		13.44	61.3
		Females 21-59y	B2		8.06	70.6
			A6		14.18	65.9
		Males 60–87y	B2		7.39	63.9
			A6		12.82	59.1
		Females 60–87y	B2		6.79	57.9
			A6		10.70	48.5
Makarczyk (1962)	Poland	Stratified quota	C1	2,387	3.60	65.0
Man (1991)	Hong Kong	Children 13y	A10	940	3.88	48.0
		Children 16y	A10	940	3.63	43.8
Moller (1988)	South Africa	Retired Zulu				
		<64y	C2	125	2.36	34.0
		>65y		127	2.73	43.3
Moller (1992)	South Africa	Black youth	C2	1,200	3.21	55.3

[1] The total sample for this study was 447 but the sub-sample N was not stated.

Scales A1 to A13 and B1 to B3 have been described below Tables I and II respectively. The other scale description is as follows:
C1 – "On the whole are you satisfied with life?" using a 5-point scale from 'definitely yes' to 'definitely no'.
C2 – "Taking all things together, how satisfied are you with your life as a whole these days?" using a 5-point scale from 'very dissatisfied' to 'very satisfied'.

Data from other Countries

Table III presents data from countries other than those included in either the normative or sub-population tables. As can be seen, the samples are extremely diverse and so no overall statistic has been calculated. It is notable, however, that the values generally lie well below the normative range.

The Satisfaction with Life Scale

A deliberate omission from previous tables has been data derived from the Satisfaction with Life Scale (Diener *et al.*, 1985). The reason is that this scale appears to yield values that lie significantly below the other scales that have been reported.

Responses on Diener's scale are made to five items on a 7-point Likert format that is anchored by 'strongly disagree' and 'strongly agree'. Thus, scale scores may vary between 5 and 35. The statements that comprise this scale are as follows:

1. In most ways my life is close to ideal.
2. The conditions of my life are excellent.
3. I am satisfied with my life.
4. So far I have gotten the important things I want in life.
5. If I could live my life over, I would change almost nothing.

In 1993, Pavot and Diener published a compilation of data derived from the use of this scale. Including only those scores derived from the countries represented in Table I & II, the 22 values aggregate to a mean of 56.2%SM with a standard deviation of 13.3. The upper range extends to only 67%SM (college students and nuns) while the lower range extends to 24.3% (male prisoners) and 22.7%SM (medical inpatients). A total of six studies report on college students (mean = 63.6%SM), three on members of the adult work-force (mean = 64.2%SM), and three on medical patients/psychological clients (mean = 40.1%SM). It seems clear that this scale is yielding data which fall at least 10%SM points below those of comparable scales.

DISCUSSION

The main result to emerge from this paper is the finding that, despite the use of very different methodologies, the combination of data from 16 unrelated studies into life satisfaction has yielded a mean of 75.02%SM and a standard deviation of just 2.74.

One explanation for this result could be the existence of a psychological, homeostatic mechanism that maintains an average level of

life-satisfaction at around 75%SM. This would be a highly adaptive device on a population basis ensuring that, under relatively stable but diverse living conditions, most people feel satisfied with their lives, thereby conferring a non-zero sum benefit on the population as a whole. Such a mechanism has been suggested previously (Headey and Wearing, 1988). However, it would clearly only apply on an averaged basis across populations. Many individuals or sub-population groups fall outside the normal range, defined by two standard deviations from the mean, as indicated by Table II.

The data presented in Table II, derived from non-probability samples, need to be interpreted cautiously. These studies do not generally have the advantage of large numbers of subjects, often were not intended to be representative even of their sub-population categorization as depicted in Table II, and have employed an even more diverse set of measuring instruments. Despite these caveats, some trends seem to be apparent.

Firstly it can be noted that some of these sub-groups do not differ from the normative range. This categorization applies to females, people who have survived an organ transplantation for a number of years, and people with a mild or moderate level of intellectual disability living in the community. In contrast, one group stands out as having a level of life satisfaction above the normative range. The three studies on people with higher-than-average income yielded a score of 83.8 ± 1.15, indicating a remarkable degree of consistency between the three sets of data.

On the downward side, people over the age of 65 years, people with a physical disability, or with a low income, are on the lower margins of the normal range. People with a chronic medical condition, on the other hand, have a mean value well below the two standard deviation range. In addition it would appear that blacks in the USA have a lower satisfaction than whites.

While there is a relatively high degree of variability within most of these sub-groups, the trends of these data do not seem surprising. This is encouraging for the idea that the %SM data are valid and that this measure of life satisfaction can be used to discriminate between people whose lives are objectively better or worse off than normal.

The two final sets of data are drawn from studies involving people from countries other than those used for the normative calculations, or from studies using the Satisfaction with Life Scale. Both yield values that lie well below the normative range. The reason for the former set is speculative at this stage. Whether people from these countries score lower due to cultural factors that influence their responding or whether their responses reflect a relatively lower standard of living remains to be determined. Similarly, the reason that Diener's scale produces such low values is unclear. It seems rather extraordinary that this scale should behave in a manner noticeably different from the multitude of others that have been cited, but the fact that it does is most interesting. Clearly there is some element in the format of this scale, such as reaction to the strong statements of positive life satisfaction that constitute the items, that causes people to respond differently from other scales.

One final statistic can be drawn from the data presented in Table II in support of an observation made by Headey and Wearing (1988). Since the human set-point for life satisfaction lies above the point of neutrality, ceiling effects should start to reduce sample variance as the sample means become higher. This has been tested through the calculation of a Spearman Rank Correlation Coefficient between the rank order of sub-population means in Table II against the rank order of their variance. This latter statistic was adjusted for absolute values by using the standard deviation divided by the mean. The resulting $r = 0.717$ ($N = 9$) is significant at $p < 0.05$. This result is further evidence of coherent variation among the sub-population %SM values listed in this table.

In terms of using the information presented in this paper to guide future studies, it is recommended that researchers standardize the form of their 'life satisfaction' question and the nature of their scale. In the absence of data indicating the superiority of one form of question over another, researchers who seek a global measure should consider using the original Andrews and Withey (1976) wording "How do you feel about your life as a whole?" It is further recommended that any variation on this wording, or the use of framing procedures for such a question, be constructed with a stated purpose. It is also recommended that, in the absence of alternative argument, researchers adhere to the original 7-point Delighted-Terrible scale.

The reason for this is that: (**a**) Adults, on average, could be expected to reliably locate their level of satisfaction on such a scale, (**b**) There are currently no data to indicate the reliability of scales that exceed 7-points. Moreover, Ramsay (1973) has argued, from a psychometric perspective, that using 7 categories on a Likert Scale provides very nearly as much precision of estimate as a corresponding continuous judgement task, and that increasing the number of categories beyond 7 is of little value in terms of increased precision of estimate. On the other hand, reducing the number of scale points below 7 is likely to reduce both validity (Andrews and McKennel, 1980) and discriminative power, (**c**) The words 'delighted-terrible' used as the scale reference-points, have been shown to produce greater discrimination than such other markers as 'very satisfied' and 'very unsatisfied' (Andrews and Withey, 1976).

It can also be observed that, if researchers are interested only in an overall life-satisfaction score, there seems little benefit in asking respondents multiple questions. It seems that a single question can yield reliable and valid data.

In considering the accuracy of the normative statistics that have been produced, it is clear that they must be viewed only as approximations. The use of multiple instruments, varying instructions to respondents, different populations and even nationalities, means that the data have been drawn from an extremely heterogeneous base. In addition, the simple arithmetic combination of mean values from studies with such differing sample sizes is clearly a crude method of averaging.

Because of these factors it is proposed that, as a working hypothesis, the life-satisfaction gold standard be considered as 75.0 $\pm$ 2.5%SM. This mean value closely approximates the calculated mean, and the range of two standard deviations includes all of the values in Table I. The utility of this proposition is that is exemplifies the approximate nature of the proposed standard in a way that is readily applied to future comparative studies.

To conclude that more research is needed seems rather obvious. At least now, however, the %SM statistic and the data presented in this paper may provide the rationale for comparative empirical studies based on population norms. One trail to the gold standard for life

satisfaction is now open. The task of elaborating the hypothesised range of 75 ± 2.5%SM remains as a major challenge.

REFERENCES

Aaronson, N. K., W. Bakker, A. L. Stewart, F. S. A. M. Van Dam, Y. Van Zandwijk, Jr. and A. Kirkpatrick: 1987, 'Multidimensional approach to the measurement of quality of life in lung cancer clinical trials', in N. K. Aaronson and J. Beckmann (eds.), The Quality of Life of Cancer Patients (Raven Press, New York), pp. 63–82.

Adams, D. L.: 1969, 'Analysis of a life satisfaction index', Journal of Gerontology 24, 470–474.

Andrews, F. M.: 1991, 'Stability and change in levels and structure of subjective well-being: USA 1972 and 1988', Social Indicators Research 25, 1–30.

Andrews, F. M and A. C. McKennel: 1980, 'Measures of self-reported well being: Their affective, cognitive, and other components', Social Indicators Research 8, 127–155.

Andrews, F. M. and S. B. Withey: 1976, Social Indicators of Well-being: Americans' Perceptions of Life Quality (Plenum Press, New York).

Atkinson, T.: 1982, 'The stability and validity of quality of life measures', Social Indicators Research 10, 113–32.

Baker, F. and J. Intagliata: 1982, 'Quality of life in the evaluation of community support systems', Evaluation and Program Planning 5, 69–79.

Bortner, R. W. and D. F. Hultsch: 1970, 'A multivariate analysis of correlates of life satisfaction in adulthood', Journal of Gerontology 25, 41–47.

Boucher, J. and C. E. Osgood: 1969, 'The Pollyanna hypothesis', Journal of Verbal Learning and Verbal Behavior 8, 1–8.

Bowling, A.: 1990, 'Associations with life satisfaction among very elderly people living in a deprived part of inner London', Social Science and Medicine 31, 1003–1011.

Bowling, A., M. Farquhar and P. Browne: 1991, 'Life satisfaction and associations with social network and support variables in three samples of elderly people', International Journal of Geriatric Psychiatry 6, 549–566.

Bramston, P.: 1994, Personal communication.

Branholm, I. B. and E. A. Degerman: 1992, 'Life satisfaction and activity preferences in parents of Down's Syndrome children', Scandinavian Journal of Social Medicine 20, 37–44.

Brickman, P., D. Coates and R. Janoff-Bulman: 1978, 'Lottery winners and accident victims: Is happiness relative?', Journal of Personality and Social Psychology 36, 917–927.

Brief, A. P., A. H. Butcher, J. M. George and K. E. Link: 1993, 'Integrating bottom-up and top-down theories of subjective well-being: The case of health', Journal of Personality and Social Psychology 64, 646–654.

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

Cantril, H.: 1965, The Pattern of Human Concerns (Rutgers University Press, New Jersey).

Cummins, R. A.: 1993, The Comprehensive Quality of Life Scale: Intellectual Disability. Fourth edition (School of Psychology, Deakin University, Melbourne).
Cummins, R. A., M. P. McCabe and Y. Romeo: 1994, 'The Comprehensive Quality of Life Scale – Intellectual Disability: Results from a Victoria Survey', Proceedings, 28th National Conference of the Australian Society for the Study of Intellectual Disability, pp. 93–98.
Diener, E., R. A. Emmons, R. J. Larson and S. Griffin: 1985, 'The satisfaction with life scale', Journal of Personality Assessment 49, 71–75.
Ferrans, C. E. and M. J. Powers: 1992, 'Psychometric assessment of the quality of life index', Research in Nursing and Health 15, 29–38.
Flanagan, J. C.: 1978, 'A research approach to improving our quality of life', American Psychologist 33, 138–147.
Gill, W. M.: 1984, 'Monitoring the subjective well-being of chronically ill patients over time', Community Health Studies 8, 288–296.
Goldings, H. J.: 1954, 'On the avowal and projection of happiness', Journal of Personality 23, 30–47.
Gover, W. R., M. Hughes and C. B. Style: 1983, 'Does marriage have positive effects on the psychological well-being of the individual?', Journal of Health and Social Behavior 24, 122–131.
Hall, J.: 1973, 'Measuring the quality of life using sample surveys', in J. Stober and D. Schumacher (eds.), Technology Assessment and Quality of Life (Elsevier, Amsterdam), pp. 93–102.
Harris, L. and Associates: 1975, The Myth and Reality of Aging in America. (National Council on Aging, Washington, DC).
Headey, B. and A. Wearing: 1988, 'The sense of relative superiority: Central to well-being', Social Indicators Research 20, 497–516.
Headey, B. and A. Wearing: 1992, Understanding Happiness: A Theory of Subjective Well-Being (Longman Cheshire, Melbourne).
Headey, B. and A. Wearing: 1994, Victorian Quality of Life Panel Study, personal communication.
Helson, H.: 1964, Adaption-Level Theory (Harper & Row, New York).
Hicks, F. D., J. L. Larson and C. E. Ferrans: 1992, 'Quality of life after liver transplant', Research in Nursing & Health 15, 111–119.
Hoyert, D. L. and M. M. Seltzer: 1992, 'Factors related to the well-being and life activities of family caregivers', Family Relations 41, 74–81.
Huxley, P. and R. Warner: 1992, 'Case management, quality of life and satisfaction with services of long-term psychiatric patients', Hospital and Community Psychiatry 43, 799–802.
Judge, T. A. and S. Watanabe: 1993, 'Another look at the job satisfaction-life satisfaction relationship', Journal of Applied Psychology 78, 939–948.
Keith, K. D., R. L. Shalock and K. Hoffman: 1986, Quality of life: Measurement and Programmatic Implications (Mental Retardation Services, Nebraska).
Kober, B., T. Kuchler, C. Broelsch, B. Kremer and D. Henne-Bruns: 1990, 'A psychological support concept and quality of life research in a liver transplantation program: An interdisciplinary multicenter study', Psychotherapy and Psychosomatics 54, 117–131.
Koch, U. and F. A. Muthny: 1990, 'Quality of life in patients with end-stage renal disease in relation to the method of treatment', Psychotherapy and Psychosomatics 54, 161–171.

Krause, J.S.: 1992, 'Life satisfaction after spinal cord injury: A descriptive study', Rehabilitation Psychology 37, 61–70.
Lough, M. E., A. M. Lindsey, J. A. Shinn and N. A. Stotts: 1985, 'Life satisfaction following heart transplantation', Journal of Heart Transplantation 4, 446–449.
Man, P.: 1991, 'The influence of peers and parents on youth life satisfaction in Hong Kong', Social Indicators Research 24, 347–365.
Makarczyk, W.: 1962, 'Factors affecting life satisfaction among people in Poland', Polish Sociological Bulletin 1, 105–116.
Mastekaasa, A.: 1992, 'Marriage and psychological well being: Some evidence on selection into marriage', Journal of Marriage and the Family 54, 901–911.
Mastekaasa, A.: 1994, Personal communication.
Moller, V.: 1988, 'Quality of life in retirement: A case study of Zulu return migrants', Social Indicators Research 20, 621–658.
Moller, V.: 1992, 'Spare time use and perceived well-being among black South African youth', Social Indicators research 26, 309–351.
Mookherjee, H. N.: 1992, 'Perceptions of well-being by metropolitan and non-metropolitan populations in the United States', Journal of Social Psychology 132, 513–524.
Moum, T.: 1994, 'Is subjective well-being a predictor of nonresponse in broad population surveys?', Social Indicators Research 32, 1–20.
Muthny, F. A., U. Koch and S. Stump: 1990, 'Quality of life in oncology patients', Psychotherapy and Psychosomatics 54, 145–160.
Nathawat, S. S. and A. Mathur: 1993, 'Marital adjustment and subjective well-being in Indian-educated housewives and working women', The Journal of Psychology 127, 353–358.
Neugarten, B. L., R. J. Havighurst and S. S. Tobin: 1961, 'The measurement of life satisfaction', Journal of Gerontology 16, 134–143.
Oliver, J.P.J.: 1988, 'The quality of life of the chronically mentally disabled', Research Progress Report, cited by P. Huxley and R. Warner: 1992, 'Case management, quality of life, and satisfaction with services of long-term psychiatric patients', Hospital and Community Psychiatry 43, 799–802.
Oppong, J. R., R. G. Ironside and L. W. Kennedy: 1988, 'Perceived quality of life in a centre-periphery framework', Social Indicators Research 20, 605–620.
Palmore, E. and C. Luikart: 1972, 'Health and social factors related to life satisfaction', Journal of Health and Social Behaviour 13, 68–80.
Pavot, W. and E. Diener: 1993, 'Review of the satisfaction with life scale', Psychological Assessment 5, 164–172.
Ramsay, J. O.: 1973, 'The effect of number of categories in rating scales on precision of estimation of scale values', Psychometrika 38, 513–532.
Roy Morgan Research: 1993, International Values Audit, 22/23 May (Roy Morgan Research Center, Melbourne).
Schalock, R. L., K. D. Keith, K. Hoffman and O. C. Karan: 1989, 'Quality of life: Its measurement and use', Mental Retardation 27, 25–31.
Schulz, R. and S. Decker: 1985, 'Long-term adjustment to physical disability: The role of social support, perceived control, and self-blame', Journal of Personality and Social Psychology 48, 1162–1172.
Schmotkin, D.: 1990, 'Subjective well-being as a function of age and gender: A multivariate look for differentiated trends', Social Indicators Research 23, 201–230.

Shmotkin, D.: 1991, 'The structure of the Life Satisfaction Index A in elderly Israeli adults', International Journal of Aging and Human Development 33 131–150.
Stensman, R.: 1985, 'Severely mobility-disabled people assess the quality of their lives', Scandinavian Journal of Rehabilitation Medicine 17, 87–99.
Taylor, S. E., J. V. Wood and R. R. Lichtman: 1983, 'It could be worse: Selective evaluation as a response to victimization', Journal of Social Issues 39, 19–40.
White, J. M.: 1992, Marital status and well-being in Canada – An analysis of age group variations', Journal of Family Issues 13, 390–399.
Ying, Y-W.: 1992, 'Life Satisfaction among San-Francisco Chinese-Americans', Social Indicators Research 26, 1–22.

School of Psychology,
Deakin University,
Victoria, 3125,
Australia