The Stability and Validity of Quality of Life Measures

Author(s): Tom Atkinson

Source: *Social Indicators Research*, Feb., 1982, Vol. 10, No. 2 (Feb., 1982), pp. 113-132

Published by: Springer

Stable URL: https://www.jstor.org/stable/27521003

TOM ATKINSON

THE STABILITY AND VALIDITY OF QUALITY OF LIFE MEASURES*

ABSTRACT. Effective social indicators must be stable when individual or societal characteristics are unchanged and dynamic when circumstances alter. Highly reliable measures may be poor indicators because they are insensitive to change. Little evidence is available on the sensitivity or validity of objective and subjective indicators. A lack of panel data has restricted the assessment of the stability of subjective measures.

This paper examines longitudinal data on a representative sample of 2162 Canadians interviewed in 1977 and again in 1979. Test-retest correlations of approximately 0.50 were obtained for satisfaction and self-anchoring ladder measures among respondents who reported no significant changes in their lives during the past two years. Correlations were substantially lower, as expected, for those reporting life changes. Comparisons of the absolute values of these subjective indicators show that very little change in quality of life measures occurs when stable circumstances are reported but the indicators rise or fall significantly when situations change with downward adjustments being more dramatic than upward modifications. Positive and negative life events had little effect on overall evaluations of life quality.

In general, these findings provide very strong evidence for the stability and validity of subjective indicators over time. These measures, with one exception, were constant in unchanging situations and sensitive to change when it occurred.

INTRODUCTION

Measurement issues have provided a central focus of social indicators research over the last decade. Those interested in developing objective quality of life (QOL) measures have puzzled over which subset of the plethora of available statistics should be included in their models, and how to combine them into general indexes comparable to popular aggregate economic measures such as the Gross National Product and Consumer Price Index. Students of subjective or perceptual indicators have been no less concerned with the selection of conceptually suitable measures but have not had to wrestle with the question of how to combine them into a general index. Global measures of perceived well-being can be easily obtained through direct measurement, and previous research has demonstrated that evaluations of specific areas of experience such as work, marriage, etc. combine in a linear fashion to produce general QOL perceptions (Campbell *et al.*, 1976; Andrews and Withey, 1976).

Instead, research on subjective indicators, particularly in North America, has been greatly concerned with the reliability and validity of these measures. Subjective social indicators have been exposed to closer methodological scrutiny than any other attitudinal or value measure with the possible exception of the *F*-scale. Several factors are responsible for this concern. First, social indicators, whether objective or subjective, are intended to enlighten many areas of social policy and the consequences of poor measurement extend far beyond the confines of disciplinary interests. Second, those surveying QOL perceptions of the general public have had to rebut attacks that argue that most people are incapable of making consistent and honest evaluations of their own lives. Third, because subjective measures are designed and utilized directly by the researchers, they can be easily modified if proven inadequate, while objective indicators are based on data gathered by various government agencies for administrative purposes and the measurement process is beyond the control of those developing indicators. Finally, most researchers working on subjective measures have been trained as psychologists and are very familiar with the assessment of reliability and validity. Sociologists and economists who develop objective measures are more adept in the construction of statistical models which delineate causal influences and tend to ignore questions about the quality of their measures.

Assessments of the reliability and validity of subjective indicators have, with one major exception, been limited to analyses of cross-sectional surveys. Published studies provide substantial evidence that several QOL measures correlate well with each other, are relatively free of method effects and are not significantly affected by response biases such as social desirability. However, the absence of panel data on large representative samples has prohibited the evaluation of the stability of QOL measures over time and their sensitivity to the changing circumstances of people's lives. Indicators which are very stable over time may be poor candidates for systems meant to monitor change.

This paper presents an analysis of a representative sample of over 2000 Canadians interviewed in 1977 and again in 1979 about perceptions of their lives in general and specific areas or domains of them. Data is presented on the stability of several QOL measures and on their behaviour when individuals experience change in their lives during the two year period. The effects of several types of significant events on QOL assessments are briefly examined.

A. Previous Research

North American research on subjective indicators has made significant contri-
butions to our understanding of the reliability and validity of several QOL
measures. With the exception of one study, the methodological research on
these measures have involved cross-sectional rather than panel data and have
often included Frank Andrews as an investigator.

First, with associates at the Institute for Social Research (Andrews and
Withey, 1976; Andrews and Crandall, 1976), and more recently with Aubrey
McKennell and this author (Andrews and McKennell, 1980; McKennel,
Atkinson and Andrews, forthcoming), he has examined reliability and validity
using a methodology based on a multi-method, multi-trait matrix approach
(Campbell and Fiske, 1959). A software package (LISREL) developed by
Joreskog and Sorbom (1978) identifies the elements of an item score which
result from trait variance, from method variance and from unique or random
variation. Validity is the trait coefficient and reliability is the summed square
of the trait and method effects.

Since reliability computed in this fashion roughly approximates standard
reliability coefficients, it is possible to compile the following summary of
item reliabilities from national surveys in Canada and the United States
(exact question formats appear in Appendix A):

Measures	U.S.	Canada
Satisfaction		
7-point	0.59	0.68[a], 0.63
11-point		
Self-anchoring Ladder	0.54	0.56
Happiness (3-point)	0.52	
Delighted-Terrible	0.67[a], 0.70	

[a] Test-retest computation.

All estimates are fairly high with the classical test-retest cofficients suggesting
reliabilities of about 0.70. Of the variance accounted for by method and trait
factors, e.g. the reliability, 15 to 20% appears to result from method factors,
the remainder from trait factors. These analyses indicate fairly high reliability
and convergent validity for QOL measures and demonstrate that the correla-
tions result primarily from trait, not method, variance.

Follow-up interviews with a small subset ($N = 285$) of a national sample in
the United States have produced the only published data on the stability of

QOL measures over time (Campbell *et al.*, 1976; Rodgers and Converse, 1976). In brief, the test-retest correlations over an eight month period were as follows:

Index of Domain Satisfaction	0.76
Index of General Well-Being	0.53
Average Domain Satisfaction	0.53
General Life Satisfaction	0.42
Happiness	0.38

Although the stability of the two indexes and the specific domain satisfaction measures were high given the fairly long time lapse between interviews, the correlations for the life satisfaction and happiness items are cause for concern. Campbell and his colleagues point out, however, that these estimates provide the lower bounds of stability over time. Given the passage of eight months between interviews, it is likely that changes occurred in the lives of many individuals thus reducing the size of the correlation. From this perspective, a very high stability coefficient, e.g. the Domain Satisfaction Index, may disqualify a measure as a social indicator because of its insensitivity to change.

The present research continues the investigation of the performance of QOL measures over time begun by Campbell, Converse and Rodgers. A large panel survey allows the stability of QOL measures to be examined over a two year period for those reporting changes in their lives and those who do not report changed circumstances. The correlation of measures in the No Change condition provides a fair test of the stability of these indicators. In addition, comparisons of the difference and mean scores for Change and No Change groups will indicate the stability of measures in absolute terms and their sensitivity to changing conditions. Both attributes must be demonstrated if these measures are to be useful indicators of well-being.

Although this analysis does not fit the hypothesis-testing mold, three results are expected:

(a) the stability correlations will be higher for the No Change than the Change respondents;

(b) the No Change group will show little mean change in QOL measures over the two year period; and,

(c) those reporting changed circumstances will experience significant changes between 1977 and 1979 QOL indicators — increases for those in improved circumstances and declines for those in deteriorating conditions.

II. METHOD

A. Subjects

In 1977, a multi-stage probability sample of Canadian dwelling units yielded a sample of 3288 respondents. The response rate for the national sample was 70%. When weighted to compensate for over-sampling in some urban areas and response rate biases, the sample is representative of the non-institutionalized, adult (age 18 and older) population (Murray and Atkinson, forthcoming). The second wave of the study was launched in 1979 with a sample design which called for reinterviews of all 1977 respondents and new interviews in all dwellings from which the previous respondents had moved. Two thousand one hundred and sixty-two of the 1977 respondents (66%) were reinterviewed. Of this number, 1669 (77%) resided in the same location while 493 (23%) had moved. Of the 34% who were lost from the original sample, 16% were located but refused the interview, 7% had moved and could not be located and 6% were ill, aged or had died, and 5% were temporarily absent or unavailable for other reasons.

B. Measures

All interviews were conducted by trained interviewers of the Survey Research Centre at York University and sessions lasted an average of two hours. The interview was divided into sections which obtained information on the 'facts' of the respondents circumstances in each of twelve domains, e.g. work, leisure, housing, and included several questions requiring evaluations of that domain or specific aspects of it. A separate section dealt with assessments of life in general.

Three evaluations of general well-being were included in both interviews (See Appendix A for complete question text). They were:

(a) A life satisfaction measure which was similar but not identical to that used by Campbell *et al.* in their 1971 survey. The two questions were identically worded, but the Canadian survey used an eleven-point response scale while the U.S. version was restricted to seven points. The Canadian measure is identical to that used in Great Britain and in the European Community surveys.

(b) A self-anchoring ladder scale which was a modification of Cantril's

(1965) measure. The major difference between this measure and Cantril's is that his was anchored by the phrase 'the best possible life for you' allowing the interpretation that the imagined conditions must be obtainable by the respondent. The use of the 'ideal or perfect life' as an anchor in the current survey eliminates that ambiguity and permits personal definitions of the desired conditions without suggesting that such a life must be obtainable.

(c) A general happiness scale which has been used by the Gallup organization for several decades and has been used in several academically-based surveys including Bradburn (1969), Campbell *et al.* and the General Social Survey in the United States.

In each domain, respondents were asked to indicate their level of satisfaction using the same format as the general life measure.

Two indexes are computed from the QOL measures. The General Quality of Life Index combines responses from the life satisfaction, ladder and happiness measures while the Domain Satisfaction Index is an average of responses to the job, financial, housing, health and marriage/romance domains. These indexes are comparable to those developed by Campbell *et al.* which were more stable over time than individual items.

In addition to these evaluation questions, each section in the 1979 questionnaire included items requiring a comparison between respondent's current situation and that two years previous. The questions asked whether the current situation, e.g. their job, was better, worse or the same as two years ago. These items provide a summary measure of change in several areas of life and allow the respondent to interpret the impact of changes.

Finally, in 1979 each person was asked whether any of sixteen events had occurred in their lives during the past five years. Seven events are assumed to be negative — divorce/separation, romantic breakup, death of a family member, death of a close friend, serious injury of illness, lose touch with a close friend and a large decrease in income — while the other nine were positive — recent marriage, having children, a romance, new friends, new job, job promotion, on honour or award, changing residences and having a large income increase. If an event had taken place, they were asked to provide the date of the most recent occurrence and to indicate its effect on their lives. In this analysis, the effects of events occurring within the last two years i.e. since the 1977 interview, on repeated QOL measures is examined. Events which had occurred in the twelve months prior to the 1979 survey (labelled

78—79) are distinguished from those which took place between one and two
years ago (labelled 77—78).

III. RESULTS

Any analysis of panel data must be cognizant of the effects of sample mor-
tality. While the use of individual differences scores in these analyses elimi-
nates the possibility that over time changes may result from differential
dropout rates, sample mortality may reduce the representative nature of the
panel if it is not randomly distributed. Table 1 compares the unweighted
characteristics of the entire 1977 sample with those of the panel which was
reinterviewed in 1979.

TABLE I

National sample and panel demographic characteristics

		1977 National sample	1977—79 panel
Gender –	Male	41%	41%
	Female	59	59
Family income –	Under $8000	22	19
	8000—13 999	24	24
	14 000—19 999	23	23
	20 000—up	31	34
Education –	Primary School or less	21	20
	Some Secondary	25	23
	Completed Secondary	30	31
	Some University	13	14
	Completed University	11	12
Employment status –	Working Full-time	46	46
	Working Part-time	10	11
	Not-Working	44	43
Marital status –	Married	67	68
	Widowed	8	8
	Divorced/Separated	7	7
	Never married	18	17
Age– 18—19		4	4
20—29		25	24
30—39		22	23
40—49		16	17
50—59		14	15
60—69		11	11
70+		7	6

Distributions shown in Table I indicate a high degree of similarity between the original sample and the 66% which was reinterviewed in 1979. The two samples are virtually identical on all attributes except family income where the panel has a slightly higher average income than the original sample ($17 950 to $17050). It is surprising that the age distributions are so similar given that younger persons are more mobile and harder to recontact while the aged are more likely to die or be incapacitated. Similarities between the two samples assure that, when weighted, both are representative of the Canadian population.

The stability of QOL measures is mot stringently tested by the zero-order correlations presented in Table II. Correlations were computed for the entire sample and separately for those who reported some changes in their circumstances and those whose situations did not change. Stability is indicated by the correlations for the latter group because these measures, if they are valid, should not correlate as highly for the change group.

TABLE II

Correlations of QOL measures repeated after two years

	Total sample	No change group	Change group	Z-score difference
Evaluations of Life in general				
Satisfaction scale	0.41	0.43	0.39	1.19
Self anchoring ladder	0.40	0.47	0.36	3.09[b]
Happiness scale	0.39	0.34	0.41	−1.91
General QOL index	0.53	0.55	0.52	0.98
Domain evaluations				
Job satisfaction	0.42	0.48	0.37	2.11[a]
Financial satisfaction	0.49	0.49	0.50	−0.34
Housing satisfaction	0.38	0.51	0.09	8.76[b]
Health satisfaction	0.53	0.56	0.45	2.96[b]
Marriage/Romance satisfaction	0.45	0.50	0.27	4.61[b]
Domain satisfaction index	0.58	0.60	0.56	1.40

[a] $p < 0.05$.
[b] $p < 0.01$.

Table II indicates that the stability of all general measures, except the happiness scale, is high — between 0.40 and 0.55 — for the sample as a whole. Domain measures were in the same range except for housing where the

greatest amount of actual change occurred. Correlations for the No Change group were usually higher than those who experienced changes. These differences are very much in evidence among the domain specific measures suggesting that such indicators are more sensitive to change than terms requiring evaluations of very general conditions such as 'your life as a whole'.

Comparisons of these results with those reported by Campbell *et al.* show that the general assessments performed in virtually identical ways in the Canadian study even though the span between surveys was two years rather than eight months as in the earlier work. Our domain measures were less strongly correlated than those in the smaller U.S. study. Average correlations for the five areas were 0.58 in the U.S. versus 0.45 for the total sample and 0.51 for the No Change group in Canada.

This patterning of correlations in the current study supports the contention that domain ratings are more firmly rooted in the specific circumstances of the individual's life. As a result, stability coefficients for such measures decline more quickly than general indicators as the time period between observations increases. Rather than reducing the attractiveness of specific measures, this feature, and the marked differences between Change and No Change groups, recommends them as reliable and sensitive measures. The two indexes were more stable than the individual terms but may not be the best measurement tools if detecting change is a major objective of the research.

The weakest item assessed here is the happiness measure — it has the lowest stability correlations of the general measures and the No Change group had lower correlation than those whose lives had changed. The poor performance of this scale may result from a methodological fault or from differences between affective measures such as happiness and primarily cognitive scales like satisfaction. While affect may not be as good a social indicator as cognition, it is more likely that the item's problems stem from the use of a three-point response scale. Future research should utilize lengthier continuums with this measure rather than replicate the short scale because of its long history. A poor measure used twenty times in twenty times a poor measure.

In general, these stability coefficients, except the happiness item, are very encouraging given the lengthy period between the two administrations. These figures represent conservative estimates of stability because, although we attempted to eliminate persons who were experiencing changed circumstances, most peoples' lives changed during a two year period if for no

TOM ATKINSON

other reason than they, their children, their houses, etc. have aged two years. Those using this type of satisfaction scale or ladder measure, i.e. ten or eleven-point scales, can be assured that they are fairly stable over long periods of time. Items which employ more abbreviated response continuums such as Campbell's *et al.* satisfaction measure and the Andrews/Withey Delighted-Terrible, probably enjoy similar reliability but may not share their sensitivity to change.

As suggested earlier, candidates for social indicator systems must not only be proven reliable by correlational tests but must have demonstrated stability and sensitivity in *absolute* terms. In Table III individual difference scores for the general QOL measures are analyzed by two measures of change — self-report change and the occurance of positive and negative events. The hypothesis that the No Change group would have more stable QOL scores than those reporting change is supported when the perceptual indicator of change is used as the independent variable but not when the event measures are involved. In most of the cases, however, the events produce difference patterns which are in the expected direction and approach statistical significance.

Differences in general life satisfaction fit the expected pattern almost perfectly with virtually identical scores for the No Change group. The negative change group suffered a larger drop in satisfaction than the gain experienced by the positive change group — a pattern which reappears throughout the following analysis. This finding in conjunction with the large proportion of persons who felt that their lives had improved over the two year period, suggests that people anticipate at least small improvements in life quality and that a deterioration produces a sharp drop in satisfaction, while improvements result in more modest increases.

It is also noteworthy that the positive change group had higher levels of satisfaction in 1977 and exceeded the average to a greater degree in 1979. Similarly the negative change group began below the mean and declined from there. The same pattern holds for all general Quality of Life assessment — the group reporting improvement between 1977 and 1979 were above average in the first survey and the margin increased during the two years. Apparently very high or low quality of life levels are the product of trends covering years rather than short term reversals of fortune. Future surveys should indicate the length of these cycles and identify the life events responsible for them.

The self-anchoring ladder scale shows the same difference patterns except

TABLE III

General QOL measures by perceived life change and significant events

	N	Means 1977	Means 1979	Differ-ence	Eta
GENERAL LIFE SATISFACTION					
Perceived change in life					
Better	1129	8.85	9.07	0.22	0.12[b]
Same	845	8.57	8.58	0.01	
Worse	155	7.68	7.05	−0.63	
Positive events					
0	557	8.67	8.69	0.02	
1 or 2	1113	8.69	8.73	0.04	0.06
3 or 4	374	8.58	8.75	0.17	
5+	86	8.29	8.86	0.57	
Negative events					
0	808	8.83	8.84	0.01	
1	741	8.65	8.79	0.14	
2	406	8.52	8.60	0.08	0.03
3+	175	8.12	8.22	0.10	
SELF-ANCHORING LADDER					
Perceived change in life					
Better	1129	7.29	7.73	0.44	
Same	845	7.12	7.45	0.33	0.17[b]
Worse	155	6.69	5.97	−0.72	
Positive events					
0	557	7.19	7.53	0.34	
1 or 2	1113	7.23	7.52	0.29	
3 or 4	374	7.06	7.40	0.34	0.02
5+	86	6.97	7.35	0.38	
Negative events					
0	808	7.28	7.66	0.38	
1	741	7.20	7.56	0.36	0.05
2	406	7.11	7.28	0.17	
3+	175	6.78	6.94	0.16	
HAPPINESS (3-POINT)					
Perceived change in life					
Better	1129	2.55	2.53	−0.02	
Same	845	2.43	2.29	−0.14	0.12[b]
Worse	155	2.24	1.98	−0.26	
Positive events					
0	557	2.51	2.37	−0.14	
1 or 2	1113	2.48	2.41	−0.07	0.05
3 or 4	374	2.46	2.38	−0.08	
5+	86	2.33	2.36	0.03	
Negative events					
0	808	2.52	2.45	−0.07	
1	741	2.47	2.41	−0.06	0.06
2	406	2.46	2.31	−0.15	
3+	175	2.37	2.24	−0.13	

Table III (continued)

	N	Means 1977	1979	Differ- ence	Eta
GENERAL QUALITY OF LIFE INDEX					
Perceived change in life					
Better	1129	8.26	8.46	0.20	
Same	845	7.99	7.96	−0.03	0.18[b]
Worse	155	7.37	6.66	−0.71	
Positive events					
0	557	8.13	8.11	−0.02	
1 or 2	1113	8.12	8.15	0.03	
3 or 4	374	8.01	8.10	0.09	0.06
5+	86	7.74	8.16	0.42	
Negative events					
0	808	8.23	8.28	0.05	
1	741	8.08	8.19	0.10	0.05
2	406	8.00	7.94	−0.07	
3+	175	7.67	7.66	−0.01	

[a] $p < 0.01$.
[b] $p < 0.001$.

that the No Change group has significantly higher scores in 1979. Again the expectation of natural improvements in life quality may lead those who experience small increments to view their lives as substantially unchanged. While they recognize that they may have moved closer to some ideal life, they are not more satisfied because as much was expected.

Finally, the happiness measure presents the hypothesized pattern but the means decline for all three groups. The decline in happiness with increased age has been noted in several national studies but these drops are too large to result from a two year increase in the sample's average age. It may be that happiness levels in the population as a whole are declining while other QOL indicators remain steady or increase slightly.

Such divergent trends for the affective and cognitive indicators might account for an apparent paradox which often threatens this type of survey research. Most QOL surveys show relatively high levels of satisfaction in the face of the image of troubled, and perhaps declining, societies presented by the North American and Western European media and social critics. The general public also reflects a feeling that somehow the quality of life has declined. It may be that in cognitive terms 'we never had it so good' — higher standard of living, better health and education, etc. — but our affective morale

has declined. We, as individuals, may not be as happy as we used to be. Put differently, we are better off now but enjoying it less. These speculations are based on the supposition of a trend which requires more than two surveys to confirm, but they argue for the maintenance of improved affective measures in QOL surveys.

TABLE IV

Satisfaction with job, finances, housing and health by perceived change and significant events

	N	Means 1977	1979	Difference	Eta
JOB					
Perceived change in job					
Better	377	8.40	8.59	0.19	
Same	533	8.49	8.30	−0.19	0.18[b]
Worse	70	7.77	6.59	−1.28	
Job promotion					
Did not occur	754	8.32	8.25	−0.07	
77−78	70	8.68	8.36	−0.32	0.03
78−79	156	8.66	8.41	−0.25	
FINANCIAL SITUATION					
Perceived change in income					
Large increase	390	7.37	8.03	0.66	
No major changes	1614	7.41	7.51	0.10	0.12[b]
Large decrease	145	6.81	6.48	−0.33	
Family income change					
Increase of 40% +	479	6.98	7.39	0.42	
Up to 40% increase	907	7.55	7.62	0.07	0.08[a]
Decline in income	477	7.49	7.42	−0.07	
HOUSING					
Perceived changes in housing					
Better	375	7.58	8.64	1.06	
Same	1718	8.47	8.39	−0.08	0.25[b]
Worse	60	7.97	6.17	−1.80	
Changed dwellings					
Did not occur	1667	8.54	8.41	−0.13	
77−78	172	7.27	8.20	0.93	0.17[b]
78−79	316	7.63	8.27	0.64	
HEALTH					
Perceived change in health					
Much better	84	6.62	7.99	1.37	
Somewhat better	137	7.16	7.24	0.08	
Same	1626	8.10	8.11	0.01	0.23[b]
Somewhat worse	256	7.12	6.17	−1.02	
Much worse	50	6.14	4.99	−1.15	

Table IV (continued)

	N	Means 1977	1979	Difference	Eta
Serious illness or injury					
Did not occur	1918	7.93	7.86	−0.07	
77–78	93	6.90	6.46	−0.44	0.04[a]
78–79	142	7.07	6.99	−0.08	
DOMAIN SATISFACTION INDEX					
Perceived change					
Better	1129	8.18	8.37	0.19	0.11[b]
Same	845	8.08	8.11	0.03	
Worse	155	7.32	6.99	−0.33	
Positive Events					
0	557	8.15	8.14	−0.01	0.11[b]
1 or 2	1113	8.15	8.17	0.02	
3 or 4	374	7.89	8.19	0.30	
5+	86	7.52	8.06	0.54	
Negative events					
0	808	8.22	8.34	0.12	0.07[a]
1	741	8.09	8.25	0.16	
2	406	7.98	7.89	−0.09	
3+	175	7.61	7.63	0.02	

[a] $p < 0.01$.
[b] $p < 0.001$.

Turning to difference patterns in satisfaction with job, financial situation, housing and health, Table IV presents group means.

Several characteristics are shared by results from each domain.

(1) As expected, the self-report domain change measure is always significantly related to changes in satisfaction levels.

(2) The reported No Change group has remarkably steady satisfaction means in all domains.

(3) Negative change groups usually differ from the No Change groups to a greater extent than positive change groups.

(4) Relationships between self-reported change and satisfaction levels are always higher than between events and changes in satisfaction.

Events related to the housing and financial domains were associated with satisfaction changes but promotions did not increase job satisfaction nor did serious illness produce the expected large drop in health satisfaction. Since those experiencing negative health events during the past two years had low satisfaction levels in 1977, the illness or injuries are probably recurring ones

TABLE V

Satisfaction with marriage and romantic relationships by perceived change and significant events

	N	Means 1977	1979	Differ-ence	Eta
Perceived change in relationship					
Much better	172	8.55	9.28	0.73	
Somewhat better	134	8.29	7.97	−0.33	
Same	1503	8.94	8.69	−0.25	0.24[b]
Somewhat worse	42	7.09	5.02	−2.07	
Much worse	17	8.11	3.94	−4.17	
Got married					
Did not occur	1828	8.80	8.54	−0.26	
77–78	34	8.69	8.56	−0.13	0.07[a]
78–79	31	8.28	9.30	1.02	
Got divorced/separated					
Did not occur	1836	8.82	8.58	−0.24	
77–78	21	6.88	8.73	1.85	0.09[b]
78–79	36	8.13	7.15	−0.98	
Started romance					
Did not occur	1753	8.90	8.56	−0.34	
77–78	63	6.52	8.15	1.63	0.17[b]
78–79	77	8.09	8.72	0.63	
Ended romance					
Did not occur	1780	8.74	8.63	−0.11	
77–78	39	6.84	7.27	0.44	0.05
78–79	74	7.89	7.32	−0.57	

[a] $p < 0.01$.
[b] $p < 0.001$.

which had their major impact prior to the 1977 survey. In the work domain, promotions have a negative impact on job satisfaction although they are viewed as a positive event by most employees. Income changes have a significant but hardly overwhelming effect but the dependent measure is satisfaction with financial situation which is defined as involving income and expenditures. Improvement in income may be balanced by increased expenditures resulting in no real advancement.

Change in satisfaction with marriage and romantic relationships are presented separately in Table V because of the centrality of this domain to most individuals' quality of life and the amount of research published on the stressful nature of events involving status changes. Since each respondent was asked to indicate his or her satisfaction with whichever form of romantic

relationship they had at the time, it is possible to compare individuals whose marital status has changed during the two year period.

Patterns observed in Table V are similar to those evidenced in other domains. One remarkable difference is the catastrophic effect of deteriorating relationships on satisfaction levels. Although the number of persons affected is small, the significance of these changes on quality of life is evident. However, the effects of the four events measure here, while usually significant, are not large. The events do not account for the large drops in satisfaction for those with deteriorating relationships. Most of the observed decline seems to be taking place within legally defined statuses. As suggested below changes in marital status come toward the end of a down trend most of which occurs without major events.

The pattern observed for those getting divorced or separated and those who have concluded a close romantic relationship is mot intriguing. In both cases, the immediate impact of the event is negative while the longer term effects, i.e. one to two years, are positive. To determine whether the apparent improvement after the first year resulted from the establishment of new romantic relationships, those who had formed new liaisons were excluded from the analysis. Persons who did not establish new relationships showed the same gains in satisfaction as those who did. Reseachers and others who view these events as inherently stressful and destructive have apparently overlooked the possibility that a bad marriage or relationship may be very stressful and release from it a positive step.

IV. DISCUSSION

These results provide considerable assurance that subjective measures possess the two attributes required of good social indicators – stability in unchanging situations and sensitivity to changing circumstances. While the endorsement does not apply to all subjective measures, this analysis demonstrates the validity of satisfaction measures for general quality of life and more specific domain evaluations. The self-anchoring ladder scale also appears to be a good indicator of general life quality and it would serve as well for domain assessments. Some doubts are in order for the happiness measure used here and it did not receive as much support as the other two indicators.

Opinions have been expressed in some quarters that subjective measures such as satisfaction are poor social indicators because they were so conditio-

ned by expectations and restricted awareness as to be insensitive to changing circumstances. It is also argued that expectations and aspirations adjust very quickly to new situations and that satisfaction and other measures revert to their original levels immediately. This position would have led to predictions that (a) very few individuals would indicate any change in their situation, and/or (b) virtually no adjustment in subjective indicators would occur when changes did occur.

Our findings contradict both hypotheses in that significant numbers of respondents perceive changes in their lives and those changes were reflected, for better or worse, in their satisfaction levels. The fact that these changes took place over a two year period indicates that, while adaptation probably does occur, it is not instantaneous and will be detected by an indicator series which utilizes fairly frequent measurements. Future waves of this study should cast some light on the extent and rate of accomodation to new circumstances.

Both indexes of QOL had higher stability than the individual items suggesting their usefulness as indicators. In particular, the General QOL Index was more sensitive to personal change than its composite items. The Domain Satisfaction Index, however, adds little to the portrait obtained from the General Index and separate domain evaluations and is not crucial.

Finally, the relationships between actual changes in the QOL measures and perceived change was always higher than their association with specific events. It is often the case that low correlations between objective variables, in this case events, and subjective measures lead some users of social statistics to question the validity of the latter. This analysis suggests that it is the objective measure which may be insufficient. A case in point is the impact of promotions on job satisfaction. While advancement is usually viewed as a positive event, its effect on satisfaction is, if anything, negative. Many factors may explain this finding − more responsibility, longer hours, etc. − but the central point is that the objective event is too gross a measure to lead to a prediction of its effects. Additional information about changes in other job characteristics are required to understand personal reactions to work changes.

A second difficulty with objective measures is that the short-term consequences of objective changes may be in one direction and the long-term effects in the other. Divorce and separation appear to have neutral or negative short term impacts followed by a positive long-term consequence. Most analyses of the impact of changing conditions have not considered the possibility of

time dependent effects and have assumed that the impact of change is constant, or at least in the same direction, over time.

Those who expect objective changes to produce large aggregate changes in subjective indicators also assume that the impact of events is the same for most individuals. Although not reported here, our analysis showed that the immediate impact of separation and divorce is either very negative or very positive. In aggregate, these opposite responses to the same event produced a modest, but significant, decline in satisfaction. The assumption that most objective changes will produce the same effects ignores individual differences in values and judgement standards. It leads to the mistaken conclusion that the oft-reported lack of correlation between objective and subjective indicators results from the inadequacies of the latter.

V. CONCLUSIONS

In sum, this analysis should dispell several doubts about the utility of subjective social indicators. The measures used here, with the possible exception of the happiness scale, were shown to be stable in unchanging conditions and sensitive to change when it occurs. Given this evidence of the methodological soundness of these measures, two types of questions can now be raised. First, what is the extent and rate of adaptation to changing circumstances; and, second, what combinations of changing objective conditions lead to the judgement that one's situation is improving or deteriorating and cause changes in subjective indicators? An understanding of the dynamics of quality of life perceptions requires answers to both questions and panel studies are a more productive way of pursuing them than single corss sectional surveys.

Finally, although primarily methodological in focus, this analysis uncovered evidence of two trends which require examination. At the aggregate level, it appears that the trend in affective variables such as happiness may be down while satisfaction and other cognitive measures remain constant. If confirmed this divergence may lead to a better understanding of the way in which current quality of life experiences seem less rewarding than in 'the good old days'. At the individual level, these data suggest that low levels of perceived Quality of Life result from the cumulation of experiences which may require several years to amass. It is apparent that major life events make only a small contribution to this downward spiral and questions arise about the other depressants which gradually take their toll. A clear picture of the

scenarios which result in low QOL conditions could lead to a better notion of how to intervene to preserve mental health.

Institute for Behavioural Research,
York University, Toronto

NOTE

BIBLIOGRAPHY

Andrews, F. M. and Crandall, R.: 'The validity of measures of self-reported well-being', Social Indicators Research 3 (1976), pp. 1–20.

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

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

Bradburn, N.: The Structure of Psychological Well-Being (Aldine, Chicago, 1969).

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

Campbell, D. T. and Fiske, D. W.: 'Convergent and discriminant validities by the multitrait-multimethod matrix', Psychological Bulletin 56 (1959), pp. 81–105.

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

Joreskog, K. G. and Sorbom, B.: 'LISREL IV: Analysis of linear structural relationships by the method of maximum likelihood' (National Education Resources, Chicago, 1978).

McKennell, A. C., Atkinson, T., and Andrews, F. M.: 'Structural constancies in surveys of perceived well-being', in Szalia, A. and Andrews, F. M. (eds.), Comparative Studies on the Quality of Life (Sage Publications, London, forthcoming).

Murray, M. and Atkinson T.: 'Gender differences in correlates of job satisfaction', Canadian Journal of Behavioural Science (forthcoming).

Rodgers, W. L. and Converse, P. E.: 'Measures of the perceived overall quality of life', Social Indicators Research 2 (1976), pp. 127–152.

APPENDIX A: SUBJECTIVE QUALITY OF LIFE MEASURES

Satisfaction scale

All things considered, how satisfied or dissatisfied are you with your life as a whole? Which number comes the closest to how you feel?

1	2	3	4	5	6	7	8	9	10	11

Completely dissatisfied   Neutral   Completely satisfied

10	THE BEST
9	YOU CAN
8	IMAGINE
7	
6	
5	
4	
3	
2	
1	
0	THE WORST YOU CAN

Self-anchoring ladder scale

Here is a picture of a ladder. At the top of the ladder is the best life you can imagine – the ideal life. At the bottom of the ladder is the worst life you can imagine – a life that is terrible. Using a number on this card, where on the ladder would you place your life at this time?

Happiness item

Generally speaking, how happy would you day you are – very happy, fairly happy, or not too happy?

IMAGINE

Delighted-terrible scale

We want to find out how you feel about your life as a whole. 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.

I feel:

7	6	5	4	3	2	1
Delighted	Pleased	Mostly satisfied	Mixed (about equally satisfied and dissatisfied)	Mostly dissatis-fied	Unhappy	Terrible