MICHAEL R. HAGERTY, ROBERT A. CUMMINS, ABBOTT L. FERRISS, KENNETH LAND, ALEX C. MICHALOS, MARK PETERSON, ANDREW SHARPE, JOSEPH SIRGY and JOACHIM VOGEL

QUALITY OF LIFE INDEXES FOR NATIONAL POLICY: REVIEW AND AGENDA FOR RESEARCH ⋆

(Accepted 29 January, 2001)

ABSTRACT. A number of governments and public policy institutes have developed "Quality of Life Indexes" – statistics that attempt to measure the quality of life for entire states or regions. We develop 14 criteria for determining the validity and usefulness of such QOL indexes to public policy. We then review 22 of the most-used QOL indexes from around the world. We conclude that many of the indexes are successful in that they are reliable, have established time series measures, and can be disaggregated to study subpopulations. However, many fall short in four areas: (1) indexes vary greatly in their coverage and definitions of domains of QOL, (2) none of the indexes distinguish among the concepts of *input, throughput, and output* that are used by public policy analysts, (3) they fail to show how QOL outputs are sensitive to public policy inputs, and (4) none have examined convergent validity against each other. We conclude that many of these indexes are potentially very useful for public policy and recommend research to further improve them.

Writers since Plato have speculated on the "good life" and how public policy can help to nurture it. Only recently have we had the resources and the science to begin measuring the "good life" and how it arises. The last 30 years have seen a great many attempts to measure Quality of Life (QOL) in many parts of the world (Ferriss, 2000). Various indexes of QOL have been proposed by public policy institutes, government agencies, and news media. However, the advantages and liabilities of each have not been systematically evaluated.

⋆ This paper is a report of the Committee for Societal QOL Indexes, ISQOLS. Thanks are due to the officers and members of ISQOLS, who provided encouragement and advice, and to the authors of the reviewed indexes, who corresponded with the committee to improve the report. The report was endorsed by a majority of the committee. However, not all authors agree with all views expressed in the report.

Social Indicators Research **55:** 1–96, 2001.

The purpose of this report is to evaluate 22 QOL indexes that have been proposed over the last 30 years, to summarize how well they measure QOL for public policy purposes, and to propose research to further improve their usefulness. Our focus on public policy distinguishes our report from previous reviews (e.g., Diener, 1994; Evans, 1994).

Section 1 establishes the criteria that we propose to evaluate the QOL indexes. Section 2 describes the indexes and assesses each on the criteria. Section 3 summarizes "best practices" from the 22 QOL indexes and describes the agenda for research to further improve QOL indexes for public policy.

1. CRITERIA FOR EVALUATING QOL INDEXES

The committee used a Delphi technique to jointly develop criteria that are important in QOL indexes for use in public policy. Three rounds of proposals, comments, and drafts were conducted, after which 14 criteria were adopted. The criteria are summarized in the first column of Table I. We next discuss and give reasons for each.

1. *The index must have a clear practical purpose, i.e., a public policy purpose*. This distinguishes our review from other excellent reviews such as Diener (1994) whose focus is on measurement and theory, but not on usefulness for public policy. Land's (2000) recent review of social indicators also emphasizes that QOL measures should aid public policy.

2. *The index should help public policymakers develop and assess programs at all levels of aggregation*. This begins at the individual level of aggregation (e.g., physicians and counselors helping individuals in need) and continues to the family or household level (e.g., social workers helping families in need), community level (e.g., town governments developing policies and programs that can enhance community QOL), state (or province) level (e.g., state bodies developing policies and programs that can assist residents of the entire state or province), the country level (e.g., national agencies developing policies and programs that can assist citizens of that country), and the international level (e.g., international agencies

TABLE I

Summary evaluation of 22 QOL indexes on 14 criteria

	1. Health-Related QOL (HRQOL)	2. WHOQOL	3. Consumer Confidence Indexes	4. Money's "Best Places"	5. Index of Economic Well Being (IEWB)	6. Genuine Progress Index (GPI)	7. American Demographics Index	8. Johnson's Quality of Life Index	9. Eurobarometer	10. Veenhoven's Healthy Life Expectancy	11. International Living	12. U.N. Human Development Index	13. Index of Social Health	14. Annual QOL in Virginia Survey	15. Estes' ISP Index	16. Diener's Basic and Advanced Index	17. Cummins' COMQOL	18. North American Social Report	19. Philippines' Weather Station	20. Netherlands LCI	21. German Social Indicators System	22. Swedish ULF	Average Score for Criterion
1. The index must have a clear practical purpose, i.e., a public policy purpose.	2	3	3	1	3	2	2	3	3	3	2	2	2	3	3	3	2	3	3	3	3	3	2.6
2. The index should help public policymakers develop and assess programs at all levels of aggregation.	2	1	2	1	2	2	2	1	3	2	1	1	3	2	3	2	2	2	3	3	3	3	2.1
3. The index should be based on time-series to allow periodic monitoring and control.	2	1	2	2	3	3	3	3	3	3	2	2	3	2	3	1	3	2	3	3	3	3	2.5
4. The components of the index should be reliable, valid and sensitive.	3	1	2	1	2	1	3	1	3	2	1	1	2	1	3	2	2	2	2	3	3	2	2
5. The index should be grounded in well-established theory.	3	1	2	1	3	2	2	1	1	2	2	1	1	1	2	3	2	2	2	1	2	2	1.8

TABLE I

Continued

	1. Health-Related QOL (HRQOL)	2. WHOQOL	3. Consumer Confidence Indexes	4. Money's "Best Places"	5. Index of Economic Well Being (IEWB)	6. Genuine Progress Index (GPI)	7. American Demographics Index	8. Johnson's Quality of Life Index	9. Eurobarometer	10. Veenhoven's Healthy Life Expectancy	11. International Living	12. U. N. Human Development Index	13. Index of Social Health	14. Annual QOL in Virginia Survey	15. Estes' ISP Index	16. Diener's Basic and Advanced Index	17. Cummins' COMQOL	18. North American Social Report	19. Philippines' Weather Station	20. Netherlands LCI	21. German Social Indicators System	22. Swedish ULF	Average Score for Criterion
6. Index should be reported as a single number, but should be able to break it down into components.	2	2	2	1	3	3	3	2	1	2	2	2	3	1	3	3	3	2	1	3	2	2	2.2
7. The domains in total must encompass the totality of life experience.	1	1	1	1	1	1	3	1	1	2	2	1	1	2	1	1	2	2	1	2	2	2	1.5
8. 8. Each domain must encompass a substantial but discrete portion of the Qol construct.	2	2	2	2	3	2	1	1	1	2	2	2	1	2	1	2	2	2	2	2	2	2	1.8
9. 9. Each domain must be able to be measured in both objective and Subjective dimensions.	2	1	1	2	1	1	2	1	1	3	1	1	1	2	1	1	3	1	1	1	1	1	1.4
10. 10. Each domain within a generic qol instrument must have relevance for all people.	2	2	2	2	3	3	3	2	2	3	2	2	2	3	2	2	2	3	2	2	2	2	2.3

TABLE I

Continued

	1. Health-Related QOL (HRQOL)	2. WHOQOL	3. Consumer Confidence Indexes	4. Money's "Best Places"	5. Index of Economic Well Being (IEWB)	6. Genuine Progress Index (GPI)	7. American Demographics Index	8. Johnson's Quality of Life Index	9. Eurobarometer	10. Veenhoven's Healthy Life Expectancy	11. International Living	12. U. N. Human Development Index	13. Index of Social Health	14. Annual QOL in Virginia Survey	15. Estes' ISP Index	16. Diener's Basic and Advanced Index	17. Cummins' COMQOL	18. North American Social Report	19. Philippines' Weather Station	20. Netherlands LCI	21. German Social Indicators System	22. Swedish ULF	Average Score for Criterion
11. Contributes unique variance to the qol construct beyond the generic domains for the target group.	N/A	N/A	2	N/A	N/A	N/A	N/A	N/A	1	N/A	2	N/A	2	N/A	1	N/A	N/A	2	N/A	2	2	2	1.8
12. Domains must be potentially neutral, positive or negative in contribution to the qol construct.	2	1	2	N/A	N/A	N/A	N/A	2	2	3	2	2	1	2	2	1	3	2	N/A	2	2	2	1.9
13. Domains differ from the dimensions of personality	2	2	2	N/A	N/A	N/A	N/A	2	2	3	2	N/A	2	3	2	2	3	2	N/A	2	2	2	2.2
14. The subjective dimension of each domain has both a cognitive and an affective component. They are measured by questions concerning 'satisfaction'	2	1	2	N/A	N/A	N/A	1	1	2	3	1	N/A	1	1	1	N/A	3	1	N/A	1	1	1	1.4
Total Score	2.1	1.5	1.93	1.4	2.4	2	2.3	1.6	1.9	2.5	1.7	1.5	1.8	1.9	2	1.9	2.5	2	2	2.1	2.1	2.1	2

developing policies and programs that can assist the world's citizens and the planet at large).

3. *The index should be based on time series to allow periodic monitoring and control*. Land (2000) points out that this is crucial for public policy in order to assess whether conditions are improving for targeted populations and to forecast future conditions.

4. *The index should be grounded in well-established theory*. By "theory" we mean the "nomological net" of concepts and causal paths that specify how QOL is related to exogenous and endogenous variables. By "well-established", we mean that its parts have been subjected to empirical test. In particular for public policy applications, the paths and mediating variables by which policy variables will affect different domains of QOL must be specified so that policy-makers can predict the effects of new programs.

5. *The components of the index should be reliable, valid, and sensitive*. As in any measurement system, components must be shown to be reliable and valid. By "sensitive" we mean that the index should show changes in response to public policy inputs. Some items studied have been carefully assessed for reliability, validity, and sensitivity for decades (e.g., surveys of income by U.S. Census Bureau, see Moffitt, 1998; surveys of happiness by Veenhoven, 1993), while others have still to be demonstrated.

6. *The index should be reported as a single number, but can be broken down into components*. The authors disagreed on including this "single number" criterion, with a majority wishing to include it. The majority's argument is that a single number would allow citizens and policymakers to determine when QOL is improving. They suggest a method similar to the "index of leading economic indicators", which is most commonly reported as a single number, but is comprised of 12 components to assess status in each component to assist policy intervention. The minority argued that the philosophical problems in combining disparate domains of life into QOL are considerable, that weights for combining components vary greatly among people, and that most public policy interventions can be achieved merely by tracking the components of QOL (e.g.,

Vogel, 1998). Michalos (1997) has highlighted some of the difficulties in constructing single comprehensive indexes. These issues are discussed in more depth in Section 3.

7. *The domains in aggregate must encompass the totality of life experience*. QOL is a term that implies the quality of a person's whole life, not just some component part. It therefore follows that if QOL is to be segmented into its component domains, those domains in aggregate must represent the total construct.

How can one test the notion that the "domains in aggregate must encompass the totality of life experience"? One simple method is to develop composite scores made up of the aggregate QOL domains and regress global measures of QOL on these scores. If the domains in the aggregate truly encompass the totality of life experience, then we would expect a strong and significant regression coefficient. That is, composite QOL scores should not only be highly predictive of global QOL scores, but also account for a useful portion of the variance. This method can be applied with the caution that the global measure must itself be valid. See Michalos (1991: vol. 1, p. 19) for a detailed discussion.

8. *Each domain must encompass a substantial but discrete portion of the QOL construct*. The number of possible domains is infinite if one regards each aspect of life as a putative domain, so parsimony is essential in order to define a small number of domains that fulfill the requirements of 7 above. This can be tested through calculations of shared variance between domains or, most appropriately, by the amount of unique variance contributed to the aggregate QOL score by each domain.

There are no absolute rules for making such judgments, but redundancy can certainly be inferred by domain inter-correlations exceeding 0.9. Moreover, the use of several domain titles that refer to components of a common single domain, such as "health", are likely to indicate redundancy.

9. *Each domain must have the potential to be measured in both objective and subjective dimensions*. Another fundamental agreement within the QOL literature is that the construct should be measured in both objective and subjective dimensions. Therefore, in

order to retain comparability between these two dimensions at the level of domains, it is necessary that each domain have the potential to be measured both objectively and subjectively.

Why is this important? Many QOL researchers have argued and demonstrated that very often objective indicators of QOL do not correlate highly with their subjective counterparts (Cummins, 1999). Furthermore, one can easily argue that subjective well-being is a necessary, but not a sufficient, condition to capture the totality of life experience. A person may report a high level of subjective well-being, despite environmental conditions bad enough to significantly shorten life expectancy, hence affecting immediate future QOL. Similarly, objective QOL conditions (e.g., health, material possessions) of a person may have very little to do with subjective well-being. For example, a person may be wealthy, yet feel very dissatisfied with life, perhaps because of a comparison with others who may have more material possessions. Hence, the argument is that both subjective and objective indicators are necessary conditions, but neither is sufficient to encompass the totality of life experiences. Lane (1991) has argued that if we take subjective indicators of QOL at face value, then we risk accepting the "wantlessness of the poor and the acquiescence of the exploited .... We also risk accepting the inauthentic self-reports that, although reported to be infrequent, are unacceptable: The housewives who do not like their status, but, because they think they should, report themselves as 'pleased' with their lives; the abused subordinate who has learned to fear expressions of dissatisfaction ..." (p. 440). Take, for example, the work of Erik Allardt on level of living and quality of life (Allardt, 1978). The level of living concept refers to material and impersonal resources that an individual has in his or her command to maintain and/or improve the quality of that life. QOL, according to Allardt, is satisfaction of social needs, e.g., need for love, self-actualization, among others. QOL can and should be assessed through subjective indicators, but the level of living has to be assessed using objective indicators. Thus, both subjective and objective indicators are needed to capture the totality of the means and ends of QOL. D'Iribarne (1974) uses the following example to illustrate the problem of using subjective indicators without understanding the objective circumstances of the individual. A person may report being cold because

of expensiveness of fuel, while another person may report being cold because they have chosen to go skiing. The former may be an indicator of low QOL, but not the latter.

Observing these differences, many QOL scholars point to the need for both objective and subjective indicators (e.g., Cummins, 1996, 1997a; Cummins et al., 1994; Firat and Karafakioglu, 1990; Samli, 1995). Typically, measuring QOL overall or within a specific life domain (at any level of analysis) has been done through either subjective or objective indicators. Subjective indicators are mostly based on psychological responses, such as life satisfaction, job satisfaction, and personal happiness, among others. Objective indicators are measures based on frequency or physical quantity. Examples include standard of living, physical health status, and personal income, among others. Michalos (1980, Ch. 1) gives more detailed analysis of subjective versus objective indicators.

10. *Each domain within a generic QOL instrument must have relevance for most people*. Some QOL instruments are designed for use with sections of the population who have special concerns, such as medical patients. These instruments then tend to include items, such as nausea, which have relevance to the target group, but not to the general population. Such instruments cannot, therefore, be used as generic measures of QOL.

This is an important issue because the validity of these measures may be in question. QOL measures designed with a specific target population in mind, in a specific social context, may not capture the totality of life experience for other populations in different contexts and settings. Hence, the validity of a generic measure of QOL has to be demonstrated across a variety of populations in different contexts.

11. *If a specific domain is proposed for a non-generic instrument, it must be demonstrated to contribute unique variance to the QOL construct beyond the generic domains for the target group*. This requirement makes the distinction between instruments that are designed to be diagnostic and generic QOL instruments. Many of the instruments that purport to measure QOL for specific groups use several items relating to some common theme, such as the number and type of relationships. While such instruments provide detailed

information concerning the domain of "relationships", each item is unlikely to contribute significant variance to QOL if the variance from a relevant generic item (e.g., how satisfied are you with your family and friends?) was extracted first.

12. *Domains must be potentially neutral, positive, or negative in their contribution to the QOL construct*. This criterion eliminates domains that only have the capacity to elicit reduced QOL. For example, domains concerning functional status or pain are to be excluded. Our reasoning is that the optimal functioning of such domains, such as full functional status or absence of pain, can only have a neutral effect on QOL, never a positive influence.

This is an important issue because QOL measures are designed to capture the totality of life experiences, both positive and negative. Many clinical psychologists, for example, deal with ill-being. However, it is important to note that ill-being is not low well-being, and the absence of ill-being is not high well-being. For example, Argyle (1996) has argued that subjective well-being is determined by three factors: (1) Happiness, (2) life satisfaction, and (3) absence of ill-being. He argued that subjective well-being cannot be experienced when people experience ill-being in the form of depression or anxiety.

13. *Domains differ from the dimensions of personality (e.g., extraversion, self-esteem), cognitive processes (e.g., cognitive dissonance), and affect (e.g., joy) in that they cannot be measured objectively*. This criterion is related self-evidentially to Criterion 9. It also acts to separate the cognitive and affective processes that lead to subjective QOL from QOL as an outcome variable. We believe that QOL is an end state of being (although knowledge and evaluations of that state will frequently be instruments or means to action). Hence, measures of QOL have to focus on this end state, not factors that may affect it. Therefore, any measure of QOL has to capture the end state of being in various life domains. Examples include the quality of work life, quality of leisure life, quality of family life, quality of community life, quality of home life, and so on (Sirgy et al., 1982). See Michalos (1995) for more on instrumental versus intrinsic value.

14. *The subjective dimension of each domain has both a cognitive and an affective component. They are measured by questions concerning "satisfaction"*. It is widely considered that the perception of QOL is a result of multiple comparative processes (e.g., Michalos, 1985). These processes compute the "gap" or discrepancy between one's perceived current circumstance and imagined other circumstances that may refer to other people, the past, etc. The response to questions in terms of satisfaction are considered the most parsimonious measure of such discrepancies and therefore of QOL (see Cummins, 1997b).

There is a great deal of research in QOL that makes a distinction between happiness and life satisfaction (e.g., Andrews and McKennell, 1980; Brief and Roberson, 1989; Campbell, 1976; Crooker and Near, 1995; McKennell, 1978; McKennell and Andrews, 1980; Organ and Near, 1985). Happiness is recognized as more of an affective state, whereas life satisfaction represents more of a cognitive state. Hence, a good QOL measure has to capture both cognitive and affective components to truly capture the totality of life experiences.

We next apply these 14 criteria to the 22 indexes of QOL to evaluate their usefulness for public policy.

2. QOL INDEX DESCRIPTIONS AND EVALUATIONS

The committee collected QOL indexes with any of three criteria: They received attention from researchers in the field, they received attention from the press, or they had public policy applications. Committee members reviewed research and the general press and invited nominations from the members of ISQOLS. The final list contained 22 different QOL indexes. A few indexes were not reviewed because they were not applicable to public policy (Spiritual Well-Being by Bufford et al., 1991) or were published too late to review (Calvert-Henderson QOL by Henderson, et al., 2000). We will review the indexes in the order listed in Table I. The ordering reflects the number of domains encompassed, with indexes focused on one or two domains near the beginning. (We review some of these indexes because they may suggest best practices for measuring those particular domains, even though they do not profess to

measure overall QOL or to compute an overall index.) Indexes that focus on many domains are listed near the end of Table I. Thus, the CDC's Health-Related QOL index focuses on only one domain, while Estes' ISP index focuses on a broad range of domains.

The body of Table I summarizes each index on the 14 criteria using a three-point scale, where 3 = "Excellent", 2 = "Satisfactory", and 1 = "Not Satisfactory".

2.1. *CDC's Health-Related Quality of Life*

The HRQOL was developed by the Center for Disease Control, Health Care and Aging Branch (Hennessy et al., 1994). It is the first American index with a subjective component that has fielded nationally representative surveys. Over 600,000 people have been surveyed using this four-item scale beginning in January 1993, with some states using an expanded scale of 14 items. The surveys will continue as part of the Behavioral Risk Factor Surveillance System (BRFSS), and are intended to supplement more traditional measures of mortality and morbidity. It is important to note that most health-related organizations in the U.S. government have now endorsed the need for subjective as well as objective measures of health. For example, the Department of Health and Human Services' "Healthy People 2010 Objectives" for national health are composed of ten objectives, half of which are subjective self-reports.

The items in the core HRQOL are:

1. Would you say that in general your health is: Excellent, Very good, Good, Fair, or Poor?
2. Now thinking about your physical health, which includes physical illness and injury, for how many days during the past 30 days was your physical health not good?
3. Now thinking about your mental health, which includes stress, depression, and problems with emotions, for how many days during the past 30 days was your mental health not good?
4. During the past 30 days, for about how many days did poor physical or mental health keep you from doing your usual activities, such as self-care, work, or recreation?

The "healthy days" index is computed as the sum of items 2 and 3 subtracted from 30 days. From 1993 to 1996 this index averaged

24.7 days, with less educated respondents reporting fewer healthy days per month.

Evaluation: We evaluate the index using the criteria from Table I. The index is quite satisfactory on criterion 1, that it have a clear practical purpose for public policy. The Department of Health and Human Services' "Healthy People 2010 Objectives" includes objectives on healthy days. The index also helps public policy makers develop and assess programs at all levels of aggregation (#2), because the index can be computed down to the individual level, and indexes for subpopulations can be easily computed. The index is also satisfactorily based on time-series to allow periodic monitoring and control (#3), though the time-series is currently short because it was initiated in 1993. The items in the HRQOL were rated as "excellent" on #4, that the components of the index should be reliable, valid, and sensitive. A state-wide study in Missouri found strong correlations with longer scales and validates well with known groups of respondents with specific impairments (Andresen, 1999). The CDC is sponsoring a larger validation study that relates "objective" measures (such as suicide, alcoholism, divorce, and unemployment) to this subjective scale. The index is rated as excellent on being grounded in established theory (#5), with a large literature on medical quality of life and on specific disabilities resulting from various diagnoses. The index is satisfactory on #6, because it can be reported as a single number but can be broken down into the components specified above. However, the domains in total do not encompass the totality of life experience (#7) because the index is restricted to the health domain. This criterion is the only one rated as unsatisfactory. The index is rated as satisfactory on the next three criteria, because each domain (in this case, the health domain) encompasses a substantial portion of the QOL construct (#8), the domain must be measurable in both objective and subjective dimensions (#9), and the domain has relevance for all people (#10). The index was rated as "not applicable" for #11, because the index does not propose to add new domains to QOL. Lastly, the domains were rated as satisfactory on the last criteria, because the domains are potential neutral, positive or negative in contribution to the QOL construct (#12), the domains differ from dimensions of personality, which change little over time (#13), and the subjective

dimension has both a cognitive and an affective component (#14).

Overall, the HRQOL index is quite satisfactory as a measure of the health domain in QOL, and has undergone considerable reliability and validity testing. Unfortunately, it is restricted to only this domain. However, it is a good candidate for inclusion into a larger instrument that measures all domains of QOL.

2.2. *WHOQOL*

During the period 1991/1992, a series of meetings in Geneva set the operational parameters for the development of a new QOL instrument under the auspices of the World Health Organization (WHO).

They adopted a definition of QOL which states:

> Quality of life is defined as an individual's perception of their position in life in the context of the culture and value systems in which they live and in relation to their goals, expectations, standards and concerns. It is a broad-ranging concept incorporating in a complex way the person's physical health, psychological state, level of independence, social relationships, personal beliefs and their relationship to salient features of the environment.
>
> This definition reflects the view that QOL refers to a subjective evaluation which is embedded in a cultural, social, and environmental context. As such, QOL cannot be simply equated with the terms 'health status', 'life-style', 'life satisfaction', 'mental state', or 'well-being'. Rather, it is a multidimensional concept incorporating the individual's perception of these and other aspects of life.
>
> The instrument will be organized into six broad domains of QOL. These are: (1) Physical, (2) psychological, (3) level of independence, (4) social relationships, (5) environment, and (6) spiritual (The WHOQOL Group, 1993: p. 1).

The following observations can be made:

1. Despite the collection of voluminous qualitative data (see below), the group had apparently made a prior decision to adopt six domains. No justification or rationale is provided for this choice. In addition, no explanation is provided for a domain named "psychological" when the whole instrument is intended to be based on "subjective evaluations". Moreover, no rationale is provided for the omission of domains based on material well-being or productivity/meaningful activity/employment, both of which are a feature of most prior QOL scales, and how a "phys-

ical domain" is consistent with subjective evaluations is also not stated.

2. The definition seems to make it clear that QOL is to be evaluated at an individual level, but subsequent text indicates this is not necessarily so. In a discussion of the different scale versions required to cope with different cultures, they indicate that all versions will contain "core questions" such that "When weighted core questions, respondents' importance weightings of individual facets, and normative data are used, comparability of scores across different centers will be possible" (p. 2). In other words, group-based data are to be used to weight individual-level data. However, in the actual formulation of the WHOQOL-100 and WHOQOL-BREF (The WHOQOL Group, 1995a, 1996a), no importance weightings are employed, so this idea has either been dropped or put on hold for a future edition.
3. Despite the fact that a coherent body of QOL research had been forming for at least 15 years prior to this project (Andrews and Withey, 1976; Campbell et al., 1976) and dozens of QOL instruments had already been developed (see Cummins, 1998b – Directory of QOL Instruments), there is no indication that the group had drawn systematically on such data to formulate its views.
4. The group, with its background in health rather than in QOL, clearly had a deficit view of the construct. For example, "... the instrument is not expected to provide a means of measuring in any detailed fashion symptoms, disease or conditions per se, but rather the effects of disease and health interventions on quality of life" (The WHOQOL Group, 1993: p. 3). This statement indicates the view of a "health-driven" QOL, such that full health equates to maximum QOL, and QOL can be considered to be discounted through ill-health. This, of course, is consistent with an instrument intended to measure "perceived health", but certainly not an instrument intended to measure the broader QOL construct. As just one example of this difference, Broe et al. (1998) surveyed elderly people (81 ± 4) living in the community who suffered from a diverse range of health-related impairments and diseases characteristic of persons in this age group. However, their subjective QOL as measured through the

Life Satisfaction Index A (Neugarten et al., 1961) was 67.5% of the scale maximum (%SM), only slightly below the normative 70–80%SM established by Cummins (1995, 1998a). So the precise nature of the measure intended by The WHOQOL Group is unclear.

5. Even though the definition makes no reference to objective measurement, the study protocol states (p. 3), "The 'perceived objective' and 'self-report subjective' questions for each facet [sub-domain] will be written in general terms and are global self- assessments of a person's quality of life." It is then stated that the "validity of the distinction" between such items was determined by agreement by independent raters. But neither the instructions provided to these raters nor the criteria for their judgment are provided. Moreover, even further confusion is introduced on page 4 where it states that the "perceived objective questions tend to be addressed by intensity and frequency." It is not explained how the perception of intensity can be considered to be objective.

This confusion, between the distinction of objective and subjective items, is demonstrated in the "example question" provided in Table I of The WHOQOL Group (1993). The "perceived objective" item is, "How well do you sleep?" and the "self-report subjective" is, "Are you satisfied with your sleep?" In fact, however, the former is a subjective evaluation of sleep quality rather than a true objective question, which would refer to the frequency of sleep disturbance, the number of times waking occurs during a night's sleep, etc.

This lack of proper objective items explains a rather curious finding (The WHOQOL Group, 1998a) that their analysis of the original WHOQOL data found correlations greater than 0.8 between their objectives and subjective items which caused them to drop the distinction for further trials. In contrast, correlations between true objective and subjective times are normally very low and insignificant (for a review see Cummins, 1999).

The process of instrument development has been described by The WHOQOL Group (1995e).

WHOQOL-100 (The WHOQOL Group, 1995a, b, c)
This version of the instrument comprises 24 facets with four questions each. In addition, there are four questions inquiring into overall QOL and general health. The six domains remain as previously described, but they contain from one (spiritual domain) to eight (environment domain) facets. In addition to the generic instrument, a separate version has been developed for each co-operating center (e.g., The WHOQOL Group, 1995d).

Each facet contains four questions and these are variously measured according to frequency (e.g., how often do you suffer (physical) pain?), intensity (e.g., how difficult is it for you to handle any pain or discomfort?), capacity (e.g., do you have enough energy for everyday life?), and evaluation (e.g., how well do you sleep?). The forms of question are not distributed evenly within the facets, and yet the responses are summed to form a total facet score. This constitutes a major methodological issue since the combination of different response modalities in the prescribed manner is likely an invalid procedure. For example, each modality requires its own scale, and it is almost certain that each scale, per se, would have its own psychometric characteristics that would be different from others. For example, questions concerning "how much" employ a unipolar scale anchored by "not at all" and "an extreme amount", whereas questions concerning "satisfied, happy, or good" employ a bipolar scale anchored by "very dissatisfied" and "very satisfied", with a mid-point of "neither satisfied nor dissatisfied". It is highly unlikely that the various forms of measurement error (see, e.g., Bardo et al., 1982) would be equivalent between such different scale formats, and the parameters of frequency and intensity are well known to exhibit different psychometric outcomes when applied to the measurement of affective states (e.g., Diener and Emmons, 1985).

WHOQOL-BREF (WHOQOL Group, 1996a, b)
This scale has been constructed as a short form of The WHOQOL-100. It contains 26 questions, comprising one item from each of the 24 facets plus one item to measure "overall quality of life" and another to measure "general health". The item chosen to represent each facet was generally the one with the highest item vs. total

facet score correlation (The WHOQOL Group, 1998b), and the six domains of the WHOQOL-100 have been reduced to four as: Physical health (incorporating level of independence), psychological (incorporating spirituality), social relationships, and environment.

Evaluation: Szabo (1996) produced the earliest description of psychometric performance. This reported, on behalf of The WHOQOL Group, that all facets of the WHOQOL-100, except the Provider facet #19, had a Cronbach alpha of 0.82 to 0.95, that all domains were able to distinguish between well and ill people, and that "all facets are fairly independent of each other" (p. 360), even though no correlational data are presented.

This latter claim, in particular, appears to have been premature, as the analysis presented in two subsequent and linked publications attest (The WHOQOL Group, 1998a, b). These commence with a principal components analysis that employed orthogonal rotation on the grounds that "there was no reason to assume that facets such as physical environment and pain and discomfort would be related to one another" (1998a: p. 1582). However, no correlations are provided to support this assertion and many, if not a majority, of the other facets would be expected to correlate quite strongly given their general orientation to "health". Moreover, the resultant analysis produced four, rather than the anticipated six, factors, indicating substantial facet co-variation between the putative domains.

This four-factor solution explained 58 percent of the variance, but an examination of factor loadings does not indicate confidence even in this reduced factor structure due to the high level of complex loadings. Specifically, using the authors' criterion for facet inclusion as a loading >0.3, Factor 1 ("health") contains eight facets and one that cross-loaded (7/8 non-complex facets), Factor 2 comprises (5/9), Factor 3 comprises (3/8), and Factor 4 comprises (2/6). It seems clear that only the first factor can be asserted with confidence.

Despite this, however, a confirmatory analysis applied to the four-factor model produced an acceptable fit after three pairs of error variances (one pair between factors) were allowed to covary. However, it is important to note that the confirmatory four-factor model and the principal components analysis do not match. For example, in the latter analysis, both "positive feelings" and "negative feelings" loaded onto single, but different, factors. In

the confirmatory diagram, however, they are represented as both contributing to the psychological domain. Again, in the principal component analysis "bodily image and appearance" loads highly and singly onto the domain "social relationships", yet in the confirmatory model it is included in the psychological domain.

Most recently Skevington et al. (1999) have reported the analysis of 144 "national" items, additional to The WHOQOL-100, thought to have special significance for specific cultural groups. The reported procedures are interesting in a number of respects. First, no mention is made of the 1998 factor analysis which found four factors (The WHOQOL Group, 1998a), and the validity of the original six domains structure is assumed. Second, working within this assumption, each new "national" item that had been identified through focus groups was allocated to "its most appropriate facet" not, it appears, on the basis of any statistical rationale, but on the basis of the authors' intuition. Third, a series of analyses then determined the extent to which the new items fitted into the original facet structure.

The resulting analysis provides further evidence that the WHOQOL-100 requires serious revision. Within the context of some national groups, new items were found that were superior facet descriptors than were the original items. While this confirmed the power of national items, it also confirms the inadequacy of the WHOQOL-100 as a standard cross-cultural instrument. Even more disturbingly, however, some new items performed better than the original items in a cross-cultural context, causing the authors to conclude that such items should be "returned to the international item pool for global reassessment as part of any revision of the core [100 items]" (p. 485).

Conclusion: The authors claim "The WHOQOL-100 presents a major advance both in the background methodology used for the development of a reliable and valid cross-cultural instrument, and in the provision of an instrument that measures a broad range of domains of quality of life" (The WHOQOL Group, 1998a: p. 1585).

This claim is doubtful for several reasons. First, while the WHOQOL Group have certainly progressed through an exhaustive and highly complex process of scale development, this exercise has been flawed in that: (a) The authors did not use their vast array of data to deduce domains, but rather adopted the six domains they

proposed at the outset; (b) the processes of data reduction to arrive at the 24 facets are opaque as far as published material is concerned, and evidence is now available to cast doubt on the adequacy of assumed facet content; and (c) three of the four domains forming the WHOQOL-100 and BREF contain considerable cross-loading facets.

Second, there are problems with the claim of a "broad range of domains". Even assuming that the scales competently measure the claimed domains, these comprise health, environment, psychological, and social relationships. Concerns with the range of these domains include the following: (a) Spiritual well-being, material well-being, and work/productivity, all of which feature as prominent domains within the QOL literature, are not separately represented; (b) the weakest of the four domains is social relationships, and yet there is a high level of agreement within the QOL literature that this domain is preeminent in its contribution to the life quality of most people; and (c) there is a very poor match, which is unexplained, between the factor content of the principal component and confirmatory analyses.

In summary, and in relation to the criteria depicted in Table I, the only criterion to be fully met is #1, that the index has a clear practical purpose. Four of the other criteria are met to some degree (#6, 8, 10 & 13) while the remaining nine criteria are not met to any reasonable degree. It is concluded that, at this stage of its development, the WHOQOL cannot be recommended as a scale to measure overall QOL. Whether the content of the "health" factor can be regarded as an adequate measure of "health-related QOL" remains to be established.

2.3. *Consumer Confidence Index (CCI)*

Consumer Confidence Indexes or Consumer Sentiment Indexes measure the "economic outlook" domain of QOL in subjective terms. Begun in 1952 by George Katona and most notably executed since then by the University of Michigan's Survey Research Center and the Conference Board, a five-question battery of questions comprise such indexes. These indexes ushered in the subdiscipline of Behavioral Economics. Much has been written by scholars on these indexes. Zagorski and McDonnell (1995) evaluated correlates

of the Australian CCI and found this index to be more strongly correlated with leading indicators of economic cycles, rather than current (coincident) indicators.

The index is comprised of questions addressing household and society levels. In addition, current conditions and those expected in the future are also included.

The US questions follow:

1. We are interested in how people are getting along financially these days. Would you say that you (and your family living there) are better off or worse off financially than you were a year ago?
2. Now looking ahead – do you think that a year from now you (and your family living there) will be better off financially, or worse off, or just about the same as now?
3. Now turning to business conditions in the country as a whole – do you think that during the next 12 months we'll have good times financially, or bad times or what?
4. Looking ahead, which would you say is more likely – that in the country as a whole we'll have continuous good times during the next 5 years or so, or that we will have periods of widespread unemployment or depression, or what?
5. About the big things people buy for their homes – such as furniture, a refrigerator, stove, television, and things like that. Generally speaking, do you think now is a good or a bad time for people to buy major household items?

The index is calculated in the following way:

For each of the five questions which comprise the Index, the proportion of unfavorable responses are subtracted from the proportion of favorable responses – to give the favorable balance of opinion – and then 100 is added to each balance. The resulting five figures are then averaged to form the Index of Consumer Sentiment at the aggregate level.

Evaluation: Sensibly used, such indexes as the CCI can provide a measurement of future economic expectations for a society. In an economic sense, these indicators gauge societal hope, and as such are one indicator of subjective QOL. Such consumer confidence indexes are useful social indicators, because they reflect subjective opinions about material living conditions. In addition, consumer

confidence indexes serve as predictors of other attitudes, such as those regarding political party preferences in national voting.

Because the CCI has a clear practical purpose for policy makers (Criterion 1), it receives the highest rating in our study on this dimension. As a representation of consumer sentiment, the CCI has had a wide variety of uses in government from setting monetary policy to forecasting fiscal revenues. However, its practicality results from its tight focus on subjective assessments of material elements of QOL As a result, the CCI receives the lowest rating for having domains which encompass the totality of life experience (#7), and for being able to be measured in both objective and subjective terms (#9).

On the remaining criteria in our study, the CCI receives a rating of "2". The primary reason for such ratings is the narrow focus of the CCI. In addition, the CCI was developed before QOL research, as we now know it, began. While improving the CCI for QOL research is not out of the question, the costs/benefits trade-off is likely to be small. The CCI could be applied directly as a measurement of the domain of expected economic conditions within studies of subjective QOL. For this purpose, the CCI would well used.

2.4. *Money's "Best Places"*

With subscription and individual sales each month of almost two million copies, *Money* Magazine could be said to be the most prolific distributor of QOL information today with its annual Best Places survey. Sadly, *Money's* editors give only a vague description of their proprietary methods for ranking 300 metropolitan areas in the US in their work. More sadly, several scholarly researchers have cited this index as an exemplar of unreliable and volatile measurement (Kotler et al., 1993; Guterbock, 1997).

Money uses a three-step process in developing its rankings each year (Guterbock, 1997). In the first stage, 250 *Money* readers are surveyed to determine the importance weights of more than 40 criteria used in choosing a city to live. In the second stage, current statistical data for each city is collected on a wide range of empirical indicators. While the full list of indicators is not disclosed, some examples offered by *Money* include the following: (1) number of doctors per capita, (2) violent crime rate from the FBI Uniform

Crime Reports, (3) the cost-of-living index from the American Chamber of Commerce Research Association, (4) recent job growth, (5) future job growth estimates, (6) typical price of a three-bedroom home and its property tax from Century 21 Real Estate brokers, and (7) housing appreciation rate over the past 12 months from Century 21. In the third stage, the individual indicators are aggregated into nine broad categories matching categories previously derived in the first stage.

For its 1999 Best Places survey, *Money* Magazine used demographic and U.S. Census data compiled by Fast Forward, a consulting firm in Portland, Oregon, to see how the 300 biggest U.S. metropolitan areas stack up in 46 different areas (Gertner and Kirwan, 1999). According to the staff writers of *Money*, the Best Places survey focused on the things that matter most to *Money's* readers: Economy, education, culture, recreation crime, and pollution. No explanation of a rating scheme within these categories or a description of a weighting system for the categories in a final computation of the Best Places was offered.

Evaluation: *Money* Magazine's "Best Places" is unsatisfactory as a QOL index. Its total score of 1.4 is the worst of the 22 QOL indexes. In Table I, *Money's* index receives the lowest rating of "1" in our study for its orientation to policy making (Criteria 1 and 2), its reliability and validity (#4), its theoretical grounding (#5), its reporting as a single number (#6), and its encompassment of the totality of life experience (#7). *Money's* index receives a rating of "2" for its repeated measurement each year (#3), its encompassing of a substantial portion of the QOL construct (#8), its measurement in both objective and subjective dimensions (#9), and its relevance for all people (#10). The *Money* index does not receive the highest rating of "3" in our study on any of the fourteen criteria.

Guterbock (1997) does a masterful job of "retro-engineering" the skimpy data provided by *Money* over ten years and succeeds in deducing the flawed weighting scheme for the variables used. Aside from being atheoretical, the problem with the index appears to be the overweighting given to the economic conditions of the 300 cities in the U.S. which are ranked in this index (Guterbock, 1997). While not acknowledged by *Money's* editorial staff or its publisher, Time-Warner, the index would be better labeled "*Money's* Finan-

cially Hot Places". This overweighting could be easily corrected. However, it appears that the index's volatility contributes to the "news" aspect of each year's new presentation of "best places". In this way, the index generates excitement similar to the announcement of each year's Academy Award nominees. With such glamour, it appears unlikely that Time-Warner (publishers) would correct the index for the purpose of being more truthful, because the results would likely be more stable (predictable and less newsy and therefore more boring). This is about fun (entertainment), not truth (science).

With this likely permanent impairment in the development of *Money's* Best Places identified, scientists could use the better methodological aspects of this index (i.e., yearly survey to assess the importance of variables used in the study) and then develop the index in a rigorous way. For example, the sample size of the survey itself needs to be boosted from 250 to 1000. Subsequently, the weighting of the variables could be corrected based on appropriate theory. The good news here is that a grant opportunity appears to be emerging because policymakers (mayors and city councilmen) do not ignore *Money's* Best Places index. Sponsoring an index decoupled from the demand of entertainment would be a public service likely to be compelling for certain funding sources.

2.5. *Index of Economic Well-Being (IEWB)*

The IEWB was developed by Lars Osberg of Dalhousie University in Halifax, Nova Scotia and Andrew Sharpe of the Ottawa-based Centre for the Study of Living Standards (see Osberg and Sharpe, 1998, 1999, 2000) and is posted at www.csls.ca. The Index is based on the view that the economic well-being of a society depends on the level of average consumption flows, aggregate accumulation of productive stocks, inequality in the distribution of individual incomes and insecurity in the anticipation of future incomes. The weights attached to each of these component of economic well-being will vary, depending on the values of different observers. It is argued that public debate would be improved if there is explicit consideration of the aspects of economic well-being obscured by average income trends and if the weights attached to these aspects were explicitly open for discussion.

The four components or dimensions of economic well-being on the index are:

- effective per capita consumption flows;
- net societal accumulation of stocks of productive resources;
- poverty and inequality; and
- economic security from job loss and unemployment, illness, family breakup, and poverty in old age;

Consumption flows encompasses marketed personal consumption flows, adjusted for the underground economy, the value of increased longevity, changes in family size which affect the economies of scale in household consumption, and regrettables or defensive intermediate consumption goods (cost of commuting, household pollution abatement, auto accidents, and crime); changes in working time; government services; and the value of unpaid work.

Stocks of wealth include the net capital physical stock, including housing stocks; the stock of research and development; value of natural resources stocks; the stock of human capital; the level of foreign indebtedness; and the net changes in the value of the environment due to CO_2 emissions.

The inequality component of the index consists of income inequality, defined as the Gini coefficient for after-tax household income and the intensity of poverty (incidence and depth), defined as the product of the poverty rate and the poverty gap, that is the difference between the average income of those in poverty and poverty line divided by the poverty line. The poverty line is defined as one half median adjusted household income.

The insecurity component of the index is based on the change over time in the economic risks associated with unemployment, illness, "widowhood" (or single female parenthood) and old age. The risk of unemployment is determined by the employment/population ratio, the employment insurance coverage of the unemployed, and the benefits ratio. The risk of illness is modelled as the percentage of disposable income devoted to health costs. The risk of single parent poverty is determined by the divorce rate and poverty intensity of single parent families. The risk of poverty in old age is a function of the poverty intensity of the elderly population.

Trends in the index are determined by the choice of variables that are included in the index, the trends in those variables and the

weights given these variables. Since the four main dimensions of economic well-being are separately identified, it is easy to conduct sensitivity analyses of the impact on perceived overall trends of different weighting of these dimensions. For discussion purposes, consumption flows have been given a weight of 0.4, wealth stocks a weight of 0.1, and inequality and economic insecurity have each been given weights of 0.25.

The sub-components of the consumption flows and wealth stocks are expressed in constant dollars on a per capita basis. There consequently is no need for explicit weighting as these dollar values represent implicit weights. In terms of the inequality/poverty subcomponents, a Rawlsian perspective assigns greater importance to poverty than to overall inequality trends, and a weight of 0.1877 has therefore been given poverty intensity and 0.0625 to the Gini coefficient. In other words, poverty is given three times the weight of inequality. The subcomponents of the economic security index are weighted by the relative importance of the specific population at risk in the total population.

Evaluation: The public policy purpose of the IEWB is to ascertain trends in overall economic well-being and the constituents of well-being, allowing policymakers to identify where problems exist and to take corrective action (#1). It has been estimated at the province, national, and international level, so it can help policymakers at these levels in program and policy development (#2). However it is difficult to disaggregate it to special populations, such as the elderly or immigrants. The IEWB has been estimated for Canada for 1971–1999 period, for the United States for the 1962–1999 period, and for Australia, Germany, Norway, Sweden, and the United Kingdom for the 1980–1996 period (#3).

The IEWB overall is reliable (#4), but it is a complex index and could be improved in a number of areas through better quality data, inclusion of additional variables (e.g. more variables on environmental degradation), and a less subjective weighting of the components. The IEWB is well grounded in economic theory (#5), grouping variables that contribute to economic well-being, using the concepts of consumption flows, stocks of wealth, inequality, and economic security. The variables themselves have sound analytical roots based on such concepts as the Sen measure of poverty intensity and the

Gini coefficient. The IEWB is reported as a single number, with the base year 100 (#6). As it is built up from components, it can easily be broken down by components.

The IEWB really only encompasses the economic domain (#7 and 8). Still, the economic domain is broadly defined to include such constructs as poverty and will eventually include crime. All variables included in the IEWB are based on objective data (#9). There are no variables on subjective perceptions of well-being (though the weights are subjective and may be changed for each individual's values). Hence, these items are rated as missing.

2.6. *Genuine Progress Index (GPI)*

The GPI is a measure of economic well-being for the United States from 1950 to the present. It broadens the conventional accounting framework to include the contributions of the family and community realms, and of the natural habitat, along with conventionally measured economic production.

The GPI takes into account more than 20 aspects of economic life that GDP ignores. It includes estimates of the economic contribution of numerous social and environmental factors, which the GDP gives an implicit and arbitrary value of zero. It also differentiates between economic transactions that add to well-being and those that which diminish it. The GPI then integrates these factors into a composite measure so that the benefits of economic activity can be weighted against the costs. The GPI is intended to provide citizens and policy makers with a more accurate barometer of the overall health of the economy, and of how the national condition is changing over time.

The GPI was developed by the San Francisco-based think tank Redefining Progress (Cobb et al., 1995). It starts with personal consumption expenditures, makes an adjustment for income distribution, and then adds or subtracts categories of spending based on whether they enhance or detract from well-being. Additions are the value of time spent on household work, parenting, and volunteer work; the value of services of consumer durables; and services of highways and streets. Subtractions are defensive expenditures due to crime, auto accidents, and pollution; social costs, such as the cost of divorce, household cost of pollution and loss of leisure; and depre-

ciation of environmental assets and natural resources, including loss of farmland, wetlands, old growth forests, reduction in the stock of natural resources, and the damaging effects of wastes and pollution. All categories are expressed in dollars so can be aggregated into one figure or bottom line, the GPI.

Evaluation: The GPI is somewhat grounded in economic theory and uses economic theory to place monetary value on many variables that are not explicitly valued by the market (e.g., ozone depletion, loss of wetlands). The GPI is reported as a dollar value like GDP. As it is built up from components, it can easily be broken down by components. The GPI has been estimated for the 1950–1997 period for the United States.

There are serious problems with the assumptions and valuation techniques used to estimate many of the resource and environmental variables in the GPI. For example, the value of the loss of wetlands becomes unrealistically larger and larger over time and gives a strong downward bias to the index. For this reason, the index in its current form is not a reliable measure of QOL or genuine progress. The public policy purpose of the GPI is to ascertain trends in overall progress and its components, allowing policymakers to identify where problems exist and take corrective action. The GPI has been estimated at the national and international level, so it could potentially help policymakers at these levels in program and policy development. It only encompasses the economic domain since it is proposed as an alternative to or replacement for GDP. The economic domain covered by the GPI covers a substantial, but discrete, portion of a QOL construct. On the other hand, it covers the complete population, not specific groups.

All variables included in the GPI are based on objective data. There are no variables on subjective perceptions of well-being. The methodology used to estimate components is, however, very subjective.

2.7. *American Demographics Index of Well-Being*

The *American Demographics* published the Index of Well-Being for the United States from February 1996 to December 1998. The Index, however, covers the period April 1990 to July 1998. It is a monthly

composite of five indicators developed, maintained, and reported by Associate Professor Elia Kacapyr, an economist of Ithaca College, Ithaca, NY. He is the author of *Economic Forecasting* (1996).

Item Selection: Professor Kacapyr selected the items upon the basis of an economist's conception of well-being, free of any paradigm or QOL theory. Discussants of the Index "often criticize the ad hoc selection of the components. I have no defense against this criticism. I tried to incorporate as many data-series as possible in my index that were available on a monthly basis" (e-mail communication 4/26/99).

Components: Eleven primary indicators relate to five sectors or components: Consumer attitudes, income and employment opportunity, social and physical environment, leisure, productivity, and technology. Each month, the composite index as well as the five components are reported and changes monthly or yearly noted. Typically, the "current" report would be for a month six months later. For example, in the October issue of *American Demographics*, indicators for May would be reported for comparing the preceding April and the May indicators for one and two earlier years. Each component is "benchmarked" to an April 1990 level of 100. Thus, it could be reported that in May 1996, when Income and Employment Opportunity stood at 102.01, that on average this component was 2.01% better than in 1990. The separate reporting of each component and the socioeconomic forces undergirding the change are an important, informative feature of the Index.

Weights: Weights for each element of each component were determined "by fitting a trend line to the series from 1983 to 1997. The larger the monthly deviations from that trend, the smaller the weight given to the data series. Specifically, the weight given to a data series is inversely proportional to the variance from its own trend. The weights are normalized so that they sum to unity" (e-mail communication 4/9/99). The author further explains, "Every component of my index gets the weight it deserves because a 10% change in consumer attitudes is equivalent to a 0.2% change in the leisure sector based on past trends. The 10% move in consumer attitudes gets a 1% weight while the 0.2% move in leisure gets a 50% weight.

TABLE II

Components and Weights for American Demographics Index

Component	Components	Weight
Consumer Attitudes (CATT)		**1%**
	Consumer Confidence Index	47%
	Consumer Expectations Index	53%
Income and Employment Opportunity (INCEMP)		**21%**
	Real disposable income per capita	39%
	Employment rate	61%
Social and Physical Environment (INEG)		**10%**
	Number of endangered species	32%
	Crime rate	43%
	Divorce rate	25%
Leisure (LSUR)		**50%**
	168 minus weekly hours worked	90%
	Real spending on recreation per capita	10%
Productivity and Technology (PROTEC)		**18%**
	Industrial production per unit of labor	69%
	Industrial production per unit of energy	31%

After applying the weights, both moves are seen to be equivalent" (ibid.).

Thus, by the above-described device, change in the Index is influenced equally by each of the five components. The indicators employed and their weights are presented in Table II.

Trend: The Index declined from 4/90 to 4/92 and since has followed an upward course. The July 1998 index reached the level of 104.11, showing that the "typical American" was 4.11% better off than in April 1990. Each component had increased over 1990 except social

and physical environment, the July 1998 readings being:

Consumer Attitudes	119.3
Income and Employment Opportunity	104.3
Social and Physical Environment	77.1
Leisure	105.9
Productivity	112.9
Index of Well-Being	104.1

Evaluation:

1. Owing to a lack of data reported monthly, the development of a monthly index would be limited in scope. Morbidity data may become available on a monthly basis eventually from the Behavioral Risk Factor Surveillance System Survey of the CDC. However, the range of available data expands considerably when an annual period is employed.
2. In evaluating QOL items, the input-throughput-output scheme may be useful. The Index of Well-Being contains items that might be classified in the three phases of the QOL process. For example, the item, "industrial production per unit of labor" is considered output in economic terms, but is considered as an input to subjective well-being. As such, it should be distinguished from items that more closely reflect output aspects of well-being, such as crime, divorce, etc. Genuine output measures are not readily available. Consideration should be given to identifying essential output measures and advocating their collection.
3. The Index of Well-Being seems to employ a weighting scheme unique among QOL studies. The purpose was to equalize the influence upon change, rather than the influence of the item upon the output of QOL. Needed is an external QOL criterion against which weights of input factors may be derived.
4. The presentation of the monthly Index was accompanied by commentary identifying changes in the individual elements of the Index. Comments also included excursions into underlying economic and social influences. Some of the latter included status of American women, changes in the physical environ-

ment, wealth vs. well-being, black market effects upon the economy, hours of work, and others. These commentaries were quite enlightening and provided an elementary economic education of ongoing economic forces. An index tracking change should include commentaries that probe the underlying causes of changes.

5. Evaluations with reference to the Index Criteria: The practical purpose of informing economic policy is **satisfactory** for this index. Similarly, it should help policy makers assess and develop programs. It is **excellent** in being based upon time series for monthly monitoring. The components of the index are recognized, widely used, indicators and in this respect are rated **excellent**. It is based upon common sense economic understandings some of which would be considered economic theory, and in this respect is evaluated **satisfactory**. However, it is **excellent** in that it is reported in a summary number and also is reported by components, as is mentioned above. The domains do not encompass the totality of life experience, but they do include the chief domains of economic activity, which is its purpose, for this reason it is rated **excellent**. While each domain measures a discreet segment of well-being, this criteria must be considered ***unsatisfactory*** since many aspects are not represented (and could not be considering the monthly requirement for the index, monthly data not being available for many items). While each domain cannot be measured both subjectively and objectively, owing to the character of the items. It is rated **satisfactory** in this respect, since questionnaire items for SWB can be written to reflect the dimensions. It is **excellent** in reflecting the entire working population. The remaining criteria for evaluation are not applicable to this index.

2.8. *Johnston's QOL Index*

This index (Johnston, 1988) is based purely on objective measures, and its construction is unusual. Instead of providing a simple index value, it provides a comparative value based on time series analysis. Describing the method of construction will make this feature clear as follows:

1. A total of 21 objective variables are presented representing nine "areas of social concern". These are health, public safety, education, employment, earnings and income, poverty, housing, family stability, and equality. Either two or three indicators represent each area. For example, the area of health comprises three items as life expectancy at birth, infant mortality rate, and days of disability per person per year.
2. Some arbitrary year is chosen to provide the initial values for each indicator. The values derived from some future time period are then expressed as percentage changes in relation to the first set of values.
3. "Multipliers" then adjust the percentage changes within a data set, which typically comprises several time periods. Each multiplier is a value within the range 1 to 100, chosen pragmatically to hold all percentage fluctuations to within the range +50 to –50. This is thus a simple linear transformation, applied to each variable, such that the transformed values can be more reasonably aggregated to form a single index.
4. The indexes for any time period can be aggregated to provide a single score.

Evaluation: The index is essentially atheoretical. No rationale is provided for the choice of "areas"; several of the "areas" would be expected to share considerable variance, such as "earnings and income" and "poverty"; and the reason some "areas" comprise two indexes while others comprise three is not stated.

The differential weighting has the unfortunate consequence that the degree of relative change observed between areas can not be simply interpreted, since some will be relatively enhanced or attenuated by the strength of their weighting.

The only psychometric datum presented in the paper is a regression of the index values against disposable income. The $R^{š}$ value of 0.45 reflects the dominance of economic indicators, or at least economically sensitive indicators, in the index. No other data are known. A literature search has failed to provide further references to the scale.

In summary, and in relation to the criteria depicted in Table I, the Index reasonably meets two criteria. These are #1 that it has a clear practical purpose, and #3, that it is based on time-series. It partially

meets four other criteria (#6, 10, 12 & 13) but fails to meet the remaining eight. In conclusion, this instrument appears to have little utility as a general QOL index. It omits any subjective measures, is heavily biased towards economic concerns, and has other problems of construction and interpretation that have been mentioned.

2.9. *Eurobarometer*

Eurobarometer (1998) public opinion surveys ("standard Eurobarometer surveys") have been conducted on behalf of the Directorate-General for Information, Communication, Culture, Audiovisual of the European Commission each spring and autumn since autumn 1973. They have included Greece since Autumn 1980, Portugal and Spain since autumn 1985, and the former German Democratic Republic from autumn 1990 onwards.

An identical set of questions is asked of representative samples of the population aged 15 years and over in each Member State of the European Union. The regular sample in standard Eurobarometer surveys is 1000 people per country except Luxembourg (500) and the United Kingdom (1000 in Great Britain and 300 in Northern Ireland). The Eurobarometer surveys are carried out by national institutes, with coordination of questionnaires through the commission. All institutes are members of the European Society for Opinion and Marketing Research and comply with its standards.

Eurobarometer thus basically is a continuing series of twice-yearly cross-sectional public opinion surveys, 1973-present – sort of an analogue of the General Social Survey in the U.S. Eurobarometer reports typically give cross-tabulations of responses to survey questions for each of the 15 Member States as well as for the total European Union community.

Two satisfaction questions are asked each year:

- On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the life you lead?
- On the whole, are you very satisfied, fairly satisfied, not very satisfied, or not at all satisfied with the way democracy works (in your country)?

In addition, numerous other opinion questions are asked in each survey, with the content varying somewhat with the current issues of the day. The general focus of many of the surveys has been on

such political issues as attitudes towards European integration, the European Union, and its Parliament.

Evaluation: The Eurobarometer series provides only a limited QOL Index. With respect to the criteria for evaluation given in Table I, the Eurobarometer satisfaction scales presumably are measured to determine the extent to which the national populations of the European Union are satisfied with their lives and with the operation of democratic governance processes in their countries. These goals have a clear public policy purpose (Criterion 1). The data also can be aggregated to various levels of analysis (#2). Sample size, however, limits the extent to which the index can be disaggregated sub-nationally. The Eurobarometer is based on time series that show meaningful variation among the Member States and with the ups and downs of the European national economies since 1973 (#3). The components of the index have the reliability, validity, and sensitivity that are typical of "overall satisfaction" sample survey questions (#4). But whatever theory the Eurobarometer surveys are built upon is opinion/marketing research theory and not QOL theory (#5). There also is little systematic attention given to the measurement of satisfaction with most specific domains of social life and thus the index is limited in the extent to which it can be broken down into components (#6). It follows that domain-specific #s 7, 8, 9, 10, 11 12, 13, and 14 largely are inapplicable.

2.10. *Veenhoven's Happy Life-Expectancy Scale (HLE)*

This scale was developed by Ruut Veenhoven (1996) and it is connected to several other works, including especially Veenhoven (1984, 1994, 1995). It is not based on a single theory, but is supported by theoretical arguments as well as empirical research. One of the most attractive features of Veenhoven's work is that he always provides his readers with up-to-date reviews of theories and hypotheses and matches them against his empirical research, usually giving theoreticians more problems. For example, he criticizes the *basic need approach* to identifying a set of indicators of the QOL on the ground that there is no clear way to empirically demonstrate what people "really" need. He criticizes *social comparison theory*

on the ground that its predictions (as he sees them) are not supported for rich and poor nations (Veenhoven and Ouwenell, 1995).

The computation of Happy Life-Expectancy consists in multiplying "standard" life-expectancy in years with average happiness as expressed on a scale ranging from zero to one. For example,

> Suppose that life-expectancy in a country is 50 years, and that the average score on a 0 to 10 step happiness scale is 5. Converted to a 0–1 scale, the happiness score is then 0.5. The product of 50 and 0.5 is 25. So happy life-expectancy in that country is 25 years. This example characterizes most of the poor nations in the present day world. If life-expectancy is 80 years and average happiness 8, happy life-expectancy is 64 years (80×0.8). This example characterizes the most livable nations in the present day world (Veenhoven, 1996: p. 29).

Evaluation: Under the title "metric quality", Veenhoven (1996: p. 44) lists the following virtues of HLE. It has "substantive validity", by which he means that it clearly measures the "degree to which people live long and happily in a country". It has discriminant validity insofar as it successfully differentiated 47 nations according to their HLE scores, and it has concurrent validity insofar as he was able to show that, "The better the living conditions in a country, the higher the happy life-expectancy". With a fairly long list of zero-order and affluence-controlled correlations between HLE and other measures (e.g., purchasing power, state expenditures in % GDP, % literate), he is able to dispel suspicions that people might have that "happiness" designates something too vague to be useful for public policy. It is unclear why he said that "HLE appeared unrelated to state welfare effort" when the correlation between HLE and state expenditures in % GDP are listed as $r = 0.57$ ($n = 34$ nations). Controlling for affluence lowers the relationship to $r = 0.15$, as one would expect. It is also unclear (in spite of his explanatory footnote) why unemployment rates would be positively associated ($r = 0.40$) with HLE.

Given the widespread use and usefulness of life-expectancy measures, any quality-of-life scale involving these measures is bound to be initially attractive. Its use in the United Nations Human Development Index, for example, is well known. Combining life-expectancy to happiness to measure "apparent" livability as input is an idea that certainly merits more work. Veenhoven is aware that life-expectancy rates change so slowly that HLE will not be useful

for monitoring the short-term impact of interventions resulting from public policy initiatives. It is also not particularly useful for identifying particular, detailed problems to be solved with public initiatives. Still, as a long-term measure, it may prove to be as useful as the Human Development Index. The idea of attaching just one more measure, happiness, to the other three items of the HDI merits serious consideration. Since it would be possible to add a happiness item to any one of the routinely collected and massive labor force surveys undertaken in most countries, it would be very easy for countries to get the required data to calculate their HLEs.

2.11. *International Living Index*

International Living is a periodical with the mission to "detail the travel, lifestyle, retirement, investment, employment, and educational opportunities in countries all around the world". Each year since 1984 it has rated over 150 countries on seven attributes and has combined these attributes into an overall index. The attributes with their weights are: economy (20%), health (12%), culture and recreation (12%), infrastructure (12%), cost of living (20%), freedom (12%), and safety and risk (12%). They state that many sources are used to rate each country on each domain, including: Economy (from GDP/capita in $PPP, GDP growth, and inflation rate) health (from calorie consumption as a percent of daily requirements, number of people per doctor, percent of people with access to safe water, infant mortality, life expectancy and public health expenditures as a percent of GDP), culture and recreation (from literacy rate, newspaper circulation, primary and secondary school enrollment ratios, acres of national parks and nature preserves), infrastructure (from highways and waterways per 100,000 inhabitants, number of airports, telephones, televisions and radios per 1,000, and number of people per motor vehicle), cost of living (from U.S. State Department's Index of Overseas Living Costs for a Western style of life in each country), freedom (from Freedom House, averaged over their political rights and civil liberties ratings), and safety and risk (from U.S. Department of State's hardship differentials and danger allowances). However, these scores can be modified by the subjective judgment of the editors, who in turn draw on their own

experience and the input of their subscribers: "We come armed with an atlas, a year's worth of newspaper clippings, notes from our contributors, and opinions formed in years of travel and overseas living" (p. 6). Hence the exact scoring procedure on each domain is subjective, though is supported by data collected by government agencies.

Evaluation: The index was rated as satisfactory on having a public policy purpose (criterion 1), but was rated as unsatisfactory in assessing programs at all levels of aggregation (#2), since only nation-wide scores are given. The index is based on time-series (#3), but received an unsatisfactory on #4. Peterson and Malhotra (1997) and Peterson et al. (1999) assess test/retest reliabilities in the form of year-to-year correlations and conclude that they are very high, averaging 0.93. However, inter-rater reliabilities are unreported and the procedure for scaling component scores is not specified and is subject to the biases and heuristics of group decision-making. Peterson et al. also factor the seven attributes, to recover a three-factor structure, which they name: benefits, costs, and sustainability. The magazine does not report predictive validity scores in predicting actual development of the country. The magazine also reports no theory for their QOL ratings (#5), though Peterson and Malhotra (1997) provide some structure. The index can be reported both as a single number and as component scores (#6) and each domain contains a discrete portion of QOL (#8). However, they do not demonstrate that the components encompass the totality of QOL (#7). The index restricts itself to "objective indicators" and does not collect perceived satisfaction from citizens (#9). The remaining criteria were rated as satisfactory.

In summary, it is important to point out that the magazine's purpose is to aid businesses and individuals (non-citizens) in immigration, corporate expansion, travel, and retirement. This differs from most QOL indexes, which evaluate the status of native citizens. However, the International Living Index is quite a useful index for this specific audience. It has the further advantage that it specifically rates almost every nation in the world. In contrast, most other indexes are currently applied to only one nation.

2.12. *U.N. Human Development Index (HDI)*

The United Nations Development Program first published a Human Development Index (HDI) in the *Human Development Report* (UNHDP, 1990). Annually thereafter, at least for four years, in response to criticisms, revisions to the methodology were made and published in the same serial publication (UNHDP, 1993).

The HDI is a combination of three indicators: Longevity, knowledge (literacy 2/3, years of schooling 1/3), and income. A maximum and minimum value is selected for each variable, and by a formula the indicators are transformed to range from zero to one, and averaged to produce the HDI. Longevity is life expectancy at birth, which is the average years of life of persons who died in the year of reference. The knowledge variable is a combination of adult literacy – the percent of adults who can read and write – and years of schooling attained by the adult population. Income originally was the log of the per capita GDP. Subsequently, the GDP/capita was modified by using an Adkinson formulation s that "the higher the income relative to the poverty level, the more sharply the diminishing returns affect the contribution of income to human development." (UNHDP, 1992: p. 91) As is noted below, some of these measures and weights were modified in response to criticisms.

The concept of human development is defined as, "A process of enlarging people's choices" (UNHDP, 1990: p. 10). The essentials for all peoples are taken to be "a long and healthy life, to acquire knowledge and to have access to resources needed for a decent standard of living" (ibid.). The definition further encompasses choices made in other areas, including political, economic, and social freedom, being creative and productive, enjoying personal self-respect, and human rights (This conception of human needs may be compared with those set forth by Diener (1995) and Schwartz (1992)).

Evaluation:

1. *Longevity*. Justification: "The importance of life expectancy lies in the common belief that a long life is valuable in itself and the fact that various indirect benefits (such as adequate nutrition and good health) are closely associated with higher life expectancy" (UNDP, 1990: p. 12). This concept has been criticized as being

of "limited value as an indicator of health, illness, and disability during life" since it is a measure of mortality (Nubler, 1995: p. 173). Many developing countries have inadequate mortality data for computation of reliable life tables; the reliability of longevity estimates for some countries, therefore, is weak (Murray, 1991). The minimum and maximum life expectancies for the 1990 calculations were 41.8 and 78.4 years. (UNDP, 1990: p. 109). The minimum/maximum for 1991 were 42.0 and 78.6 years. The 1994 calculations used minimum/maximum of 25.0 and 85.0 years. The same minimum/maximum were used in 1996. These changes undoubtedly were in response to criticisms (Luchters and Menkhoff, 1994: p. 12). The index value of a country is computed by dividing the actual life expectancy value minus the minimum value by the maximum value subtracted from the minimum value. This number is given the value 1/3 in the final summation of the index (UNHDP, 1996: p. 106).

The use of minimum and maximum values is faulted. "The minimum country will always have a value of one and the maximum country a value of zero for the deprivation in this dimension." The case is cited of a country that raises its life expectancy to increase the minimum value, with the maximum country remaining constant, the transformed values would still range the same and would not reflect the leap in longevity (Trabold-Nubler, 1991: p. 239). The solution suggested for this problem is to select minimum and maximum values that are absolute (constant) and will not be surpassed by the developing countries over the next decade or two. Zero for the minimum and 100 for the maximum was suggested (ibid. 241).

2. *Literacy*. Their measure is "... only a crude reflection of access to education .... But literacy is a person's first step in learning and knowledge building, so literacy figures are essential in any measurement of human development" (UNDR, 1990: p. 12). Literacy is weighted 2/3 and years of schooling 1/3. No justification is given for this differential weighting, and Nubler (1995: p. 173) argues that because infant mortality is negatively correlated with years of schooling, but not adult literacy, years of schooling is more critical and should be given greater weight. He also finds that since the average number of years

of schooling relates only to the population 25 years of age and older, the measure reflects an increase in the level of education only after a lag of up to 20 years (Nubler, 1995: p. 174). He also finds that, "The threshold is ... very low, and differences in the level of reading and writing ability are not taken into account." He finds that 80% of all countries are in the upper third. "The literacy rate therefore has a limited ability to differentiate" (ibid., p. 174). Minimum and maximum values are chosen from 1987. The minimum value for literacy was 12.3% and the maximum 100%, the values for Somalia and Japan, respectively. Trabold-Nubler argues that, "The raw data are simply added and not transformed into a 0–1 scale" (1991: p. 239). In the 1991 HDI literacy plus the mean years of schooling were used, weights being 2/3 and 1/3, respectively. The minimum/maximum were set at 9.1 and 70.1, respectively. The 1992 report combines literacy 2/3 plus average years of schooling 1/3, and uses a minimum/maximum figure of 0.00 and 3.00 (sic) (Attachment A for 1992). The 1994 report uses for adult literacy the minimum/maximum of 0.0 and 100.0. For mean years of schooling, the minimum/maximum are 0.0 and 15.0. The same minimum/maximum were used in 1996.

The 1996 HDI used a combination of adult literacy (2/3 weight) and combined primary, secondary, and tertiary enrollment ratios (1/3 weight). Adult literacy minimum and maximum were 0 and 100%, as also were the combined enrollment ratio minimum and maximum. In effect, the observed value is divided by 100 in each case. In 1996, however, the schooling variable was changed to the combined primary, secondary, and tertiary enrolment ratio index. One criticism of this measure is that the average years of schooling of the adult population reflects the educational product of, say, 20 years ago, rather than the present (Murray, 1991). Murray recommends the use of age-specific mean years of schooling multiplied by population in each age to produce an aggregate average years of schooling, a procedure obviously impossible for many countries with insufficient statistical bases.

3. *Income*. Justification. Income, it is said, provides "command over resources needed for a decent living" (UNDP, 1990: p. 12). Per capita income is rejected as "not very useful" because it

involves "non-tradable goods and services and the distortions from exchange rate anomalies, tariffs, and taxes." Consequently, HDI uses purchasing-power-adjusted real GDP per capita. The adjusted GDP used were developed in the International Price Comparison Project. The initial 1990 HDI employed the log of the real GDP per capita. The 1990 minimum/maximum were 2.34 and 3.68 (UNDP, 1990: p. 109). The Atkinson formula for the utility of income was employed (Trabold-Nubler, 1991: p. 240). The 1991 HDI modified the procedure for calculation, employing a poverty line and enforcing decreasing utility as income increases. The 1991 minimum/maximum for adjusted real GDP per capita were 350 and 5,070, respectively. For 1992 the maximum was increased to 5,079, and the minimum remained at 380. For the 1993 HDI the minimum/maximum were again changed to 367 and 5,075, respectively, with no explanation for the change. The method of treating income was changed again for the 1994 HDI as follows: "For income, the threshold value is taken to be the global average real GDP per capita of PPP$5,120. Multiples of income beyond the threshold are discounted using a progressively higher rate." Income minimum/maximum were PPP$220. and $40,000. Another form of the Atkinson formula was used in the 1996 HDI.

4. *Other measures*. In addition to the general measure, as outlined above, the UN Development Programme also developed a gender-related development index (GDI) and a gender empowerment measure (GEM) for countries where adequate data were available.
5. *Trends and Associations*. The HDI shows variation between developed and developing nations and among regions. The different progress of various nations in change in HDI is illustrated by these trends for particular states. The trend from 1960 to about 1990 in the basic indicators for the Gulf States shows appreciable progress. The relation between HDI and GNP is only moderate. Some nations show the HDI ahead of income, while others show income ahead of the HDI. Finally, for 130 countries, the female HDI as a percentage of the male HDI illustrates the disparity among nations in the status of women.

6. *Summary Evaluation*. The HDI is **satisfactory** in having a clear public policy focus: that of assessing trends in development and comparing countries. It is **excellent** in the general level of aggregation in its purpose of providing an assessment of development, limited, of course, by the domains encompassed. It is **satisfactory** in providing a time series for monitoring trends. While the items of the index are as reliable and valid as their statistical systems, they are slowly responsive to programmatic changes, for example, the lag in time between the action of a program and its response in the education measures and in longevity. Generally, the index is **not satisfactory** in these respects. Neither can it be said that it is satisfactory in theory, although its defenders present arguments that ground the measures in practical wisdom. It is **satisfactory**, however, in being reported as a single number and having three domains, also reported. The domains do not encompass the scope of life experience, although longevity, education and income are undoubtedly important elements of life experience, but, as has been shown above, they do not measure the totality of life experience: **not satisfactory**. Neither can the domains be represented as subject to both objective and subjective measurement: objective, yes, but not subjective. The index, however, represents *satisfactorily* the entire population, given that all are subject to the phenomena measured. It may be said, however, that each domain contributes positively to the QOL concept: longer life, better education, and more income. The remaining three criteria are not applicable to this objective measure.

2.13. *Miringoffs' Index of Social Health*

The Index of Social Health was developed by Marc. L Miringoff and Marque Luisa Miringoff of the Fordham Institute for Innovation in Social Policy (1996, 1999). This index is one of the few that (1) evaluates several domains using reliable, objective measures, and (2) integrates these into a single index of QOL. They include 16 measures as time series since 1970, composed of: Infant mortality (as reported by the National Center for Health Statistics), child abuse (from National Committee to Prevent Child Abuse), children in poverty (measured by the Census Bureau), teenage suicide, drug

abuse (percent of teenagers using any illicit drug in the past 12 months, measured by the federally-sponsored study "Monitoring the Future"), high-school dropout rate, teenage births, unemployment, average weekly earnings, health insurance coverage (now measured by the Census Bureau), poverty among those over 65, life expectancy at age 65, violent crime rate, alcohol-related traffic fatalities, housing affordability (measured by the housing affordability index of the National Association of Realtors), and gap between rich and poor (measured by the Gini coefficient from the Census Bureau). See Miringoff et al. (1999) for complete details.

Note that these 16 components are *not* organized into the usual domains. Instead, they organize the components by age groupings, with the first three pertaining to children, the next four to youth, the next three to adults, the next two to the aging, and the last five to all age groups.

Evaluation: The index has a clear public policy purpose (criterion 1) and is rated excellent in helping public policy makers develop and assess programs at all levels of aggregation (#2) and in using time-series to allow monitoring and control (#3). Most of the components are quite reliable (#4) because the measures were developed by federal government agencies using very large samples of the U.S. population. However, the authors fail to address the question of whether these measures are valid (#4). That is, how well do these 16 components correlate with peoples' experienced quality of life? This is probably the weakest part of their project. In their recent book, they devote only one page to discussing why they chose the 16 components of their index (#5). They list seven guidelines, four of which are among our own criteria, but three of which differ from ours. The guidelines that we share are: (1) Components must be reliably measured as a time series since 1970, (4) address major issues of public policy, (5) can be disaggregated by race, gender, age, etc., and (7) they are sensitive and have changed substantially over time. However, their other three criteria are not shared by us, including: (2) Components must represent a distribution over the age spectrum, (3) reflect social (health, longevity, education, public safety) as well as economic dimensions (poverty, wages, employment, inequality), and (6) can be compared to other countries who have collected similar data. Most importantly, they have no reason for validity of

these measures: Are these the best measures for predicting peoples' perceived QOL?

The index received an unsatisfactory rating on criterion 7, that domains must encompass the totality of life experience. They do not discuss or cite whether their components span the totality of life experience. The index also received unsatisfactory ratings on criteria 8 and 9, because they did not demonstrate whether each component is a discrete portion of QOL (they did not examine correlation among components) and they do not attempt subjective measures of any component.

The index applied equal weights to all 16 components after (roughly) standardizing each. By standardizing, we mean that they attempt to put the components on a comparable scale, ranging from zero (worst performance since 1970) to one (best performance since 1970). But instead of using the usual statistical method of computing z-scores (subtract the mean and divide by the standard deviation), they subtract the minimum and divide by the range. Statisticians do not use this procedure because it has poor statistical properties: The minimum is not a stable estimator because it is vulnerable to outliers, and will vary with the number of years in the sample (Hagerty 1998). Similarly, the range is not an efficient estimator of the standard deviation. In summary, the authors could improve the reliability of their index by using traditional standardization measures. On the other hand, explaining their index to lay people is easier than explaining standardized scores.

2.14. *State-Level QOL Surveys*

There are many local, statewide, regional, and national polls that monitor QOL of certain geographic populations. This section focuses on QOL surveys conducted in a number of states in the U.S. (e.g., Virginia, North and South Carolina, Wisconsin, Rhode Island, Nebraska, Minnesota, California, Wisconsin, New Hampshire, New Jersey, Georgia, Florida, Alabama, Tennessee, Kentucky, among others). Information about these state-level QOL surveys can be obtained from the National Network of State Polls (NNSP), a data archive of state polls housed at the Institute for Research in Social Science (IRSS) at the University of North Carolina (*http://www.irss.uncc.edu/nnsp/*). These state polls contain a semb-

lance of items and measures that some may view as QOL related. However, none of these polls reports on QOL as a construct that is clearly and conceptually defined. It is beyond the scope of this paper to review all the state polls that have a "hodgepodge" of social indicators some regard as QOL related. Instead, we will select one of these state-level surveys, namely the Annual Quality-of-Life in Virginia Survey, and evaluate it as a representative of all the state-level QOL surveys.

The annual quality-of-life in Virginia Survey

The Center for Survey Research at Virginia Polytechnic Institute and State University (Virginia Tech) began conducting annual state-level surveys in 1992 (Virginia Tech Center for Survey Research, 1998). A stratified random sample of Virginia residents (approximately 2,000) are contacted by phone. From these 2,000, approximately less than half are successfully interviewed. The survey includes measures of the following constructs:

- *Overall quality of life in Virginia* through a series of rating questions, such as good/bad as a place to live, good/bad as a place to get a good college education, good/bad as a place to take a vacation, good/bad as a place to settle down and retire, good/bad as a place for young people to get a good elementary and secondary education, good/bad as a place to find a job.
- *Satisfaction with family life* through a series of rating questions, such as satisfied with family relationships, satisfied with child care services, frequencies of worries about family members getting sick and losing their job.
- *Satisfaction with work and employment* through a series of rating questions, such as satisfied with their job, proud/not proud of the work that they do at their job, whether they have flexibility in the way they do their job, whether their best friends are from their place of work.
- *Happiness and personal satisfaction* through a series of rating questions, such as satisfied with their income and financial situation, satisfied with friendships, satisfied with the communities which they live in.

- *Satisfaction with health and health care* through a series of rating questions, such as satisfied with their present physical health, satisfied with the quality of their present medical care.
- *Satisfaction with education* through a series of rating questions, such as ratings of education in terms of amount of money spent on public schools, overall education of their own community public schools, teaching of math and science in the public schools, the value of the non-teaching contributions of public universities in the state.
- *Satisfaction with law enforcement and public safety* through a series of rating questions, such as belief of increases/decreases in crime in their community, feelings of being safe/unsafe when walking at night in their neighborhood.
- *Satisfaction with the environment* through a series of rating questions, such as perceptions of enough/not enough spending to protect the environment, ratings of water quality, ratings of air quality, extent to which they recycle things.
- *Opinions on selected issues*, such as the economy, race relations, religious acceptance, abortion, violence, and the death penalty.

Evaluation: Evaluation of the index on the 14 criteria follows. The index has a clear public policy purpose (Criterion 1). Prior to 1992, policy decisions in the State of Virginia have been based solely on objective measures and data collected independently by different state governmental agencies. Examples of these data include individual income levels, consumer purchases, health care expenditures, rates of unemployment, school enrollment statistics, birth and death rates, crime rates, housing statistics, environmental pollution measures, agricultural land use statistics, number of visitors to recreational areas, etc. These social indicators are reported in such sources as the *Virginia Statistical Abstract* (University of Virginia Center for Public Service, 1996a) and the *Virginia's Local Economies* (University of Virginia Center for Public Service, 1996b). These social indicators are essentially based on objective measures. The Annual QOL in Virginia Survey complements the objective indicators by providing policymakers with subjective indicators of QOL at the local, regional, and state level.

The survey provides data for policymakers at the local, regional, and state levels. However, it can be used to measure QOL at other levels of analyses, such as the individual level, the family, the country, and the world (#2). However, it should be noted that some items are designed to measure satisfaction with state-level issues (e.g., perceptions of Virginia as good/bad as a place to live), while other items focus on community-level issues (e.g., overall education of their own community public schools), while still other items focus on individual-level issues (e.g., satisfaction with one's income and financial situation). This is methodologically problematic because if a composite QOL measure were to be formulated from the single indicators, then the question would be: QOL of whom and in relation to what? Can one form a composite QOL measure that can apply to all levels by mixing items directed to different QOL concerns in relation to the individual, family, community, regional, and state? The answer is "perhaps not". One should not mix items addressing QOL concerns at different levels. One possible solution is to focus on individual-level concerns in the various life domains. Aggregating the individual-level QOL responses by family would provide a measure of QOL of families; aggregating the individual-level responses by a geographic community would provide a measure of QOL of communities; and so on. Thus, QOL for any demographic segment can be captured by aggregating the individual-level responses of the individuals identified in that segment. Of course, this may not be the ideal solution for state-level public policy officials who are more interested in state residents' concerns with issues related to state agencies, programs, and services. As such, our recommendation is that this criterion may not be ideal for public policy decision-making. We conclude by saying that a single QOL measure may not effectively capture QOL issues to aid in policymaking at all levels – local, regional, and state. Separate and customized QOL measures may have to be designed differently for public policymakers operating at different levels.

With respect to the period-monitoring criterion (#3), we noted that annual surveys have been conducted since 1992. The annual survey includes approximately the same items and measures on

an annual basis. Therefore, periodic monitoring and control are possible.

The Annual QOL in Virginia Survey does not provide the user with a specific index that captures the overall state of QOL of Virginians. The items and measures included in the survey reflect a wide variety of experiences and opinions about the economy, health care, public education, the environment, local community services, and personal well-being. The items included in the survey are not guided by a specific theory of QOL, but by practical considerations of what public policymakers at the state and local levels would like to see measured (#4). One may argue that public policymakers may find more utility in measures that are disaggregated than those that are aggregated. To that extent, one can argue that #4 should be relaxed.

Further, no studies or data are reported concerning the reliability, validity, or sensitivity of this measure (#5). The Annual QOL in Virginia Survey does not allow the user to assess QOL using a single number. Similarly, the components of QOL (e.g., satisfaction with education, health and health care, public safety) are not constructed in a composite-like fashion, and they cannot be assembled into an overall QOL index (#6).

Although there are many life domains captured by the measure, there are a number of significant domains missing (#7). Examples include leisure and recreation, religion and spiritual life, economic well-being and standard of living, consumer and consumption aspects, and housing and neighborhood. Again, one can argue that this criterion should be relaxed. This is because the various life domains paramount for state/government decision-makers (e.g., taxation) may be quite different from those who operate at the community level (e.g., education, crime, and environmental pollution). Furthermore, professionals operating at the family and individual level (e.g., social workers, legal counselors, clinical psychologists, and psychiatrists) are concerned with life domains not considered important by community and state policymakers. Examples of these life domains include emotional life, family life, sex life, and social life.

The survey does capture a substantial and discrete portion of the QOL construct (#8). It focuses on family, work, health, education,

public safety, and the environment. These life domains are indeed substantial in most formulations of the QOL construct. The Annual QOL in Virginia Survey measures both objective and subjective dimensions of family, work, health, education, public safety, and the environment (#9). In that respect it is satisfactory. We did not rate the measure as excellent because the objective and subjective aspects of each domain do not seem to be comparable in most respects. The survey targets all adult Virginia residents. Hence, it is quite generic and is relevant for all Virginia residents. It is conceivable that the survey instrument could be easily adapted to make it relevant to all adults outside the State of Virginia, in the U.S., and overseas, too (#10).

No data are presented or discussed at this point that focus on specific domains, and the results are not reported as such to identify the unique variance contributed by non-generic domains (#11). The Annual QOL in Virginia Survey has many items that at best can be neutral but never positive (#12). Many of these items are related to health, medical care, and public safety. But on the other hand, many of the remaining items have properties that allow people to express a full range of affect. It is clearly designed in ways that differ from the dimensions of personality, cognitive processes, and affect in that they cannot be measured objectively (#13). One can argue that this criterion has to be relaxed, too. This is because there are certain life domains quite important to policymakers and other professionals that are likely to generate more dissatisfaction than satisfaction (e.g., health and finances). Conversely, other public policymakers and professionals are concerned with domains likely to generate more satisfaction than dissatisfaction (e.g., spiritual, leisure, social, and cultural). We simply cannot dismiss life domains that contribute significantly to QOL because these domains have to be "neutral" in their capacity to generate both positive and negative affect.

Finally, the survey does not distinguish between cognitive and affective dimensions of satisfaction (#14). This is because the measure does not focus specifically on domain satisfaction constructs. Again, one may argue that the inclusion of both cognitive and affective dimensions may account for greater variance in domain satisfaction. However, this fact has to be tempered by

the fact that public policymakers may not be so concerned with the predictive validity of the satisfaction constructs. They are more concerned about demonstrating the effectiveness of specific programs, services, and/or policies directly related to their jobs.

2.15. *Estes' Index of Social Progress (ISP)*

In a series of publications dating back to 1984, Richard J. Estes (1984, 1988, 1995, 1996a, 1996b, 1997, 1998) has developed an "Index of Social Progress" (ISP) and applied it to a number of nation-states around the world as well as to groups of states in particular regions of the world. The purpose of the ISP is to:

- identify significant changes in the "adequacy of social provision" occurring throughout the world; and
- assess national and international progress in providing more adequately for the basic social and material needs of the world's growing population.

"Adequacy of social provision" refers to the changing capacity of governments to provide for the basic social, material, and other needs of the people living within their borders (e.g., for food, clothing, shelter, and access to at least basic health, education, and social services, etc.) (Estes, 1984).

The ISP consists of 46 social indicators that have been subdivided into ten sub-indexes: Education, Health Status, Women Status, Defense Effort, Economics, Demographic, Geography, Political Participation, Cultural Diversity, and Welfare Effort. All 46 indicators "are known to be valid indicators of social development; indeed, the majority of the ISP's indicators are employed regularly by other scholars of socioeconomic development" (Estes, 1997: p. 3). All of the 46 component indicators of the ISP are "objective" indicators, such as "percent adult illiteracy", "life expectation in years", "real gross domestic product per head", and "violations of political rights index." All of these indicators are either taken from official government or international organization data sources or obtained from independent scholars and data-gathering organizations. Estes has computed the ISP on ten-year and five-year intervals from 1970 to 1995.

Due to the number and redundancy of the component indicators of the ISP, Estes has subjected them to a two-stage varimax factor

analysis in which each indicator and sub-index was analyzed for its relative contribution toward explaining the variance associated with changes in social progress over time. Exactly how this "changes in social progress over time" criterion for the factor analysis is defined is not specified in Estes's articles. To determine this, Estes (1988) must be consulted. Standardized scores of the component indicators then were multiplied by the factor loadings to create weighted sub-index scores which then were summed to obtain the "Weighted Index of Social Progress" (WISP).

Estes states that the WISP differs from other measures of social development in the number, range, and relevance of the indicators used in its construction. He also states that the WISP is judged to be a more comprehensive, valid, and reliable instrument for assessing changes in social development over time than such other standard indexes of national and international progress as:

- the Gross National Product (GNP);
- the Gross Domestic Product (GDP); and
- the United Nations Development Programme's Human Development Index (HDI).

Evaluation: With respect to the criteria for evaluation of existing scales in Table I, it can be noted that the WISP is grounded in a theory. But the theory is one of social development, which is not necessarily the same as an explicit QOL theory. The WISP also appears to be excellent with respect to Criteria 2, 3, 4, 5, and 6 in the table. But the index does not encompass the totality of life experience (#7), and it is not clear that each domain encompasses a substantial, but discrete, portion of a QOL construct (#8). The WISP also contains only objective dimensions (#9) and thus falls short on other subjective domains (#13) containing both a cognitive and an affective dimension (#14). Each of the domains in the WISP appears to have relevance, at least in some generic sense, for all people (#10). It is not clear that specific non-generic domains have been demonstrated to contribute unique variance (#11), but the domains (sub-indexes) seem capable of making positive, neutral, or negative contributions to overall social progress.

2.16. *Diener's Basic and Advanced QOL Indexes*

This instrument (Diener, 1995) comprises two separate indexes, the Basic QOL Index and the Advanced QOL Index. Their purpose is to discriminate between countries in terms of objective QOL criteria, except that the Advanced Index also contains one subjective component. The terms "basic" and "advanced" refer to the scales' differential application to developing and developed countries, respectively.

In the development of these indexes, the author identified shortcomings of previous scales which these instruments are intended to overcome. These are as follows: (1) That previous indexes have been developed in an unsystematic and atheoretical manner, with the result that different instruments applied to the same population yield non-comparable statistics, (2) that the choice of scale variables often means that scales discriminate better between developing or between developed countries, but not equally across the full spectrum of development, (3) that indexes based on rank-order comparisons make some countries appear to be more different than they actually are using parametric comparisons, and (4) that scale variables often do not reflect "universal values".

In order to address these concerns, Diener turned to Schwartz (1992) who had published a list of "etic" values recognized across cultures. He cites the examples of self-discipline, social justice, enjoying life, success, and protecting the environment. "These values represent three universal requirements of human existence: Meeting biological needs, coordinated social interaction, and the survival and welfare needs of groups" (Diener, 1995: p. 109).

Schwartz had listed 45 of these "universal values" and arranged them fairly evenly around a two-dimensional circumplex structure. He then divided this circle into seven pie-shaped regions that seemed to represent sets of similar values, and labeled these sections: Hierarchy, mastery, affective autonomy, intellectual autonomy, egalitarian commitment, harmony, and conservatism. Diener then selected two variables to represent each of these seven regions, one for inclusion in the Basic Index and the other for the Advanced Index. The former were chosen to reflect "lower order needs" as conceptualized by Maslow (1970), such as physical needs and

security, while the latter reflected "higher order needs", such as respect and helping others.

The resultant two indexes thus each contain seven variables. However, some of these variables are single measures, such as "physicians per capita" to represent "mastery", while others are composites. For example, "mastery" for the Basic Index is measured through the aggregate of five variables. Adding the standardized values derived from each of the seven variables creates a single scale score.

Evaluation: The construction of these two indexes meets the previously-identified problems of scale construction well. It is based on theory, items represent universal values, it is designed to yield parametric data, and the two forms of the scale are intended to cover the spectrum of national developmental status. It therefore represents a systematic advance on previous scales. However, the following issues are likely to compromise the adoption of this instrument as a standard in its current form:

1. The choice of variables to represent each of the seven "regions" is an important theoretical and empirical issue that has not been addressed. Neither has it been demonstrated that the sum of the seven "parts" represents the whole construct of QOL, even as envisaged by Schwartz.
2. The choice of some variables to represent the Basic, rather than the Advanced Index, is sometimes questionable. For example, the region of 'Affective autonomy' is measured in the Basic Index by the rate of suicide, yet high rates of suicide are more characteristic of developed than undeveloped countries (Diener and Diener, 1995; Eckersley, 1999).
3. Only one subjective measure is employed. This is the use of subjective well-being to represent the region of affective autonomy in the Advanced Index. Two issues are raised: First, that the dimension of subjective QOL is under-represented in these Indexes; second, that the combination of data from both objective and subjective variables is not simply interpretable. It is now well established within the QOL literature that subjective and objective variables are largely independent of one another (see Cummins, 2000 for a detailed discussion),

and so combining such variables within a scale is unlikely to be psychometrically acceptable.

4. In relation to the psychometric performance of these two indexes, Diener demonstrates that they are successful in discriminating between data derived from a range of different countries. However, the Advanced Index was found to be superior in this regard. It was able to discriminate at all developmental levels, whereas the Basic Index did not discriminate between the advanced countries. Moreover, the Basic Index items were dominated by income which accounted for 66% of the variance within the Basic Index, and the average correlation between variables for the Basic index was 0.27. This is indicative that the items are essentially so independent from one another that their product should not be considered to represent some common construct (Boyle, 1991).

In summary, and in relation to the criteria depicted in Table I, the Index meets well three criteria. These are #1, that it has a clear and practical purpose, #5, that it is grounded in theory, and #6, that the data can be reported either at the level of domains or in aggregate. Additionally, it partially meets five other criteria (#2, 4, 8, 10 & 13) while failing to reasonably meet the remaining six. It is concluded that these two indexes are one of the more interesting attempts at scale construction since they are based on theory and are designed to accommodate the substantial differences in developmental level between countries. However, the Advanced Index, which appears to be the superior of the two scales, has the problem of combining subjective and objective data. Additional data are also required to determine the general psychometric adequacy of this instrument.

2.17. *Cummins' Comprehensive Quality of Life Scale (ComQol)*

This scale was developed by Robert A. Cummins and work on it is explained in several places including Cummins (1995, 1996, 1997a, 1999). The scale includes measures for seven domains: Material well-being, health, productivity, intimacy, safety, community and emotional well-being. Each domain has three objective and two subjective measures. The three objective measures for each domain can be summed, and the domain sums can also be summed to obtain a single measure of objective well-being. One of the subjective

measures of each domain is a satisfaction measure (seven-point delightful/terrible item), and the other is an importance measure (five-point item). The product of the satisfaction times the importance score for each domain is a measure of the domain's subjective quality, and the seven products can be summed to obtain a single measure of subjective well-being.

Evaluation: ComQol is not based on any single theory, but Cummins offers many relatively theoretical arguments in support of it. For example, his criticisms of the concept of "health related quality of life" are theoretical/conceptual and accurate. On the other hand, he says that "If someone is asked to express their level of satisfaction with their financial situation, their response has little meaning as a measure of their life quality unless they are also asked how important they regard their finance" (Cummins, 1999: p. 21). On the contrary, Andrews and Withey (1976: pp. 119–120) experimented with satisfaction x importance scores and reported:

> Although a number of questions remain with respect to the nature and meaning of the importance measures, we have an unambiguous answer to our original question: Data about the importance people assign to concerns did not increase the accuracy with which feelings about life-as-a-whole could be predicted. . . . our results point to a simple linear additive, one in which an optimal set of weights is only modestly better than no weights (i.e., equal weights). We confess to both surprise and pleasure at these conclusions.

There appears to be no empirical demonstration in any of the Cummins' publications listed above that is inconsistent with the Andrews and Withey findings.

There are several reliability scores given for the summed objective and subjective measures and for the individual domain scores. Cronbach alphas score for the ComQol as a whole were unavailable, but in Cummins (1997a: pp. 45–48) one finds alpha = 0.54 for the sum of the objective measures, alpha = 0.65 to 0.69 for the sum of importance measures, and alpha = 0.73 to 0.81 for the sum of satisfaction measures. No alpha is given for the sum of the seven products. Five-month test/retest correlations are indicated to be 0.60 for the sum of importance measures and 0.36 for the sum of satisfaction measures.

Cummins (1996: pp. 307–308) notes that the seven domains of ComQol capture 83% of the reported domain descriptions cited in

32 studies, which is good empirical evidence for the content validity of his scale. The specific lists provided in that paper under each domain name (Table I) and the details of the studies (Table II) are very useful for researchers planning to build on his work.

The usefulness of any scale and its subscales for public policy varies with researchers' purposes and resources. For example, if one is interested in discovering people's feelings about criminal victimization with a view to initiating some kind of program of intervention in some community, then the ComQol subscale items would probably not be much help. One would need many more items and much more detail in order to decide what action might be appropriate. Similarly, because the ComQol subscale for subjective health correlates with the Short Form 36 Physical Functioning scale at r = 0.45 and with the SF-36 Mental Health scale at r = 0.60 (Cummins, 1997a: p. 50), it is not likely one would prefer ComQol to SF-36 for health research aimed at intervention. For broad-based comparisons among cities or countries, ComQol and its subscales would be more useful. If one had relatively limited resources to do a survey, the 35 items of ComQol could provide a useful overall view of the QOL in a community, city, or country. Unless there is strong empirical evidence in support of the importance x satisfaction scores, one could safely drop the seven importance items. It is likely that some reduction could also be made in the 21 objective items.

Since ComQol does not make use of any routinely-collected administrative data, special surveys are required for its application. In the interest of motivating people to fund such surveys, it would be helpful to have more validation studies in which ComQol scores are correlated with such things as crimes rates, mortality and morbidity rates, educational achievement, job and marital satisfaction, and so on.

2.18. *Michalos' North American Social Report*

The North American Social Report (Michalos 1980–1982) compares quality of life in Canada and the U.S. from 1964 to 1974. The study uses 126 social indicators grouped into 12 domains: Population Structure, Death, Disease, and Health Care, Crime and Justice, Politics and Organizations, Science and Technology, Education, Recreation, Natural Environment and Resources, Transporta-

tion and Communication, Housing, Economics, and Morality and Social Customs.

Annual stock and flow indicators were presented for each indicator. A stock indicator is the current amount of an indicator (e.g., the current violent crime rate), whereas a flow is the current change in the stock (e.g., the change in violent crime rate). The overall index was computed by weighting all indicators equally. A country received one point for each indicator on which it scored better than the other nation. For example, Canada made one point for a lower violent crime rate than the USA in 1964 and another for a slower growth rate in violent crime from 1964 to 1965. Using this method, Canada received 884 points, while the USA received 775 points over this period.

Evaluation: The index was rated as excellent on criterion 1, because it has a clear public policy purpose. For example, it compares the inputs (costs) of health programs as well as their output (health of citizens), which helps policy analysts evaluate the effectiveness of the programs. It was rated as satisfactory on criteria 2–8. The index reports only aggregated national scores (#2) but many indicators could be disaggregated further. The index was based on time-series through 1974 (#3), and updating would benefit public policy makers. Many of the indicators have been tested for reliability in large government studies (#4) but the intercorrelations among indicators were not assessed. Hence it is difficult to assess how many independent factors exist among the 126 indicators. This raises an important issue for the 'equal weighting' that is claimed for the index. If two indicators measure substantially the same factor, then the two indicators are "double counting" and receive an effective weight of two. To take an extreme example, if GDP and GNP are both included as indicators, they would be counting the same concept twice. The index is satisfactory in being grounded in theory (#5), though no formal theory is presented. The index could be reported as a single number (#6), but was effectively broken down into the 12 components. The domains appear to encompass a large part of life experience (#7) but no research was presented on this. Each domain encompassed a substantial but discrete portion of QOL (#8) but again no research supported this. The index was rated unsatisfactory on #s 9 and 14, because not all items were measured

subjectively and satisfaction was rarely measured, though the author made use of survey data when it was available. Each domain had relevance for all people (#10).

2.19. *Philippines' Weather Station*

The Social Weather Station of the Philippines conducts survey research and analysis and maintains a databank of Filipino surveys. Headed by Mahar Mangahas, the organization grew out of research activities in the Development Academy of the Philippines, which produced *Measuring Philippine Development*, edited by Mangahas (1976). The non-profit corporation was established in 1985 for the "purpose of generating survey data for social advocacy". Omnibus surveys were conducted semestrally 1986–1990 and quarterly from 1991. Topics include quality of life items, poverty, criminal victimization, satisfaction with government performance on selected issues, satisfaction with the President, and voting intentions, when appropriate, and "exit polls". The accuracy of its political predictions has gained the confidence of the public in its polling and its treatment of controversial issues.

Household interviews are conducted in eight languages of the voting-age natives. The May 1986 sample size was 2,000 (error margin of 2.2%), and, beginning biannually, with October 1986, it has been 1,200 (error margin of 2.9%) for the total estimates. Typically, regional estimates are published for Metro Manila and urban and rural for each (Luzon et al.). SWS supports an active reporting program (Mangahas and Guerrero, 1998; Arroyo, 1990). The address: Social Weather Stations, Inc., 52 Malingap St., Sikatuna Village, Quezon City, 1101, Philippines. E-mail: sws885@mozcom.com.

The QOL is assessed through two items and prospectively through one. "Comparing your quality of life these days to how it was 12 months ago, would you say that your quality of life is ... better, the same, worse?" From this question the respondent is classified as either a gainer or loser. No actual level of living is determined except by "socioeconomic class", which is determined by the interviewer following guidelines of the Marketing and Opinion Research Society of the Philippines pertaining to dwelling appearance and amenities. The latter results in five categories (a)

rich, (b) upper class, (c) middle class, (d) poor, (e) very poor. In reporting, (a), (b), and (c) are combined, usually. The "worsened" responses are subtracted from the "improved" to show the net gain or loss. Reporting also is by ecological area, socioeconomic class, and age.

The prospective item is similar: "In your opinion, what will be the quality of your life in the coming 12 months? Would you say that it will be better, same, worse?" From this confidence question the respondent is classified as an optimist or pessimist.

QOL appears to respond to the movement of the economy, a larger percentage of losers being reported in the depression years of 1984 and 1985. Similarly, the upper social classes report that the last 12 months were "better" more than the "poor", but the relationship is not perfect.

This subjective approach is favored by the social scientists because it is easily obtained, can be quickly reported, and is easily understood. It is considered a superior means of monitoring QOL to the Federally-sponsored quinquennial, Family Income and Expenditure Survey.

Poverty: Closely related to QOL is a self-rating of "poor", "not poor", or "borderline." A card is presented for the determination (Mangahas, 1995). The technique was first used in 1974 in a survey of Batangas province. In surveys since 1983, the percent rating themselves as "poor" has varied between 74 and 43%. In 1989, the total was 63%. When classified by socioeconomic class, 26% of the (combined) ABCs said they were "poor", while 84% of Class E, the very poor class, considered themselves "poor". The rural population was 71% poor, while Metro Manila registered 38%. By occupation, the range was 27% for professionals to 75% for farmers.

Respondents also are asked to estimate the pesos per month needed for home expenses. These values increase with family size up to nine persons and are higher for Metro Manila than "other urban" with "rural" being lowest.

In a regression analysis, inflation proved to have a worsening effect upon self-rated poverty. Unemployment proved of secondary importance to inflation, and GNP per capita "failed to demonstrate a meaningful ability to alleviate self-rated poverty" (Mangahas, 1995). Altogether, the regression accounted for almost half of the

variation in self-rated poverty. For Class E population, the equation with inflation alone accounts for one-half the variance in self-rated poverty. Thus, the self-rated approach appears to produce significant ratings. Other discussions include Arroyo (1989, 1990) and Mangahas (1986, 1987).

The SWS surveys also ask questions on crime victimization (Mangahas, 1996; Sandoval and Laylo, 1993; Arroyo, 1989). While crime victimization affects the quality of life, treatment of SWS surveys of it here is omitted, because this review concerns general measures of QOL, not influences upon it.

Evaluation: If all the population holds a common conception of poverty and the QOL, the use of a subjective self-report would appear satisfactory. However, in a complex social system, a common conception is unlikely. If parsimony and simplicity of data-gathering is desired, the technique described above could be used, tolerating its inherent reporting error.

The determination of social class, as well as the QOL, through observation of the characteristics and amenities of the structure of the house is a direct and simple method of classification. It reflects the work of Sewell (1940) in determining the level of living of Oklahoma farm families. With access to the living room, the observational method can be further refined to establish social class as a component of QOL (Chapin, 1935). In a large complex social system, however, both methods would require adjustment to account for cultural variation in what is considered desirable in housing and in living room décor, as well as adjustment for climatic difference in structure of the home.

Summary Evaluations: In terms of our 14 standard criteria for indices, the poverty index is *excellent* in providing a clear policy-relevant measure. The poverty measure should provide policy-makers with tangible evidence of their poverty-related programs, especially as concerns inflation – *satisfactory*. Since the survey periodically produces the poverty index over time, it is *excellent* in this respect. It is also *excellent* in the validity of its one component, as is discussed above. While not stated in its literature, the index represents an aspect of the QOL (poverty) which is theoretically accepted as undesirable; consequently it is *satisfactory* in this respect. The index is *satisfactory* in being reported as a single

number. It does not however encompass the breadth of life experience, hence is *not satisfactory* in this respect. It is *satisfactory* in representing the single domain, poverty, but it does not pretend to encompass the range of QOL sectors. It is measured subjectively, in that the respondent reports his/her sense of improvement or decline in well-being and the interviewer assesses the SES of the household, and it is subject to being measured objectively (in terms of income) – *satisfactory*. It also is ***satisfactory*** in being relevant to all Filipinos. The resulting index rises and falls and is thereby a positive or negative trend in well-being: *satisfactory*. The remaining evaluation categories are not applicable to the poverty index.

2.20. *Netherlands Living Conditions Index (LCI)*

The Netherlands Social and Cultural Planning Office (Boelhouwer and Stoop, 1999) has developed the Living Conditions Index (LCI). It was initiated in 1974, with reports every year since then. It was designed for the specific purpose of public policy "to reflect conditions in areas that are influenceable by government policy" (p. 51). The LCI index is reported as a single index (=100 in 1997), but can be broken down into its components of: Housing, health, purchasing power, leisure activities, mobility, social participation, sport activity, holidays, education, and employment. Current domains and indicators for each domain are shown in Table III. The specific indicators have changed over the years to address new public policy problems. The composition of earlier versions is given in Boelhouwer and Stoop (1999).

Evaluation: The index has a clear public policy purpose (criterion 1), and can be disaggregated to any level (#2). The continuing study has a well-defined time series (#3) and can be broken down into individual components (#6). The authors argue strongly that only objective indicators (#9) should be included in the index, because only these are controllable by public policy. Nevertheless, they also collect measures of overall happiness in order to validate their LCI against perceived happiness. These simple correlations in 1997 were all significant and in the expected direction (see their Table III). Further, their LCI is more reliable (#4) than the separate components, because the correlation of LCI with happiness is higher

TABLE III

Domains and Indicators for Netherlands Living Conditions Index (LCI)

Domains	Indicators
1. Housing	a Type of ownership
	b Type of building
	c Number of rooms
	d Area of living room
	e Scary spot in the neighbourhood
	f Year of construction
2. Health	a Number of psychosomatic symptoms
	b Number of nine serious illnesses
	c Number of other illnesses
3. Purchasing power	a Number of household appliances
	b Number of hobby articles
4. Leisure activities	a Number of hobbies
	b Number of nondomestic entertainment activities
	c Organisational membership
5. Mobility	a Car ownership
	b Season ticket for the railway
6. Social participation	a Active contribution to activities of organisation
	b Volunteer work
	c Social isolation
7. Sport activity	a Number of times sporting a week
	b Number of sports
8. Holiday	a Holiday trip in past year
	b Holiday trip in foreign country
	c Number of holiday articles

than the correlation of any of the separate components. Hence, the separate domains are not redundant, but provide some additional predictive validity. However, they should report a multiple regression to sort out which domains add significant explanation to LCI and happiness. They assign components unequal weights in computing LCI by factoring the components and using the loadings on the first factor as weights. However, this could be improved by

using the weights from a multiple regression in predicting happiness. The resulting weights would make LCI the best forecast of subjective happiness.

2.21. *German System of Social Indicators*

The German System of Social Indicators was developed as early as in the mid-seventies within the framework of the SPES Project (*Sozialpolitisches Entscheidungs – und Indikatorensystem*; Sociopolitical Decision and Indicator System), directed by Wolfgang Zapf (1997). Since 1978, this system has been continuously updated (#3) and further developed by the *Abteilung Soziale Indikatoren* (Department of Social Indicators) of the *Zentrum für Umfragen, Methoden und Analysen* (ZUMA; Center for Survey Research and Methodology), located in Mannheim (Noll and Zapf, 1994).

In its present form the German System of Social Indicators covers the following 13 life domains: Population, Socioeconomic Status and Subjective Class Identification, The Labor Market and Working Life, Income and Income Distribution, Supply and Consumption of Goods and Services, Transportation, Housing, Education, Health, Participation, The Environment, Public Safety and Criminality, and Leisure and Media Consumption.

Each domain is subdivided into a number of dimensions and subdimensions. The health module, for example, consists of three welfare or goal dimensions: (1) state of health; (2) health care coverage; (3) living conditions that affect health. For each of the three dimensions there are 4 to 5 subdimensions, which are for example, in the case of the state of health: (1) Life expectancy, (2) causes of death, (3) healthiness of life, and (4) subjective satisfaction with general health. Each of the subdimensions is again measured by one or sometimes two or more indicators. A complete list of dimensions, subdimensions, and indicators is given for the domain "Participation" in Table IV. These multiple measures contribute to reliability of the constructs (#4).

The approximately 300 indicators and almost 1000 time series currently comprising this system of social indicators provide the user with an empirical view of the changes in the living conditions of the general population and the transformation of the macrostructures of the society (criterion 1). The period of observation stretches from

TABLE IV

Dimensions, subdimensions,and indicators for the domain of "Participation" from the German System of Social Indicators

Dimension	Subdimensions	Indicators
Participation in the Political Context	Voter Turnout at the Federal Elections	Percentage of those eligible to vote who voted in the elections for the German Parliament (Bundestag)
Participation in the Political Context	Voter Turnout at the Elections for the European Parliament	Percentage of those eligible to vote who voted in the elections for the European Parliament
Participation in the Political Context	Voter Turnout at the Elections for the State (Bundesland) Parliaments	Percentage of those eligible to vote who voted in the elections for the individual state parliaments (Landtage)
Participation in the Political Context	Percentage of Political Party Members (According to the Welfare Survey)	Percentage of respondents aged 18 or above that declared to be a member of a political party
Participation in the Political Context	Percentage of Political Party Members (According to the Yearbook of the Allensbach Institute for Public Opinion Research)	Until 1991: The question was posed: "Would you be willing to join a political party?" Percentage of those who responded 'I am already a party member", 1993: The question was posed: "Would you be interested to become active in a political party?" Percentage of those who responded "I am already active in a party"

TABLE IV

Continued

Dimension	Subdimensions	Indicators
Participation in the Political Context	Percentage of Respondents Who Are Involved in Citizens' Action Groups, Political Parties and Local Politics	Percentage of the respondents who answered 'every week' or 'every month' to the question "in which of the following activities are you involved in during your free time? Please tell us how often you do one of the following: participate in citizens' action groups, political parties or local politics."
Participation in the Political Context	Percentage of Potential Party Members	Until 1991: Percentage of respondents answering "Yes" to the question: "Would you be willing to join a political party?", 1993: The question was posed: "Would you be willing to become active in a political party?" Percentage of those who responded "I would be interested in becoming active in a political party"
Participation in the Political Context	Percentage of Those Strongly Interested in Politics	Percentage of respondents answering "Strongly" or "Very Strongly" to the question: "To ask a very general question: How strongly are you interested in politics?"
Participation in the Employment Context	Rate of Unionisation	Percentage of persons in dependent employment and of the unemployed who are union members

TABLE IV

Continued

Dimension	Subdimensions	Indicators
Participation in the Employment Context	Companies with a Works Council	Number of companies that had a works council election in the respective year
Participation in Other Social Contexts	Percentage with a Religious Affiliation	Percentage of the resident population that has affiliations with a specific religious denomination
Participation in Other Social Contexts	Percentage Attending Religious Services	Percentage of all church members who attend Sunday service
Participation in Other Social Contexts	Percentage of Members of Associations/Clubs	Percentage of respondents who are members of at least one association or club
Participation in Other Social Contexts	Percentage of Voluntary Workers	Percentage of respondents who answered the question: "In which of the following activities do you participate in during your free time? Please tell us how frequently you participate in this activity." for the response item "Voluntary work for clubs, associations or social services" with "every week" or "every month"

TABLE IV

Continued

Dimension	Subdimensions	Indicators
Participation in Other Social Contexts	Percentage of Respondents that Provide Assistance to Their Relatives and Neighbours	Percentage of respondents who answered the question: "In which of the following activities do you participate in during your free time? Please tell us how frequently you participate in this activity." for the response item "Helping out friends or neighbours if there is something to do". with "every week" or "every month"
Subjective Evaluation	Satisfaction with the Possibilities for Political Participation	The question was posed: "How satisfied are you with the overall possibilities for political participation in our country?" Respondents' average answer in terms of a scale ranging from "0" ("completely dissatisfied") to "10" ("completely satisfied")
Subjective Evaluation	Satisfaction with the Church	The question was posed: "How satisfied are you – overall – with the church?" Church members' average answer in terms of a scale ranging from "0" ("completely dissatisfied") to "10" ("completely satisfied")
Subjective Evaluation	Satisfaction with Democratic Institutions	The question was posed: "How satisfied are you – overall – with the democratic institutions in our country?" Respondents' average answer in terms of a scale ranging from "0" ("completely dissatisfied") to "10" ("completely satisfied")

the beginning of the fifties until the present and thus encompasses the entire period of existence of the Federal Republic of Germany within its former boundaries.

Evaluation: The system of indicators (#5) is a systematic selection of indicators chosen according to both theoretical and sociopolitical criteria to measure individual welfare. Individual welfare is conceived both in terms of the objective living conditions that define it and in terms of its subjective assessment on the part of citizens. The majority of the indicators in the system presented here measure the degree to which goals and values of society relevant to individual welfare have been realized in each of the 13 life domains or policy areas. There are many similarities with the OECD system of social indicators. In both systems, social indicators are conceived as measures of individual welfare or well-being, and social indicators should be output-oriented, be relevant to policy, be applicable over a long period of time, form part of a comprehensive grid, portray all areas of social concern, correspond closely to the social concern to which they relate, and form an integrated framework of definitions, specifications, statistical guidelines and disaggregations (OECD, 1982).

The process of establishing this system of social indicators included several steps. Once goal areas had been selected, goal dimensions and goal values within the selected areas were analyzed as a first step. This was done by investigating laws and other kinds of regulations as well as government programs, the major political parties, and important associations like the trade unions (#1, 2). In a second step, a systematic goal catalogue with various dimensions and subdimensions were constructed for each life domain and finally appropriate indicators were selected and constructed. The latest publication is a volume of 260 pages presenting the time series in table format (Noll and Wiegand, 1993). Series are disaggregated by sex, age, income group, and others (#2). Although the system achieves many of the 14 criteria, the reports do not attempt to compute a single composite index of QOL (#6), instead preferring each reader to decide their own weights for combining domains.

2.22. *Swedish ULF System*

ULF (*Undersokningar av Levnadsfordhallanden*) is a general survey system run by the Welfare Analysis Program of Statistics Sweden, starting 1974 with a fixed and standardized survey, later expanded and converted into a system of integrated surveys, combined with utilization of public registers. Since the late '70s, ULF is designed as an annual data collection in four thematic surveys, each running for a period of 24 months (sample size for two years: 15200 persons aged 16–84 years). After eight years the entire program is repeated. Each survey covers a base program of slightly more than 120 social indicators within 12 social domains. Along with a large set of background data, the social indicator module covers about half of the interview time. The second half of the interview (30 minutes) covers detailed information within a selection of the domains, which is included during two years, and then repeated eight years later in the next cycle. The four thematic modules concern working life, health and health care, social relations, and environment (these headings imply a grouping of domains).

The Nordic welfare concept is structured hierarchically along 12 social dimensions (#7) or welfare domains (education/training, social mobility, employment, working environment, material living standards, housing, transports, leisure, social contacts, participation, victimization, and health), within which a set of 125 standardized social indicators is identified to cover the 12 domains. Further detailed information relating to sub-domains and special research topics has been added flexibly. In the case of ULF system the program is extensive (700 variables). Vogel and Häll (1997) present the standard roster used in the latest general social report based on ULF.

In the general framework of the Nordic welfare research tradition, welfare indicators are usually defined as *resources* to be utilized and converted in various *arenas*. Access to and conversion of resources relates to basic welfare policy issues (Criteria 1 and 2). Welfare indicators are primarily chosen as *objective measures* of factual conditions relevant to *influence by social planning*.

Evaluation: In the Nordic tradition QOL research is closely linked to social planning of the welfare state (criterion 1), and data can be disaggregated to any desired level (#2). The data is based on one of the longest social time series in the world (#3). The data are grounded in social science theory, and each module is developed by experts in economics, sociology, etc. (#5). However, there have been no serious attempts in constructing global or domain indexes, neither by the research community, nor at Statistics Sweden (#6). Their position is that there is no obvious theoretical foundation for such constructs, which could be based on a consensus among researchers, nor among society at large. "Simple additive indexes without theory may not contribute to our understanding, but could obscure reality. The idea of indexes is *reduction* and *transparency*. This reduction should be based on empirical experience as well as theoretical considerations. We search for a very short list for social indicators (less than ten), which will speak to the public and attract attention in the administration as well as the media."

The domains appear to represent the totality of life experience, though they present no evidence of this (#7 and 8). Statistics Sweden has given priority to objective facts (#9) which fit in econometric and other models, and relate directly to goals in social and economic planning of transfers and services, labor market policy, housing planning, training, etc. Hence, subjective indicators (perceptions, satisfactions, welfare demands, happiness, etc.) are of secondary interest, which distinguishes Swedish and Nordic welfare research from the mainstream QOL research. They cite two major reasons. First, subjective perceptions of living conditions are always affected by perceptions and levels of aspiration, in general terms reflecting the power structure and media. Data relating to perceptions and preferences are, of course, an interesting field of research in itself and sometimes included as supplements. Second, there is a strong Nordic tradition of using welfare statistics to influence general policy and provide information for the public debate on the progress in living conditions. In both cases objective information on factual conditions is required. However, other domains usually not considered directly in social planning, such as participation and social networks are included to attempt to measure more subjective domains. The remaining criteria received satisfactory ratings.

3. SUMMARY AND AGENDA FOR RESEARCH

A summary of the performance of the 22 QOL indexes is shown in the last column of Table I. It displays the average score achieved for each criterion by the indexes. It shows that the indexes (on average) achieved some criteria quite well (scores that are close to 3), but that they failed to incorporate other criteria (scores close to 1). The criteria that were achieved best were: (1) The index has a clear practical purpose for public policy, (2) the index should be based on time series to allow periodic monitoring and control, and (3) the index should be reported as a single number, but can be broken down into components. In contrast, the criteria that were NOT well achieved (on average) were: (1) The domains must encompass a substantial, but discrete, portion of QOL, (2) the index must be grounded in well-established theory, and (3) the composite index should be reliable, valid, and sensitive. We wish to expand on these last three areas to suggest solutions for the shortcomings of existing indexes and to suggest an agenda for future research. We discuss each in turn.

3.1. *Domains must encompass a substantial, but discrete, portion of QOL*

Ever since the ground-breaking publications of Andrews and Withey (1976) and Campbell, et al. (1976), many researchers have followed these authors' lead in considering QOL to comprise a number of domains. This convention is both intuitive and pragmatic. The use of such terms as "health" and "relationships" to designate defined sections of life experience is a strategy common to all languages. Such designations are also useful to researchers because they segment the grand QOL construct into components that are easier to conceptualize and measure. However, despite agreement among QOL researchers at this general level, there are very divergent opinions concerning the number and character of such domains. An extraordinary diversity of domains has been proposed, as a few examples will illustrate – natural endowment (Wills, 1995), sexual functioning (O'Boyle, 1994), personal growth (Ryff, 1995), absence of depression (Viney et al., 1994), use of public services (Pirfo et al., 1994), meaning of life (Kuyken et al., 1994), and so on.

The fact that almost any measure of human experience or living conditions could be co-opted as a QOL domain was noted almost 20 years ago by Van Dam et al. (1981) who suggested that the term "QOL" be replaced by individual dependent variables identified by their name. But their advice has gone unheeded as authors seemingly compete to identify yet another variable that can be labeled as a QOL domain.

QOL instruments can be designed at four levels, as follows:

1. No domains. There are two versions of such scales. The first and simplest is the single question, "How satisfied are you with your life as a whole?" (Andrews and Withey, 1976). The second is those scales with multiple questions that are intended to measure the single latent construct of "life as a whole". A well-known example is the five-item Satisfaction with Life Scale (Diener et al., 1985b). The advantage of such scales is their simplicity and brevity. Their disadvantage is that they provide no comparative information on the component parts of the QOL experience and they can only be constructed to measure the subjective domain. See Michalos (1991) for a review of correlations among diverse SWB scales.
2. Single domain. Here the intention is to generate a multiple-item scale that measures a latent construct envisaged as a single QOL domain. Such scales may be either objective or subjective. For example, the Health Index (Rosencranz and Pihlblad, 1970) measures types of medical illnesses, health problems, and length of confinement, which sum to an index of objective health, while the Spiritual Well-Being Scale (Paloutzian and Ellison, 1982) combines religious and existential sub-scales to form a single subjective index for this domain.
3. Multiple, intentionally overlapping domains. These are the diagnostic instruments that measure the details of life within some narrow specific context. They may be either objective or subjective. Examples of the former may lie at the level of social indicators, such as the American Demographics Index of Well-Being (Kacapyr, 1997) where the five domains all share considerable variance within a socioeconomic framework, and the Personal Assessment of Intimacy in Relationships Scale

(Schaefer and Olsen, 1981), which uses five related sub-scales to achieve an index of perceived relationship adequacy.

4. Multiple, maximally independent domains. The construction of such instruments poses the greatest challenge of test developers in this area. In addition to the usual psychometric requirements of reliability, validity, and sensitivity, such instruments must be maximally parsimonious while also encompassing the entire QOL construct. So, one way to judge the adequacy of such instruments is to seek the ones with the smallest number of domains and then ask whether those domains are adequate, in aggregate, to define QOL. This can be tested through the use of multiple regression, as previously described.

The optimal number and character of domains in this context is not yet fixed. However, two prior reviews of research have used meta-analysis to narrow the possibilities considerably.

The first review incorporated data from 32 articles that provided data on "satisfaction" with some 173 differently-named aspects of life (Cummins, 1996). Each term was then classified as to whether or not it could be placed within one of the seven domains that formed the Comprehensive Quality of Life Scale (Cummins, 1997a), and it was found that 68% were able to be categorized under one of the domain headings. Moreover, due to the use of the same term by multiple authors, a total of 351 terms had been employed, and, of these, the classification system incorporated 83%.

The second review (Cummins, 1997b) was based on an analysis of 27 QOL definitions that had named domains and also a number of surveys that had asked respondents to rank QOL domains in terms of their perceived importance. If the frequency data from these two reviews are standardized, then combined, and their product expressed as a percentage of the most influential domain (relationships with family and friends = 100), the order of the remaining domains is emotional well-being (98), material well-being (77), health (67), work and productive activity (61), feeling part of one's local community (29), and personal safety (27).

Recommendation: We recommend a standardized terminology for QOL domains. Our review concludes that seven domains can adequately span the space of perceived QOL. Their names (and

approximate importance weights) are: Relationships with family and friends (weight = 100), emotional well-being (98), material well-being (77), health (67), work and productive activity (61), feeling part of one's local community (29), and personal safety (27).

Several comments pertain to this list as follows:

1. While these seven domains seem to span the QOL space that is common to all, other "supplementary domains" may be important to particular populations. For example, "leisure" is a domain that several national indexes measure and is especially appropriate for developed countries where leisure is an issue. (The leisure domain loaded most highly on the "emotional well-being" domain found in the meta-analyses.) In contrast, a domain measuring "political participation" may be important to measure in countries that have only recently adopted democratic institutions. In all cases, the test that we propose is whether the supplementary domain adds unique variance beyond the first seven domains in the target population.
2. The relative degree of influence that has been calculated for the seven domains is on a normative basis only and would certainly be revised within specific groups. For example, while "personal safety" has the lowest ranking, this could quite possibly become the pre-eminent domain in a situation of personal danger. In addition, the domain that combines "family and friends" might justifiably be separated into two separate domains in situations where either family or friends were considered largely irrelevant.
3. Using the seven domains recommended here, Table V displays how some of the QOL indexes fit into the classification. It shows that many domains of other indexes can easily be renamed to provide consistent terminology. For other indexes, Table V highlights the domains they lack.

In summary, this paper is intended to provide a starting point for theoretical and empirical investigation into the domain structure of QOL. The criteria we have suggested are intended to limit the designation of variables as "domains" through the specification of characteristics we believe are central to the QOL construct such domains are intended to represent. The subsequent identification of seven domains that meet the aforesaid criteria is intended to provide

TABLE V

Domains of QOL measured by some indexes

COMQOL	Index of Economic Well-Being	American Demographics Index	Johnston's QOL Index	U.N. Human Development Index	Index of Social Health	Virginia QOL Survey	Estes index of Social Progress	Swedish ULF
Relationships with family and friends (weight = 100)		Divorce rate	Family stability		Child abuse	Satisfaction with family life		Social contacts
Emotional well-being (98)					Teenage suicide	Happiness		
Material well-being (77)	GDP, housing stock, capital investment, poverty rate	Income and employment, Consumer expectations; Productivity	Earnings and income, housing	PPP/capita	Average weekly earnings, access to affordable housing		Economic	Material living standards, housing, transport, leisure
Health (67)			Health	Longevity	Infant mortality, health insurance	Satisfaction with health	Health status	Health

TABLE V

Continued

COMQOL	Index of Economic Well-Being	American Demographics Index	Johnston's QOL Index	U.N. Human Development Index	Index of Social Health	Virginia QOL Survey	Estes index of Social Progress	Swedish ULF
					coverage, out-of-pocket health costs for >65	and health care		
Work and productive activity (61)	Employment		Employment, education		Unemployment, high school dropouts	Satisfaction with work and employment, satisfaction with education	Education	Employment, working environment
Feeling part of one's local community (29)		Number of endangered species	Poverty, equality	Literacy	Poverty rate >65, gap between rich and poor	Satisfaction with environment	Women status, political participation, cultural	Social mobility, participation

TABLE V
Continued

COMQOL	Index of Economic Well-Being	American Demographics Index	Johnston's QOL Index	U.N. Human Development Index	Index of Social Health	Virginia QOL Survey	Estes index of Social Progress	Swedish ULF
							diversity, demographic	
Personal safety (27)	Crime rate	Crime rate	Public safety		Homicides, alcohol-related traffic fatalities	Satisfaction with law enforcement	Defense effort, geography	Victimization
Subjective QOL						Good place to live, go to school, go to college, retire, get job		

no more than a conceptual baseline for future development. It is thus our hope that those ideas contribute to the cohesion of future research within the QOL field, in addition to providing a rationale for scale evaluation in this paper.

3.2. *The Index Must be Grounded in Well-established Theory*

Most of the indexes reviewed failed to specify any well-established theory behind the index. By theory, we mean the "nomological net" of concepts and causal paths that specify how QOL is related to exogenous and endogenous variables. By "well-established", we mean that its parts have been subjected to empirical test.

Since the existing indexes performed poorly on this criterion, we propose a solution that is a systems-theory approach to QOL. Figure 1 displays the path analysis diagram of our approach. It was first suggested by Veenhoven (1998) and Cummins, et al. (1998). The first column of the figure contains *input* variables, which denote exogenous environmental variables affecting citizens' QOL. Common measures of this are "objective indicators" of QOL, and examples are listed in that column as GDP/capita, extent of freedom, income inequality, etc. Many of these indicators can be *controlled* by public policy to improve QOL and are much studied by policy analysts to learn how to improve them. The second column of the figure contains *throughput* variables, which describe the *individual's response* to this environment (e.g., education achieved, marriage choice). These are also often measured as "objective indicators", but reflect the *individual's choice* in response to the environment and to public policy. The fourth column contains *output* variables, which is the result of input and throughput. Veenhoven summarizes these ultimate outputs as: Happiness (or SWB), personal survival, and "contribution to the human heritage". Finally, the third column is our partition of SWB into domains of life, discussed in the previous section.

The distinctions in Figure 1 among input, throughput, and output are not only familiar to the systems theorist and to the biologist (Veenhoven, 1998), but also to the public policy analyst (Hoos, 1972). The main difference in Figure 1 from a policy analyst's approach is that the policy analyst ignores our output columns and considers the *throughput variables as final output*. For example,

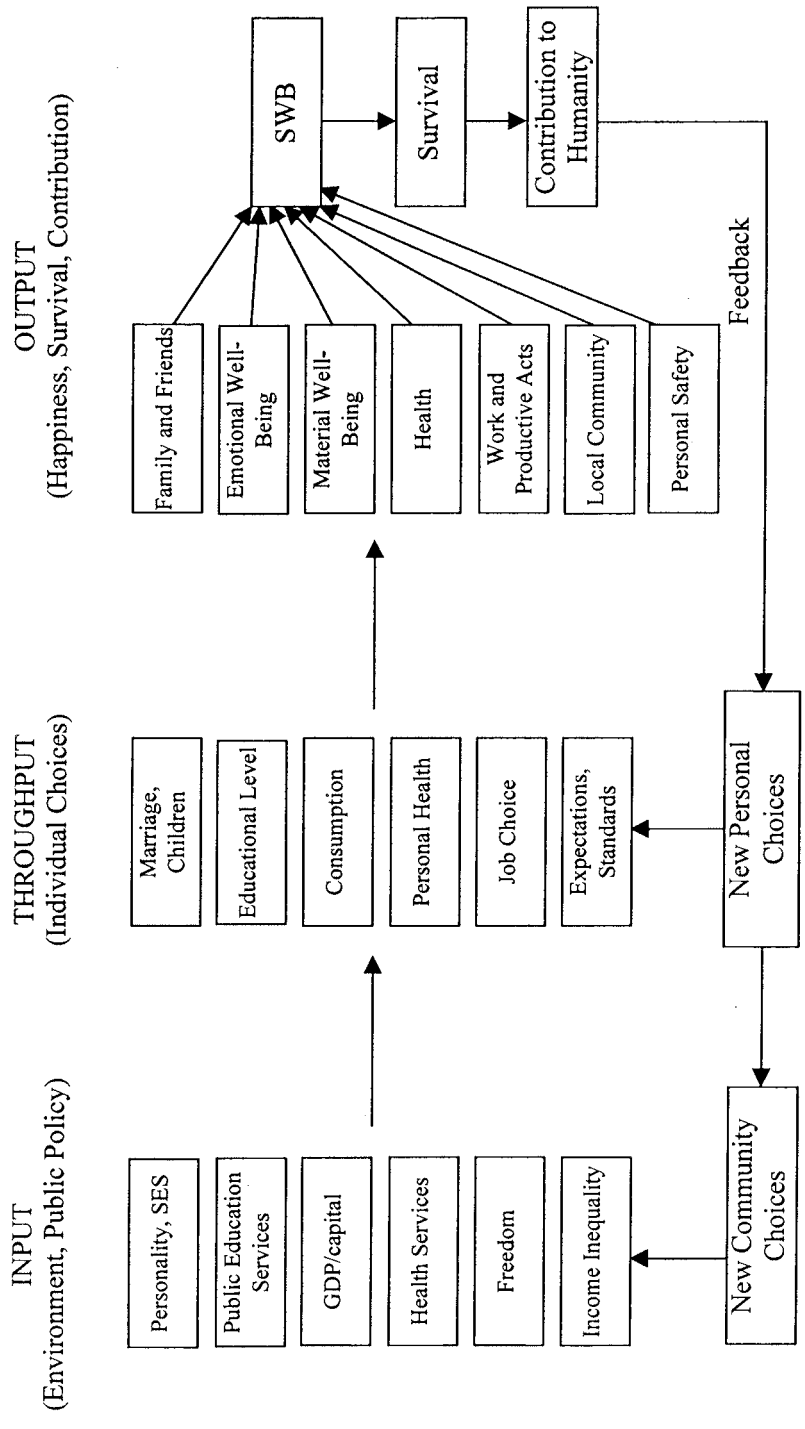

Figure 1. Systems theory structure of Quality of Life concepts and causes.

recent public policy research on "welfare to work programs" examines the effect of tax incentives and child care availability (inputs that can be controlled by public policy) on individuals' choices to get more education and to enter the work force (throughputs in our approach, but outputs in their approach). Public policy analysts sometimes cite anecdotal evidence that workfare "increases people's self esteem", but most don't realize that they can formally measure such subjective factors (output columns). Maintaining the distinctions of Figure 1 increases the usefulness of QOL indexes to public policy analysts because they structure policy analysis similar to Figure 1.

Recommendation: A QOL index for public policy should contain distinctions between input, throughput, and output variables to aid in clarifying and estimating effects of public policy inputs.

Most of the indexes we have reviewed do not distinguish among these four columns, and therefore attempt to add "apples to oranges" by aggregating indexes from various columns together. For example, Diener (1995) aggregates "SWB" (in the fourth column) with "physicians per capita" (in the first column) to get his QOL index. Yet to make an index useful to public policy, it is necessary to clearly separate *input* (which can be influenced by public policy) from *output* (the ultimate goal) and to show policymakers the causal links between the two. None of the indexes we have reviewed have presented this entire "systems view".

Nevertheless, each can be useful in measuring different stages of the system. For example, Veenhoven's and Cummins' indexes measure *only output* (SWB and survival). Estes' index and the Swedish ULF measure only input and throughput, but no output (except survival in the form of life expectancy). Therefore, integration of several of these indexes would greatly improve the theory and provide clearer controls for public policy.

3.3. *The Composite Index Should be Reliable, Valid, and Sensitive*

Most of the indexes have been well tested for reliability. The components that are objective measures (e.g., GDP/person) have long been judged reliable, while the components that are subjective measures (e.g., satisfaction and SWB) have shown acceptable

validity in this review. Further, most of the indexes have demonstrated some *convergent validity* (they correlate with other measures of the same concept. See Michalos, 1991 for a review).

However, all of the indexes fall short on *sensitivity*. By "sensitivity", we mean their ability to predict future "outputs of the system", sometimes termed *predictive validity*. In this case, it means that public policymakers need the ability to predict the effect of public policies (first column of Figure 1) on the outputs (last columns of Figure 1, including domain satisfaction, overall SWB, and survival), and that the outputs must be sensitive to variations in input, yet very few of the indexes have been subjected to this test of predictive validity.

Recommendation: A QOL index for public policy should be subjected to predictive validity tests, especially validating how the inputs of public policy affect the outputs.

The QOL indexes that have been subjected to some predictive validity tests are the Eurobarometer and Veenhoven's Happy Years Index. Research on these has successfully validated some of the arrows of Figure 1. For example, Diener et al. (1993), Moller (1997), Veenhoven (1998), and Hagerty (1998) have all found evidence for the effect of income or GDP/person on life satisfaction and happy years. Hagerty (2000) has documented the link between income inequality and life satisfaction. Veenhoven (1993) and Michalos (1991) have summarized the substantial link between marriage (or living together) and happiness. Hagerty et al. (2000) have documented the feedback link by showing that declining outputs result in voters punishing the governing party in hopes of eliciting better public policy. Vogel (1998) has examined the link between government and family welfare systems with objective living conditions (column 2) in European countries. We encourage other researchers to evaluate the predictive validity of QOL indexes from public policy inputs to predicting throughputs and outputs.

We point out that many of the arrows between the first column (public policy) and the second column (individual choice) are already being validated by economists and public policy analysts. Hence, the systems theory approach that we propose will already be comfortable to public policy analysts.

3.4. *The Index Should Weight Domains Appropriately*

An important topic related to reliability and validity is how to define importance weights for each domain. Such weighting would be used in the computation of a weighted average of the domains to produce a final composite score for QOL.

Not surprisingly, weighting domains in the computation of composite indicators is a debated topic. Perhaps to avoid such controversy, some QOL indexes do not weight and do not provide an explanation for this approach. However, no weighting is still a form of weighting – equal weighting. In fact, Estes' ISP used equal weighting in computing the subindex scores. In this way, Estes' subindex of health is an average of the standardized scores for (1) male life expectancy at birth, (2) rate of infant mortality per 1,000 live-born, (3) population in thousands per physician, and (4) per capita daily calorie supply as a percentage of requirement.

In another illustration of no weighting, International Living (IL) used no weights for the seven QOL dimensions it uses in rendering a final score for nations in its country QOL survey (Peterson and Malhotra, 1997). In this case, no weights were likely less trouble for researchers who had already devoted much toil to the assessment of objective QOL dimensions in more than 150 countries. Recently, IL has adopted an apparently atheoretical weighting of these dimensions.

The general controversy about weighting is joined by some researchers who claim equal weighting is not far behind the performance of purportedly optimal weighting schemes (Andrews and Withey, 1976: pp. 119–120). In Andrews and Withey's study, two approaches were used to obtain weights. In the first approach, the overall QOL variable was regressed onto eight dimensions of QOL to derive weights for each dimension when computing summated scores for individual respondents. A small increase in predictive ability resulted from this approach (the correlation between the unweighted summated scores and overall QOL was 0.67, while the correlation between the weighted summated scores and overall QOL was 0.71).

In Andrew and Withey's second approach, respondents were asked to rate the importance of the QOL dimensions. These importance ratings were used in effect as additional variables in two

different multiple regressions of overall QOL. Neither of these regression analyses posted better adjusted R^2s than a multiple regression without the importance ratings included. Despite the apparent lack of improvement in modeling results when including importance ratings, it appears that an expectancy-value modeling approach in its truest sense was not employed by Andrews and Withey (Fishbein and Ajzen, 1975). In such an expectancy-value approach, each respondent's rating of a QOL dimension would have been weighted by the respondent's importance rating for the QOL dimension prior to summing. In conclusion, a true expectancy-value approach might improve modeling results in future research. However, in the case of IL, Andrews and Withey's claims appear to be supported, as the overall QOL ratings for countries are highly consistent over the years of equal and non-equal weighting (Malhotra et al., 1998).

Unfortunately, what is intended to be "no weighting" does not always result in equal weighting. In apparently being unaware of how weighting can unintentionally occur, *Money's* Best Places appears to become fatally flawed. Guterbock (1997) analyzed *Money's* Best Places ratings and concluded the volatility of these annual ratings resulted simply from more items being used to measure the economic dimension. If Guterbock's conclusions are true (*Money* does not disclose its weights for dimensions of its place ratings), this suggests a profound flaw in *Money's* methods in rating overall QOL. Guterbock recommends the study would better be called "*Money's* Financially Hot Places", instead of "Best Places". In this example, a regrettable misrepresentation of overall objective QOL could now likely be in wide dissemination.

Recommendation: We propose using two advanced statistical methods to improve weighting – two-stage factor analysis and conjoint analysis.

Two-stage factor analysis is a structural equation model (Bollen, 1989; Joreskog and Sorbom, 1988) that could be used to better gauge the relative importances of indicators to each domain and each domain to overall QOL. The following hypothetical example illustrates this proposed approach, using Estes' (1998) ISP index. Here, indicator x_{36} for the country of Canada might receive a rating of 50 on a standardized scale ranging from 0 to 100. In the first

confirmatory factor analysis, the factor loading of x_{36} on the tenth dimension might be found to be 0.90. In the second-stage confirmatory factor analysis, the tenth dimension might have a loading of 0.33 on its second-order QOL factor which accounts for 10% of the variance in this second stage. Thus, the contribution of x_{36} to the weighted average estimate of overall QOL for Canada would be $50 \times 0.90 \times 0.33 \times 0.10 = 1.5$. Similar computations would be made for the remaining 35 indicators. Summing the 46 products from a set of 36 multiplications would then render an estimate of overall QOL score for each country being rated in the study.

Despite the direct nature of this higher-order factor analysis, methodological issues and opportunities are presented by the two-stage factor analysis approach just described. As an example of such an issue, the factor scores produced in each stage of the analysis would not be deterministic, but would be estimates because of the nature of common factor analysis on which the structural equations approach is based (Bollen, 1989). This shortcoming seems to be slight because a more truthful underlying structure would have likely been identified using this approach in lieu of a principal components approach. As an example of an opportunity provided by this approach, an additional factor analysis of the factor scores for the three dimensions could be executed, if researchers felt one general factor underlies QOL. Such a third-order factor analysis would be similar to the derivation of "g" in the psychometric measurement of intelligence (Morrow and Morrow, 1973).

The second approach to deriving weights could be undertaken using conjoint analysis (Hair et al., 1998). This decompositional modeling technique is based on the premise that humans evaluate objects based on the separate amounts of value provided by each attribute. The large number of attributes in studies such as Estes' might preclude conjoint analysis' use in objective QOL studies. However, if domains of subjective QOL could be kept to a set of 6 to 9 variables, as in the Virginia QOL study (i.e., family life, work, personal satisfaction, health care, education, public safety, environment) a hybrid conjoint modeling approach could be employed to derive part-worths or relative importances for the variables and their respective levels ("satisfied" – "tolerable" – "worse than tolerable" – "unacceptable") (Green, 1984). Such hybrid or computer-

administered methods such as ACA are likely to demand large sample sizes, and optimum designs (Kuhfeld et al., 1994).

In sum, methodological developments with factor analysis techniques using structural equations and with conjoint analysis could provide insights into the relative importances and weights to be used in QOL studies. While not without challenges, these techniques could provide valuable boosts to much-needed theory development in QOL research.

4. CONCLUSIONS

The last 30 years have seen a profusion of QOL indexes, many of which are reviewed here. We summarize their achievements first, then their weakness, and recommend further research to improve the indexes.

1. Most indexes are designed with a clear public policy purpose.
2. We now have *time series* data from many countries that extend back 25 years. The data include measures of happiness, satisfaction with domains, survival statistics, and many other social indicators.
3. Most can be *disaggregated* to examine problems of inequality and distribution of QOL in subgroups.
4. The component measures are largely *reliable*, and show some convergent *validity* with other reasonable measures.

We therefore conclude that many of these indexes are *potentially* very useful for public policy and planning. However, several unresolved problems must be researched before we can endorse any of these indexes.

5. The indexes vary greatly in their coverage and definitions of domains of QOL. We recommend a domain structure that all indexes can use to improve communication and validation among researchers.
6. None of the indexes distinguishes among the concepts of *input, throughput, and output* that are used by public policy analysts.
7. More studies of *predictive validity* (sensitivity) are needed to map the causal paths from public policy (input variables)

through citizen choice (throughput) to output (SWB, survival, and contribution to human heritage).

8. None of the indexes tests full *convergent validity* against any other QOL index. Hence researchers do not yet know which indexes are redundant with others (correlate highly over time and countries), and which explain new variance in QOL.

We strongly encourage researchers to take up these deficiencies in QOL indexes, to make them more relevant and useful for public policy.

REFERENCES

Allardt, E.: 1978, 'The relationship between objective and subjective measures in the light of a comparative study', in R. F. Tomassen (ed.),Comparative Studies in Sociology (Greenwich, CT: JAI Press), p. 1.

Andresen, E. M.: 1999, 'Population measurement of Health-Related Quality-of-Life (HRQoL)', Quality of Life Newsletter, Jan. 21.

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

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

Argyle, M.: 1996, 'Subjective wellbeing', in A. Offer (ed.), In Pursuit of the Quality of Life (New York: Oxford University Press), pp. 18–45.

Arroyo, D. M.: 1989a, 'Crime patterns and public safety', Social Weather Bulletin 89-9 (Quezon City: Social Weather Stations, Inc.).

Arroyo, D. M.: 1989b, 'Self-rated poverty and poverty thresholds,' Social Weather Bulletin 89-3 (Quezon City: Social Weather Stations, Inc.).

Arroyo, D. M.: 1990a, 'Monitoring changes in the Filipino's standard of living: Gains and losses, optimism and pessimism among broad groups', Social Weather Bulletin 90-13 (Quezon City: Social Weather Stations, Inc.).

Arroyo, D. M.: 1990b, 'Poverty means more than low incomes', Social Weather Bulletin 90-5 (Quezon City: Social Weather Stations, Inc.).

Bardo, J. W., S. J. Yeager and M. J. Klingsporn: 1982, 'Preliminary assessment of format-specific central tendency and leniency error in summated rating scales', Perceptual and Motor Skills 54, pp. 227–234.

Boelhouwer, J. and I. Stoop: 1999, 'Measuring well-being in the Netherlands: The SCP index from 1974 to 1997', Social Indicators Research 48, pp. 51–75.

Bollen, K. A.: 1989, Structural Equations with Latent Variables (New York: John Wiley & Sons).

Boyle, G. J.: 1991, 'Does item homogeneity indicate internal consistency or item redundancy in psychometric scales?', Personality and Individual Differences 12, pp. 291–294.

Brief, A. P. and L. Roberson: 1989, 'Job attitude organization: An exploratory study', Journal of Applied Social Psychology 19, pp. 717–727.

Broe, G. A., A. F. Jorm, H. Creasey, D. Grayson, D. Edelbrock, L. M. Waite, H. Bennett, J.S. Cullen and B. Casey: 1998, 'Impact of chronic systemic and neurological disorders on disability, depression and life satisfaction', International Journal of Geriatric Psychiatry 13, pp. 667–673.

Bufford, R. K., R. F. Paloutzian and C. W. Ellison: 1991, 'Norms for the spiritual well-being scale', Journal of Psychology and Theology 19(1), 56–70.

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

Centers for Disease Control and Prevention: 1998, Self-reported Frequent Mental Distress Among Adults (United States, 1993–1996) 47, pp. 325–331.

Chapin, F. S.: 1935, The Measurement of Social Status by Use of a Social Status Scale 1933 (Minneapolis: The University of Minnesota Press).

Cobb, C., T. Halstead and J. Rowe: 1995, The Genuine Progress Indicator: Summary of Data and Methodology (San Francisco, CA: Redefining Progress).

Crooker, K. J. and J. P. Near: 1995, 'Happiness and satisfaction: Measures of affect and cognition?', in H. L. Meadow, M. J. Sirgy and D. R. Rahtz (eds.), Developments in Quality-of-Life Studies in Marketing (DeKalb, IL: Academy of Marketing Science and the International Society for Quality-of-Life Studies 5), pp. 160–166.

Cummins, R. A.: 1995, 'On the trail of the gold standard for life satisfaction', Social Indicators Research 35, pp. 179–200.

Cummins, R. A.: 1996, 'The domains of life satisfaction: An attempt to order chaos', Social Indicators Research 38, pp. 303–328.

Cummins, R. A.: 1997a, 'Comprehensive Quality of Life Scale', Adult Manual, Fifth Edition. (Melbourne: School of Psychology, Deakin University), pp. 1–51.

Cummins, R. A.: 1997b, 'Assessing Quality of Life', in R. Brown (ed.), Quality of Life for People with Disabilities (Cheltenham: Stanley Thornes), pp. 116–150.

Cummins, R. A.: 1998a, 'The second approximation to an international standard for life satisfaction', Social Indicators Research 43, pp. 307–334.

Cummins, R. A.: 1998b, Directory of Instruments to Measure Quality of Life and Cognate Areas, Fourth Edition (Melbourne: School of Psychology, Deakin University).

Cummins, R. A.: 1999, 'A psychometric evaluation of the comprehensive Quality of Life Scale, Fifth Edition', in L. L. Yuan, B. Yuen and C. Low (eds.), Urban Quality of Life: Critical Issues and Options (Singapore: University Press), pp. 20–33.

Cummins, R. A.: 2000, 'Objective and subjective quality of life: An interactive model', Social Indicators Research 52, pp. 55–72.

Cummins, R. A., Ross Andelman, Richard Board, Laura lee Carman, Abbott Ferriss, Phil Friedman, Alex Michalos, Josh Samli, Stan Shapiro, Joe Sirgy, Joar Vitterso and Ruut Veenhoven: 1998, 'Quality of life definition and terminology: A discussion document from the International Society for Quality of Life Studies', The International Society for Quality-of-Life Studies (ISQOLS).

Cummins, R. A., M. P. McCabe, Y. Romeo and E. Gullone: 1994, 'The Comprehensive Quality of Life Scale (ComQol): Instrument development and psychometric evaluation on college staff and students', Educational and Psychological Measurement 54(2), pp. 372–382.

Diener, E.: 1994, 'Assessing subjective well-being: Progress and opportunities', Social Indicators Research 31, pp. 103–157.

Diener, E.: 1995, 'A value based index for measuring national quality of life', Social Indicators Research 36, pp. 107–127.

Diener, E. and C. Diener: 1995, 'Income and the quality of life of nations', Social Indicators Research 36, pp. 275–286.

Diener, E. D., R. A. Emmons, R. J. Larsen and S. Griffin: 1985, 'The satisfaction with life scale', Journal of Personality Assessment 49, pp. 71–75.

Diener, E., J. Horwitz and R. A. Emmons: 1985a, 'Happiness of the very wealthy', Social Indicators Research 16, pp. 263–274.

Diener, E., R. J. Larsen, S. Levine and R. A. Emmons: 1985b, 'Intensity and frequency: Dimensions underlying positive and negative affect', Journal of Personality and Social Psychology 48, pp. 1253–1265.

Diener, E., E. Sandvik, L. Seidlitz and M. Diener: 1993, 'The relationship between income and subjective well-being: Relative or absolute?' Social Indicators Research, 28, pp. 195–223.

D'Iribarne, P.: 1974, 'The relationship between subjective and objective well being', in B. Strumpel (ed.), Subjective Elements of Well Being (Paris: OECD).

Eckersley, R.: 1999, 'Rising youth suicide in Western nations: An indicator of social decline, or improvement?' (submitted for publication).

Estes, R. J.: 1984, The Social Progress of Nations (New York: Praeger).

Estes, R. J.: 1988, Trends in World Social Development (New York: Praeger).

Estes, R. J.: 1995, 'Social development trends in Africa', Social Development Issues 17(1), pp. 18–47.

Estes, R. J.: 1996a, 'Social development trends in Asia', Social Indicators Research 37(2), pp. 119–148.

Estes, R. J.: 1996b, 'Social development trends in central and South America', Social Development Issues 18(1), pp. 25–52.

Estes, R. J.: 1997, 'Social development trends in Europe, 1970–1994: Development prospects for the new Europe', Social Indicators Research 42(1), pp. 1–19.

Estes, R. J.: 1998, 'Social development trends in transitional economies, 1970–1995', in Ronald Hope Kempe, Sr. (ed.), Challenges of Transformation and Transition from Centrally Planned to Market Economies (United Nations: United Nations Centre for Regional Development), pp. 13–30.

Estes, R. J.: 1988, Trends in World Social Development: The Social Progress of Nations, 1970–1987 (New York: Praeger Publishers).

Euro-barometer: 1998, 'EuroBarometer', Brussels: Commission of the European Communities, Online access: http://europa.eu.int/en/comm/dg10/infcom/epo/eb.html.

Evans, D. R.: 1994, 'Enhancing quality of life in the population at large', Social Indicators Research 33, pp. 47–88.

Ferriss, A.: 2000, 'A brief history of social indicator developments' (unpublished manuscript).

Firat, A. F. and Karafakioglu: 1990, 'Quality of life issues and marketing as a social change agent', in H. L. Meadow and M. J. Sirgy (eds.), Quality-of-Life Studies in Marketing and Management (Blacksburg, VA: Virginia Tech, International Society for Quality-of-Life Studies), pp. 623–632.

Fishbein, M. and I. Ajzen: 1975, Belief, Attitude, Intention and Behavior: An Introduction to Theory and Research (Reading, MA: Addison-Wesley).

Fordham Institute for Innovation in Social Policy: 1999, 1998, Index of Social Health (Tarrytown, NY: Fordham Graduate Center).

Gertner, J. and R. Kirwan: 1999, 'The best places to live', Money (November), pp. 130–138.

Green, P.: 1984, 'Hybrid models for conjoint analysis: An exploratory review', Journal of Marketing Research 21, pp. 155–169.

Guterbock, T. M.: 1997, 'Why *Money Magazine's* "best places" keep changing', Public Opinion Quarterly 61, pp. 339–355.

Hagerty, M. R.: 1999a, 'Unifying livability and comparison theory: Cross-national time-series analysis of life-satisfaction,' Social Indicators Research 46 (June).

Hagerty, M. R.: 1999b, 'Comment on the construction of the fordham index of social health', Social Indicators Network News 57, p. 10.

Hagerty, M. R.: 2000, 'Social comparisons of income in one's community: Evidence from national surveys of income and happiness', Journal of Personality and Social Psychology (April) 78(4), pp. 764–771.

Hagerty, M. R., P. Naik and C. L. Tsai: 2000, 'The effects of quality of life on national elections: A multicountry analysis', Social Indicators Research (Feb.) 49(2).

Hair, J. F. Jr., R. E. Anderson, R. L. Tatham and W. C. Black: 1998, Multivariate Data Analysis, Fifth Edition (Upper Saddle River, NJ: Prentice Hall).

Henderson, H., J. Lickerman and P. Flynn: 2000, Calvert-Henderson Quality of Life Indicators: A New Tool for Assessing National Trends (Bethesda, MD: Calvert Group).

Hennessy, C. H., D. G. Moriarty, M. M. Zack, P.A. Scherr and R. Brackbill: 1994, 'Measuring health-related quality of life for public health surveillance', Public Health Reports 109, pp. 665–672.

Hoos, I. R.: 1972, Systems Analysis in Public Policy; a Critique (Berkeley, University of California Press).

Johnston, D. F.: 1988, 'Toward a comprehensive "quality of life" index', Social Indicators Research 20, pp. 473–496.
Joreskog, K. and D. Sorbom: 1988, LISREL 7: A Guide to the Program and Applications, Second Edition (Chicago, IL: SPSS Inc.).
Kacapyr, E.: 1996, '*Economic Forecasting*, M. E. Sharpe'.
Kacapyr, E.: 1997, 'Are we having fun yet?' American Demographics (October), pp. 28–29.
Kotler, P., D. Haider and I. Rein: 1993, Marketing Places (New York: The Free Press).
Kuhfeld, W. F., R. D. Tobias and M. Garratt: 1994., 'Efficient experimental designs with marketing research applications', Journal of Marketing Research 31, pp. 547–557.
Kuyken, W., J. Orley, P. Hudelson and N. Sartorius: 1994, 'Quality of life assessment across cultures', International Journal of Mental Health 23, pp. 5–28.
Land, K.: 2000, 'Social indicators', in Edgar F. Borgatta and Rhonda V. Montgomery (eds.), Encyclopedia of Sociology, Revised Edition (New York: Macmillan, forthcoming).
Lane, R. E.: 1991, The Market Experience (Cambridge, England: Cambridge University Press).
Luchters, G. and L. Menkhoff: 1994, 'The fourth premiere of the Human Development Index,' Intereconomics, January/February.
Malhotra, N. K., M. Peterson and S. B. Kleiser: 1998, 'Assessing the reliability and validity of international secondary data used to design competitive strategies for global marketing', in Manrai A and L. Manrai (eds.), Research in Marketing: Designing Competitive Strategies for Global Marketing (Stamford, CT: JAI Press) 14, pp. 185–228.
Mangahas, M.: 1976, Measuring Philippine Development (ed.) (Manila: Development Academy of the Philippines).
Mangahas, M.: 1986, 'Dimensions of poverty', SWS Occasional Paper (Quezon City: Social Weather Stations, Inc.).
Mangahas, M.: 1987, 'Rural poverty, social surveillance and democratic development', SWS Occasional Paper (Quezon City: Social Weather Stations, Inc.).
Mangahas, M.: 1995, 'Self-rated poverty in the Philippines, 1984–1992', International Journal of Public Opinion Research 7/1, pp. 40–55.
Mangahas, M.: 1996, 'Review of surveys on crime', Social Weather Bulletin 96-11/12 (June) (Quezon City: Social Weather Stations, Inc.).
Mangahas, M. and L. L. Guerrero: 1998, 'Self-sustained quality of life monitoring: The Philippine social weather reports', SWS Occasional Paper (Quezon City: Social Weather Stations, Inc.).
Maslow, A: 1970, Motivation and Personality (Second Edition) (New York: Harper).
McKennell, A. C.: 1978, 'Cognition and affect in perceptions of well being', Social Indicators Research 5, pp. 389–426.

McKennel, A. C. and F. M. Andrews: 1980, 'Models of cognition and affect in perceptions of well being', Social Indicators Research 8, pp. 257–298.
Michalos, A. C.: 1980–82, North American Social Report 1-5, D. Reidel.
Michalos, A. C.: 1985, 'Multiple Discrepancies Theory (MDT)', Social Indicators Research 6(1), pp. 347–413.
Michalos, A. C.: 1997, 'Combining social, economic and environmental indicators to measure sustainable human well-being', Social Indicators Research 40, pp. 221–258.
Michalos, A. C.: 1991, Global Report on Student Well-being 1-4 (New York: Springer-Verlag).
Michalos, A. C.: 1995, A Pragmatic Approach to Business Ethics (Sage Publications, Thousand Oaks).
Miringoff, M. L. and M. L. Miringoff: 1999, The Social Health of the Nation: How America is Really Doing (New York: Oxford Univ. Press).
Miringoff, M. L., M. L. Miringoff and S. Opdycke: 1996, 'Monitoring the nation's social performance: The Index of Social Health', in E. F. Zigler, S. L. Kagan and Nancy W. Hall (eds.), Children, Families, and Government (New York: Cambridge Univ. Press), pp. 10–30.
Moffitt, R.: 1998, Current Population Surveys: March Individual-Level Extracts, 1968–1992 (2nd ICPSR version) [Computer file]. Madison, WI: University of Wisconsin, Institute for Research on Poverty [Producer]. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [Distributor].
Moller, V.: 1997, Quality of Life in South Africa (Kluwer).
Morrow, R. S. and S. Morrow: 1973, 'The measurement of intelligence', in Benjamin Wolman (ed.), Handbook of General Psychology (Englewood Cliffs, NJ: Prentice-Hall, Inc.).
Murray, C. J. L.: 1991, 'Rational approaches to priority setting in international health', Journal of Tropical Medicine and Hygiene 93.
Noll, H. H. and W. Zapf: 1994, 'Social indicators research: Societal monitoring and social reporting', in I. Borg and P. Mohler (eds.), Trends and Perspectives in Empirical Social Research. (Walter de Gruyter, Berlin/New York), pp. 1–16.
Noll, H. H. and E. Wiegand: 1993, 'System Sozialer Indikatoren für die Bundesrepublik Deutschland-Zeitreihen 1950–1991', Tabellenband. ZUMA (Mannheim).
Nubler, I.: 1995, 'The Human Development Index revisited', Intereconomics (July/August), pp. 171–176.
O'Boyle, C. A.: 1994, 'The Schedule for the Evaluation of Individual Quality of Life (SEIQoL)', International Journal of Mental Health 23, pp. 3–23.
OECD: 1982, 'The OECD list of social indicators', The OECD Social Indicator Development Programme (Paris).
Organ, D. W. and J. P. Near: 1985, 'Cognition vs. affect in measures of job satisfaction', International Journal of Psychology 20, pp. 241–253.
Osberg, L. and A. Sharpe: 1998, 'An index of economic well-being for Canada', Research Paper R-99-3E, Applied Research Branch, Human Resources Development Canada, Ottawa, Ontario.

Osberg, L. and A. Sharpe: 1999, 'An index of economic well-being for Canada and the United States', paper presented at the annual meeting of the American Economic Association, New York, New York, January.

Osberg, L. and A. Sharpe: 2000, 'International comparisons of trends in economic well-being', paper presented at the annual meeting of the American Economic Association, Boston, Mass., January.

Paloutzian, R. F. and C. W. Ellison: 1982, 'Loneliness, spiritual well-being and the quality of life', In L. A. Peplau and D. Perlman (eds.), Loneliness: A Sourcebook of Current Theory, Research and Therapy (New York: John Wiley and Sons) pp. 224–237.

Peterson, M. and N. K. Malhotra: 1997, 'Comparative marketing measures of societal quality of life: Substantive dimensions in 186 countries', Journal of Macromarketing 17(1), pp. 25–38.

Peterson, M., N. K. Malhotra and J. Wagner: 1999, 'Country quality of life and foreign direct investment decisions', Global Outlook 11, pp. 1, 51–62.

Pirfo, E., C. Albera, Il. Arizio, S. Catapano, M. Cortese and C. Romano: 1994, 'Job preparation and improvement of the quality of life of schizophrenic patients in today's metropolis', International Journal of Mental Health 23, pp. 11–22.

Rosencranz, H. A. and C. T. Pihlblad: 1970, 'Measuring the health of the elderly', Journal of Gerontology 25, pp. 129–133.

Ryff, C. D.: 1995, 'Psychological well-being in adult life', American Psychological Society 4, pp. 99–104.

Samli, A. C.: 1995, 'QOL research and measurement: some parameters and dimensions', in H. L. Meadow, M. J. Sirgy and D. Rahtz (eds.), Developments in quality-of-life studies in marketing (DeKalb, IL: Academy of Marketing Science and the International Society for Quality-of-Life Studies) 5, pp. 173–180.

Sandoval, G. A. and P. R. Laylo: 1993, Crime Victimization in the Philippines: Demographic Patterns and Attitudinal Correlates (Quezon City: Social Weather Stations, Inc.).

Schaefer, M. T. and D. H. Olson: 1981, 'Assessing intimacy: The PAIR inventory', Journal of Marital and Family Therapy 24, pp. 47–60.

Schwartz, S. H.: 1992, 'Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries', in M. Zanna (ed.), Advances in Experimental Social Psychology 20 (Orlando, Florida: Academic).

Sewell, W. H.: 1940, 'The constructin and standardization of a scale for the measurement of the socio-economic status of Oklahoma farm families', Technical Bulletin No. 9 (Stillwater, OK: Oklahoma Agricultural Experiment Station).

Sirgy, M. J., A. C. Samli and H. L. Meadow: 1982, 'The interface between quality of life and marketing: a theoretical framework', Journal of Public Policy and Marketing 1, pp. 69–84.

Skevington, S. M., J. Bradshaw and S. Saxena: 1999, 'Selecting national items for the WHOQOL: conceptual and psychometric considerations', Social Science and Medicine 48, pp. 473–487.

Szabo, S.: 1996, 'The World Health Organization Quality of Life (WHOQOL) assessment instrument', in B. Spilker (ed.), Quality of life and pharmacoeconomics in clinical trials, Second Edition (Philadelphia: Lippincox-Raven Publishers), pp. 355–362.

Trabold-Nubler, H.: 1991, 'The human development index: A new development indicator?', Intereconomics (Sept./Oct.), pp. 236–243.

United Nations Human Development Program (UNHDP): 1990, Human Development Report (N.Y. and Oxford: Oxford University Press).

United Nations Human Development Program: 1993, 'Human Development Index: Survey of recent reviews', Human Development Report (N.Y. and Oxford: Oxford University Press).

United Nations Human Development Program: 1996, 'Human Development Index: Survey of recent reviews', Human Development Report (N.Y. and Oxford: Oxford University Press).

University of Virginia Center for Public Service: 1996a, Virginia's Local Economies: Statistical Update, 1996 (Charlottesville, VA: University of Virginia Center for Public Service, University of Virginia).

University of Virginia Center for Public Service: 1996b, Virginia Statistical Abstract, 1996–1997 (Charlottesville, VA: University of Virginia Center for Public Service, University of Virginia).

Van Dam, F. S. A. M., R. Somers and A. L. Van Beek-Couzijn: 1981, 'Quality of life: Some theoretical issues', Journal of Clinical Pharmacology 21, pp. 166–168.

Veenhoven, R.: 1984, Conditions of Happiness (Dordrecht: Reidel).

Veenhoven, R.: 1993, Happiness in Nations, Rotterdam: Risbo. Data available from web-site: http://www.eur.nl/fsw/research/happiness.

Veenhoven, R.: 1994, 'Is happiness a trait?', Tests of the theory that a better society does not make people any happier, Social Indicators Research 32, pp. 101–160.

Veenhoven, R.: 1995, 'The cross national pattern of happiness. Tests of predictions implied in three theories of happiness', Social Indicators Research 34, pp. 33–68.

Veenhoven, R.: 1996, 'Happy life expectancy: A comprehensive measure of quality of life in nations', Social Indicators Research 39, pp. 1–58.

Veenhoven, R.: 1998, 'Quality-of-life and happiness: Not quite the same', in G. DeGirolamo, et al. (eds.), Health and Quality-of-Life (Il Pensierro Scientifico, Rome).

Veenhoven, R. and P. Ouweneel: 1995, 'Livability of the welfare-state. Appreciation-of-life and length-of-life in nations varying in state-welfare-effort', Social Indicators Research 36, pp. 1–49.

Viney, L. L., B. M. Walker, T. Robertson, B. Lilley and C. Ewans: 1994, 'Dying in palliative care units in hospital: A comparison of the quality of life of terminal cancer patients', Journal of Consulting and Clinical Psychology 62, pp. 157–160.

Virginia Tech Center for Survey Research: 1998, Quality of Life in Virginia: 1998 (Blacksburg, VA: Center for Survey Research, Virginia Polytechnic Institute and State University).

Vogel, J.: 1991, 'Social report for the nordic countries', Living Conditions and Inequality in the late 1980s (Copenhagen).

Vogel, J.: 1993, 'Wohlfahrtssurveys und Sozialberichterstattung in Nordeuropa, Systematische Dauerbeobachtung der objektiven Lebensverhältnisse', in W. Glatzer (ed.), Einstellungen und Lebensbedingungen in Europa. Campus.

Vogel, J.: 1995a, The Social Indicator Tradition (Luxembourg: Eurostat, Doc E2/IS/2/95).

Vogel, J.: 1995b, Towards an Integrated Household Survey Programme. The Swedish Experience (Luxembourg: Eurostat, Doc E2/IS/2/95).

Vogel, J.: 1997a, 'Living conditions and inequality in the European Union', Eurostat: Population and Social Conditions (Luxembourg: E/1997-3).

Vogel, J.: 1997b, 'The European welfare mix: institutional configuration and distributive outcome in Sweden and the European Union. A longitudinal and comparative perspective', paper presented at the 1997 conference of the International Society for Quality of Life studies (Charlotte, NC, USA, 20–23.11.1997).

Vogel, J.: 1998a, 'Welfare surveys and social reporting in the nordic countries', I Statistik im Dienste der Öffentlichkeit. Bundesamt fuer Statistik (Bern, Schweiz).

Vogel, J.: 1998b, 'Coping with the European welfare mix. Welfare delivery systems, family formation and material inequality between types of families in the European Union', paper presented at the 1998 Siena Group Meeting on Families at the end of the Twentieth Century (Sidney, Australia, 7-9.12.1998).

Vogel, J. and L. Häll: 1997, 'Välfärd och ojämlikhet i 20-årsperspektiv 1975–95 living conditions and inequality 1975–95', Report 91, series Living Conditions, Statistics (Sweden).

Vogel, J., H. Sauli and A. Andersen: 1993, Welfare Surveys and Social Reporting in the Nordic Countries. Experiences and Potential (Eurostat: doc E3/IS/5/93).

The WHOQOL Group: 1993, WHOQOL Study Protocol (Geneva: WHO, Division of Mental Health).

The WHOQOL Group: 1995a, WHOQOL-100: Domains and Facets questionnaire (Geneva: WHO, Division of Mental Health).

The WHOQOL Group: 1995b, WHOQOL-100: The 100 questions with response scales (Geneva: WHO, Division of Mental Health).

The WHOQOL Group: 1995c, WHOQOL-100: Facet definitions and questions (Geneva: WHO, Division of Mental Health).

The WHOQOL Group: 1995d, WHOQOL-100: Australian version (Geneva: WHO, Division of Mental Health).

The WHOQOL Group: 1995e, 'The World Health Organization quality of life assessment (WHOQOL): Position paper from the World Health Organization', Social Science and Medicine 41, pp. 1403–1409.

The WHOQOL Group: 1996a, WHOQOL-BREF: Field trial version, Questionnaire (Geneva: WHO, Division of Mental Health).

The WHOQOL Group: 1996b, WHOQOL-BREF: Field trial version: Introduction, administration, scoring and general version of the assessment (Geneva: WHO, Division of Mental Health).
The WHOQOL Group: 1998a, 'The World Health Organization quality of life assessment (WHOQOL): Development and general psychometric properties', Social Science and Medicine 46, pp. 1569–1585.
The WHOQOL Group: 1998b, 'Development of the World Health Organization WHOQOL-BREF quality of life assessment', Psychological Medicine 28, pp. 551–558.
Zagorski, K. and J. S. McDonnell: 1995, 'Consumer confidence' Indexes as Social Indicators', Social Indicators Research 36, pp. 227–246.
Zapf, W.: 1977, 'Soziale Indikatoren – eine Zwischenbilanz', in H.-J. Krupp and W. Zapf (eds.), Sozialpolitik und Sozialberichterstattung (Campus Verlag, Frankfurt a.M./New York), pp. 231–246.

Graduate School of Management
University of California
Davis, CA 95616
USA
E-mail: mrhagerty@ucdavis.edu