
Issues in Composite Index Construction: The Measurement of Overall Quality of Life

9

Michael R. Hagerty and Kenneth C. Land

“Are you better off now than 4 years ago?” Asked by numerous politicians since Ronald Reagan famously posed it as a candidate in the 1980 Presidential Election in the United States, this question invites voters to make a complex and perilous series of judgments to estimate *overall quality of life (QOL)* in their country. Each citizen must, at least informally, engage in the following activities: (1) Select the indicators that are most important to her or him, (2) obtain data from social reports or other news sources on the progress of those indicators, and (3) integrate those indicators across disparate domains to achieve a judgment of overall progress on QOL.

It is clear that science can help citizens with the first two tasks by collecting reliable and valid indicators related to QOL and by disseminating those indicators widely in social reports to facilitate citizens’ judgments. But for some very good reasons, social scientists have been reluctant to help in the third task of summarizing those indicators into a composite QOL index (for brevity below, the term “QOL index” often is used to reference a “composite QOL index”). These reasons are: First, constructing a QOL index requires “comparing apples to oranges” because the indicators have no common unit (how does one combine longevity measured in years with income measured in dollars?); second, it requires knowledge of how each

citizen selects and weights indicators to arrive at their overall judgment; and third, it requires that citizens are sufficiently homogenous that a single QOL index would be accepted by a majority of citizens, because it approximates their own individual QOL judgments. And underlying these concerns is a worry that constructing a QOL index could become “politicized” or manipulated for short-term political gain at the expense of long-term scientific credibility.

This chapter outlines the progress that social science research has made on these questions in the last 50 years, and proposes some principles for developing QOL indices that help assure acceptance by citizens and resistance to politicization. We begin by reviewing three composite indices from economics that have achieved these goals in the United States (similar indices have been developed in many other countries). We then describe seven principles to guide the developments of QOL indices. This is followed by a review of 14 existing QOL indices which we evaluate with respect to these principles. We then state several common criticisms of composite indices and solutions thereto. The chapter concludes with several recommendations with respect to the construction of QOL indices.

Three Composite Indices of the Economy

The use of composite indices of various aspects of economic activity and conditions has a long history from which a number of lessons can be learned. We review three economic indices: the Dow-Jones Industrial Average, the Consumer Price Index, and the Consumer Sentiment Index.

M.R. Hagerty
University of California, Graduate School of Management,
Davis, CA, USA
e-mail: mrhagerty@ucdavis.edu

K.C. Land (✉)
Department of Sociology, Duke University, Durham, NC, USA
e-mail: kland@soc.duke.edu

The Dow-Jones Industrial Average (DJIA). The Dow-Jones Industrial Average is probably the index most quoted by print and electronic media in the USA. It was created in 1882 for investors who wanted to “see the forest instead of the trees” as a simple way to gauge overall movement in the New York Stock Exchange. It originally included only nine railroad stocks, a steamship line, and a communications company. Today, it includes 30 large “industrial” companies (including WalMart). The bundle of stocks is picked *subjectively* by editors of the Wall Street Journal, and is modified periodically to “reflect the current economy.” Originally, the calculation was a simple average of the prices of the stocks, divided by their number. Today, the editors have modified the simple average to a “price-weighted” average, with adjustments for prices when a stock “splits” (e.g., a two-for-one split in which owners of 100 shares have 200 shares after the split). Among all stock market indices, the DJIA is the least representative and uses the simplest calculation. Despite these weaknesses, it is nevertheless the most frequently quoted and most easily understood.

The Consumer Price Index (CPI-U). The Consumer Price Index-Urban is published by the Bureau of Labor Statistics as an estimate of the cost of living in urban areas in the USA. It is probably the second most-often quoted index by the media in the USA, not because it is less important than the DJIA, but because it is updated over a month rather than every day. The BLS monitors prices for 211 items in 38 geographic areas, and maintains 8,018 “disaggregated” price indices, including, for example, the price of apples in Chicago. The creation of an index would seem straightforward, since all items are measured in the common unit “dollars.” But despite the common metric, the *weights* for each item must be determined. The proper weight should reflect the fact that apples contribute far less than household rent in calculating CPI, because a “representative” family allocates far more of their budget to rent than to apples. The natural weight would be the proportion of the average family’s budget that is devoted to that item in that month. The problem is that full-scale surveys of family budgeting and “representative consumer baskets of goods” occur only every 2–5 years. Hence the “constant weights” must be modified by estimates to predict how families will allocate their purchases every month. The index must be shown to be robust to errors in estimating how families allocate their purchases

each month, and considerable research has investigated the properties of various indices when faced with real consumer decisions. This research has focused on two areas: the problem of how consumers substitute purchases when the *price* of one item rises and the problem of how consumers substitute purchases when the *quality* of one item rises.

An example of price substitution is that consumers will purchase less beef when its price increases and will instead purchase more of substitutes, such as chicken. Hence, a simple constant-weight average index overstates the true cost of living. Research has shown that geometric averages/means are more robust to substitution errors than arithmetic averages/means, so the CPI-U now uses geometric averages¹ of prices in each of the 8,018 price series. The CPI-U however still uses fixed weights based on surveys that are at least 2 years old. To remedy this, the BLS now offers a “chained” CPI-U that updates weights dynamically each month using the most recent batch of surveys.

It is worth noting that the CPI is a weighted average, where the weights have been constant for long periods. Only recently has the CPI made incremental improvements by estimating dynamically changing weights. The QOL indices that we review later are similarly of this general form: weighted averages with constant weights. It is also worth noting that CPI research has not allowed “the perfect to be enemy of the good,” but has published indices for over 50 years that had known errors, because constant weights are a first approximation to dynamic weights. Finally, the CPI research has sought to compare proposed indices with how individuals actually behave, and whether the aggregate index successfully predicted families’ substitution behavior. A similar approach with respect to QOL indices will be described below.

The Consumer Confidence Index (CCI). The CCI was introduced in 1952 by the University of Michigan’s Survey Research Center. It consists of a five-question battery in monthly consumer surveys. The questions are:

- “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?”

¹The geometric mean \tilde{X}_n of n positive numbers X_1, X_2, \dots, X_n is the positive n th root of their products, that is, $\tilde{X}_n = \left(\prod_{i=1}^n X_i\right)^{1/n}$.

- “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?”
- “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?”
- “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?”
- “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 CCI is calculated in the following way: For each of the five questions which comprise the index, the proportion of unfavorable responses is 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 with equal weights to form the Index of Consumer Sentiment.

These three economic indices draw on over 100 years of experience and research. In the next section, we generalize from this research and over 30 years of QOL studies to formulate desirable principles for constructing QOL indices. Later sections will apply these criteria to existing QOL indices.

Principles for Constructing a Composite QOL Index

We state *seven principles for constructing QOL indices* and then consider their implications. Some are well known, but some are relatively new.²

The first principle is that each of the subseries that compose an index should be reliable and valid. This criterion is well known, and a review of QOL indices by Hagerty et al. (2001) concludes that most

social reports can now boast many reliable and valid indicators. In the case of the DJIA, validation implies that it must correlate with the overall movements of the New York Stock Exchange and with gross domestic product 6 months in the future. In the case of QOL indices, the subseries could be validated to assure that they correlate with global measures of QOL, such as surveys of citizens’ average happiness, frequency of smiling, lack of revolutionary or separatist political movements, and eventually, with brain imaging that displays positive emotion.

Second, to improve transparency, the QOL index should not be reported alone, but as part of a report that shows each underlying subseries. For example, the subseries of the CPI are reported at the same time as the CPI itself, and many users calculate an alternate “core” CPI by deleting the more volatile food and fuel series, because previous research shows that the core CPI is more stable and is a better predictor of next month’s CPI.

Third, the QOL report should disaggregate the index for population subgroups. The CPI-U is calculated for all urban dwellers (the best known CPI series), but it is also calculated for rural dwellers and wage earners. This is likewise important in QOL reports, because informed citizens, policy decisions, and government programs require knowledge of whether some groups (e.g., the elderly, children, minorities, immigrants) are disadvantaged and may require help.

Fourth, an index should be robust to incomplete data or other data problems. In the CPI, research has shown that the chained CPI is a robust index even when updated surveys of family purchases are not available. In a QOL index, some series may be available monthly (e.g., earnings per family), but others are available only yearly (e.g., inequality), and each is reported with varying accuracy. Research to determine the robustness of a QOL index in these situations is warranted.

Fifth, the index should reflect the best model of how people actually make QOL judgments for themselves. Among economic indices, the CCI assumes that people can form judgments about their likelihood of earning and spending more money next year, and that a simple average of these perceptions predicts families’ future purchasing. In the case of QOL indices, we can rely on research over the last

²These principles are consistent with guidelines for constructing composite indicators of all types such as those summarized in Nardo et al. (2005), but are focused on QOL measurement and its unique, substantive features.

30 years into how individuals make their personal judgments of QOL, described later.

Sixth, the index should reflect the weights that citizens give to individual subseries. The CPI achieves this by national surveys of families and the proportion of their budget spent on each category. In the case of a QOL index, if citizens tend to place high importance on the health domain and only moderate importance on inequality, then a composite QOL index should reflect this, with a unit change in (standardized) health causing a larger change in the composite index than a unit change in (standardized) inequality.

Seventh, an index should be accepted by a large majority of citizens. By accepted, we mean that most citizens trust it and endorse its use by political decision makers, because the index is a good approximation to the QOL judgments that the citizens themselves would make. In the case of the DJIA, vast numbers of investors show acceptance by using it to make individual buy/sell decisions daily (even though it is known to represent only a few large stocks and is an imperfect predictor of future activity), and the Bureau of Economic Research uses such indices to predict future economic activity. Despite its flaws, the DJIA shows acceptance by millions of decision makers.

The first four principles for constructing QOL indices are widely known and honored. In a review of extant QOL indices, Hagerty et al. (2001) proposed similar criteria and showed that many existing QOL indices conform to these goals. However, the last three criteria have been formalized fairly recently, in response to a call by Land (2004) for “evidence-based” QOL indices. These last three criteria (especially that the QOL index be accepted by a majority of citizens) have always been implicitly employed by past researchers, but only recently have the methods and measurements for predicting citizen acceptance been formalized. Hence we elaborate in more detail the last three principles for constructing QOL indices.

Principle 5: How people actually make QOL judgments for themselves. A long stream of research has concluded that a simple weighted average model predicts individuals’ overall judgments of their QOL from satisfaction with their individual domains (Campbell et al. 1976; Cummins 1996; Veenhoven 1993). Though in some studies the direction of causality is ambiguous (“top-down” models predict that higher overall affect

causes higher ratings of individual domains), researchers agree that if actual conditions in a domain improve, then the change in overall rating of QOL is well predicted by a linear (weighted) additive model (Lucas et al. 2003; Sastre 1999). Another caveat is that if the weights contain excessive error in measurement (e.g., if weights are measured at the individual level rather than aggregated over larger samples), then an equal-weight model will perform as well or better than a weighted model.

Acknowledging these caveats, we accept the weighted average model as a good description of citizens’ QOL judgment model, and define citizen i ’s importance weight as w_{ik} and citizen i ’s overall QOL judgment for country n as Q_{in} . Then we can predict their QOL judgments with the *weighted average model* (WAM):

$$Q_{in} = \sum_{k=1}^K w_{ik} x_{kn}, \quad w_{ik} > 0, \quad \text{for } n = 1, \dots, N \text{ countries,} \quad (9.1)$$

where x_{kn} is the score for the k th social indicator of country n , and K is the total number of social indicators that citizens use to make their judgments of QOL.³ Adopting this additive model also benefits the fourth principle in constructing QOL indices, since it is well known that additive models are quite robust to errors in measurement.

Note that the weighted average model of Eq. 9.1 can be viewed as a logarithmic transformation of the weighted product method for composite index construction studied by Munda and Nardo (2003; see also Nardo et al. 2005). Using notation similar to that of Eq. 9.1, the *weighted product* (WP) model can be written as

$$Q_{in} = \prod_{k=1}^K x_{kn}^{w_{ik}}, \quad w_{ik} > 0, \quad \text{for } n = 1, \dots, N \text{ countries.} \quad (9.2)$$

In a recent contribution to the literature on methods of composite index construction, Zhou et al. (2010) studied the WP model and proposed a multiplicative optimization extension thereof by application of data envelopment analysis (DEA)-type methods to determine

³As research progresses, this model may be modified to include substitutability or complementarity between social indicators that would require modeling *interactions* among indicators. For example, an individual with higher average income may consider life expectancy more important than an individual with very low income (as life becomes more “worth living,” longer life may be more valuable).

the values of weights of individual indicators in a composite index such as the life expectancy, education, and gross domestic product per capita indicators used to calculate the Human Development Index (described later in this chapter). The DEA method originally was developed for efficiency analysis in economics and management science (Charnes et al. 1978, 1994; Land et al. 1993). It transforms a multiplicative optimization problem into a series of linear programming problems (Dantzig 1963) in which weights for composite scores are determined by internal comparisons of each of a set of entities with each other with respect to their efficiency in producing outputs (e.g., consumer products) from given levels of inputs (e.g., labor, capital).

Zhou et al. (2010) applied DEA to calculate two sets of weights for the component indicators of a composite QOL index—a set of “best” weights for each entity calculated in comparison to the “best practice” entity or entities on each specific indicator and a set of “worst” weights calculated in comparison to the “worst practice” entity or entities on each specific indicator. They then calculate composite index scores for each entity being compared as weighted averages of logarithmic transformations of the two sets of weights, and, in the absence of “decision makers or analysts [having] no particular preference” (Zhou et al. 2010, p. 173) for one set of weights or the other, suggest equal weighting as a “fairly neutral choice.” Zhou et al. suggest that this extension of the WP method can provide an alternative to subjectively determined weights for composite indices. In an empirical application, Zhou et al. show that the ranks of most of 27 countries in the Asia and Pacific region given by the conventional Human Development Index remain unchanged when they are ranked by composite indices based on the multiplicative optimization method. This relatively new approach to the development of weights for composite indices merits additional analysis and study. For instance, given the logarithmic relationship between the models of Eqs. 9.1 and 9.2, it is entirely possible that citizens as well as decision makers and analysts use an informal version, or at least some approximation thereto, of the equal weighting of “best practice” (distance from the best performing unit(s)) and “worst practice” (distance from the worst performing unit(s)) relative rankings to arrive at composite index scores/summary judgments.

Principle 6: Citizens’ importance weights for sub-series. Given that we know *how* citizens form QOL

judgments, the next obvious question is *which* social indicators do citizens use to determine QOL. Fortunately, the answer has been found to be roughly consistent over 30 years and in over 30 studies reviewed by Cummins (1996). Table 9.1 displays some of these studies and gives the mean importance weights averaged from surveys of US citizens. Column (a) contains the domains of life and mean importances (weights of relative importance) from the pathbreaking study by Campbell et al. (1976). Consistent with later studies, they found that health tends to be rated highest life domain (area), followed by family life, extent of civil rights allowed by the national government, friendships, housing, job, community, and leisure activities. To address the concern that the “stated importance” of domains might differ from the “real” importance, Campbell et al. (1976) showed a close correlation between the stated importances in Table 9.1a and regression coefficients predicting stated QOL from life domains, demonstrating convergent validity for the weights in Table 9.1.

Table 9.1b contains the average importance weights from the US responses to an international online survey in 2005 of current readers of *The Economist* magazine. Respondents were asked to rate the importance of ten social indicators on a 5-point scale, where 5 denoted “Very important” and 1 denoted “Unimportant.” Finally, Table 9.1c contains mean importance weights for US citizens from the World Values Survey (WVS) (Inglehart 2000), which asks respondents in 50 countries to rate the importance of: family, friends, leisure time, politics, work, and religion. The exact wording to the questions in 1995 was, “Please say, for each of the following, how important it is in your life. Would you say xxx is very important (3), rather important (2), not very important (1), or not at all important (0)?” The scale is usually assumed to be equal interval, (hence the codes are equal interval), and the anchoring at “not at all important” may be assumed to represent a weight of near zero.

In summary, a fairly useful and predictive model of how people make their own QOL judgments using the weighted average model has been developed and validated empirically. And importance weights have been found to be relatively consistent in 30 years of surveys. According to the fifth and sixth principles, then, a QOL index should be a weighted average of the major domains, with weights approximating those in Table 9.1. The final principle uses this information to estimate acceptance of the QOL index by citizens.

Table 9.1 Mean ratings of importance of domains in USA from (a) Campbell, Converse, and Rodgers (CCR) (1976 Table 3–5), (b) The Economist Intelligence Unit (EIU) (2005), and (c) World Values Survey (WVS) (Inglehart 2000)

(a) CCR (1976)		(b) EIU (2005)		(c) WVS (1999)	
“Health”	3.63 ^a	“Your health”	4.68	–	
“Marriage”	3.56	– ^b		–	
“Family life”	3.54	“Family relations”	4.47	“Family”	2.94
“National government”	3.46	“Degree of political and civil liberty in your country”		“Politics”	1.68
		“Degree of social equality in your country”			
		“Degree of gender equality in your country”			
“Friendships”	2.92	–		“Friends”	2.65
“Housing”	2.90	“Material well-being”	3.50		
“Job”	2.81	“Job satisfaction”		“Work”	2.31
		“Job security”			
“Community”	2.79	“Social and community activities”	3.51	–	
“Religious faith”	2.65	–		“Religion”	2.37
“Nonwork activities”	2.21	–		“Leisure time”	2.29
“Financial situation”	2.06	–		–	
“Organizations”	.99	–		–	
N		994		1,502	

^aMean ratings from Campbell, Converse, and Rodgers were reversed by subtracting them from five so that higher ratings indicate higher importances, to be consistent with other studies in table

^bIndicates that domain was not rated in study

Principle 7: Assuring acceptance by citizens. Final acceptance of any index by the public is a complex process of demonstrating unbiasedness, credibility, and usefulness to citizens, together with extensive publicity. Moller and Dickow (2002) outline how this was achieved in South Africa during its democratic transition. While some of these factors are outside the control of social scientists, the properties of unbiasedness and usefulness can and should be built into a QOL index by adopting the following proposition as closely as possible: An index will be unbiased and useful if the index summarizes a large amount of data in a way that closely mimics the judgment of a citizen if she were to read the entire report and make her own judgment of QOL.

Hagerty and Land (2007) formalized this proposition by defining a quantitative measure of agreement between an index and a citizen i 's actual judgment of QOL. They considered several measures and recommend the simple correlation coefficient between the citizen i 's actual QOL judgments and the index's ratings. As this correlation increases, agreement between

the two increases, with maximum agreement yielding a correlation of +1. They denote this correlation as A_{Qi} , for the agreement (correlation) between a QOL index and person i 's actual judgments of QOL. Critical values of this measure are +.7 (the common requirement for reliability between two raters) and 0 (the point above which the QOL index at least agrees *in direction* with the individual's actual ratings). Hence, if agreement A_{Qi} is at least above zero, then the QOL index agrees in direction with the individual's ratings, and both would agree on whether “things are getting better or worse.”

Ideal data to calculate agreement would use surveys of citizens' actual judgments of various countries' QOL. Then the agreement A_{Qi} could be calculated as the simple correlation between the proposed index and each individual's actual QOL judgments. To our knowledge, such data do not yet exist (though they would be relatively easy to collect). However, Hagerty and Land use next-best data to calculate agreement with some real QOL indices—survey data on *importance weights* that citizens report.

The research then extrapolates citizens' QOL judgments using the linear model in Eq. 9.1 which is known to fit well.

Using this method, they calculated average agreement between the Human Development Index (which uses equal weights) and the actual weights surveyed from a sample of 1,502 US citizens in the World Values Survey (Inglehart et al. 2000). As noted above, the average weights from the WVS are given in Table 9.1c. Mean agreement between the HDI index ratings of QOL and the 1,502 individuals' ratings (predicted from their weights) was +.97 (standard error of estimate = .04).

This is remarkably high, and Hagerty and Land go on to probe why agreement should be so high even though the equal weighting in the HDI differs from the unequal weights that citizens report in the WVS. Using the weighted average model of QOL judgments of Eq. 9.1, they prove mathematically that several factors unexpectedly affect agreement for any index. Specifically, they show that agreement will be higher when (1) the index is based on cross-sectional data rather than time-series data, (2) the distribution of citizens' weights is unimodal rather than bimodal (as in abortion where conflict is much higher because weights are extreme and bimodal), (3) the distribution of citizens' weights is not negatively correlated across indices (people who highly value one indicator always place a very low value on another indicator), and (4) citizens' weights are all positive (or all negative) for each indicator. The HDI and the WVS conform to all four of these properties. Hence the agreement induced by the equal weight in HDI is quite high compared to the unequal weights that citizens report in the WVS.

Why should these four properties influence distortion so greatly? The first property states that cross-sectional indices (such as the HDI, Estes' Index of Social Progress, and Veenhoven's Happy Life Expectancy, all of which are described in the next section of the chapter) will show high agreement, regardless of differences in citizens' weights. The intuitive reason behind this is that all citizens are likely to agree (regardless of their importance weights) that Somalia currently has lower QOL than Canada. Hence, any citizen with positive weights (greater than zero and less than one) will create high agreement and high correlation with QOL ratings by the index. The technical

reason behind this is that agreement A can be written as a simple matrix product:

$$A_{Qi} = \mathbf{W}_Q^* \mathbf{R}_x \mathbf{W}_i^*, \quad (9.3)$$

where \mathbf{R}_x is the correlation matrix between the K social indicators, \mathbf{W}_Q^* are the weights (standardized) that the QOL index uses, and \mathbf{W}_i^* are the importance weights (standardized) applied by person i . Equation 9.3 shows that the correlation A_{Qi} between the index and any individual is a function *not only of the weights*, but also is moderated by the correlations among the social indicators \mathbf{R}_x . When the intercorrelations \mathbf{R}_x are high (as they are in the HDI and other cross-sectional indices), Hagerty and Land prove that agreement will be high *regardless* of whether the weights for the index differ much from the weights for the average citizen, as in the case for the HDI.

Even though Hagerty and Land's first property states that cross-sectional QOL indices create the highest agreement, it is crucial for policy makers to also have QOL indices that are based on time series for a single country, because national debates more often focus on time-series analyses ("Are you better off than 4 years ago?") than on cross-sectional analyses ("Are we better off than Somalia?"). This type of data results in many more negative correlations among indicators, which tend to decrease agreement in QOL indices.⁴ Therefore, Hagerty and Land assessed distortion for a time-series index, the Index of Social Health (ISH) by Miringoff and Miringoff (1999). They show that the correlation among the 16 social indicators often yielded large negative correlations (e.g., life expectancy above age 65 is negatively correlated ($r = -.85$) with average weekly earnings in the USA since 1970). The question then is whether these negative correlations give rise to a QOL index with agreement too low for a majority of citizens to endorse. Hagerty and Land first examined a

⁴The reason for negative correlations is due in part to "restriction of range" problems (e.g., life expectancy varied far less in the USA since 1970 than it does in a cross-sectional sample of nations, where Somalia has a life expectancy of only 40 years). Negative correlations are also due to preferences of individual nations. For example, the USA seems to prefer higher GDP/capita at the expense of some loss in equality, compared to European nations. Such a policy could result in negative correlation between these indicators as inequality is pushed up in order to gain GDP/capita.

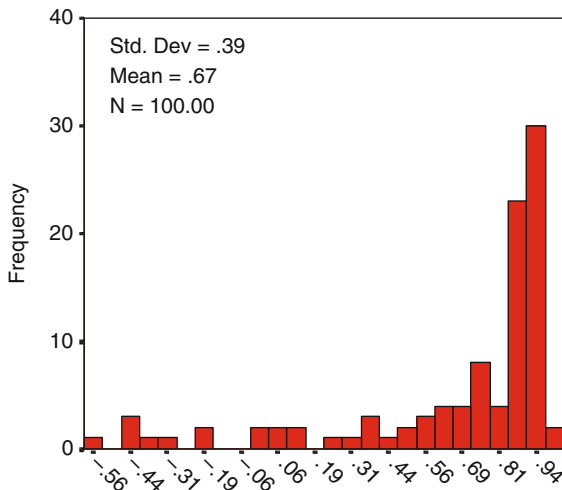


Fig. 9.1 Distribution of agreement $A_{E,i}$ between the equal-weight QOL index of the Index of Social Health and 100 simulated individuals with uniformly distributed weights

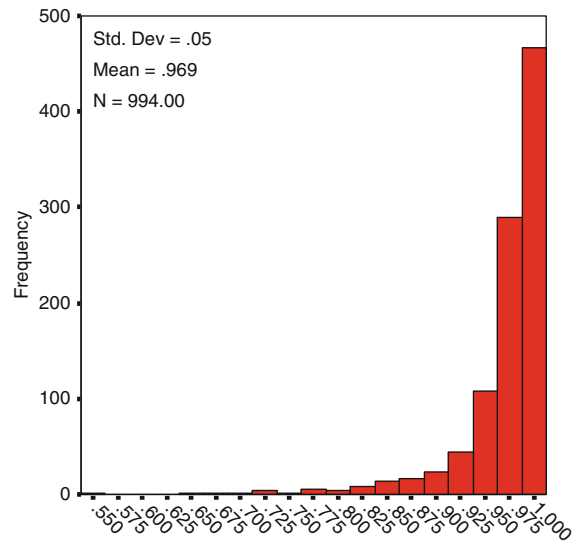


Fig. 9.2 Distribution of agreement $A_{E,i}$ between the equal-weight QOL index of the Index of Social Health and the 994 actual US respondents of the EIU survey

“benchmark” case simulating 100 citizens’ weights to be uniformly distributed⁵ across each of the 16 attributes. The results are shown in Fig. 9.1, where the distribution of correlations between the QOL index (with equal weights) and the 100 simulated individuals is plotted. Despite the fact that the correlations *among the indicators* are negative due to the time-series nature of the index, the correlations between the QOL index and the 100 individuals show that most have very high agreement with the QOL index with equal weights. The average agreement $A_{E,i}$ is .67, and over 50% of simulated individuals have agreement $A_{E,i}$ greater than +.7, the typical value that psychologists chose to show high reliability between raters. Hence, the equal-weighting index for the ISH would induce sufficient agreement to correctly capture more than 50% of these simulated citizens’ QOL judgments.

Hagerty and Land compared this “benchmark” case of uniformly distributed weights to actual sur-

veys of weights from the WVS and *The Economist Intelligence Unit* (EIU). Figure 9.2 shows the distribution of correlations $A_{E,i}$ between the QOL index for ISH (with equal weights) and the 994 US respondents to the EIU survey. Mean agreement is +.96, and over 90% of respondents displayed correlations higher than .7. Hence, not only a majority, but a supermajority of the EIU respondents would accept this equal-weighted index for ISH. Figure 9.2 shows higher agreement between the QOL index and respondents because the real respondents in Fig. 9.2 are not uniformly distributed, but have sharply unimodal distributions.

The second property that increases agreement is whether the distribution of citizens’ actual weights is unimodal as opposed to bimodal. The intuitive reason behind this is that, when weights are unimodal, a single index can be constructed near the mean to capture the weights of most citizens. In contrast, a polarized indicator such as “number of abortions performed” is likely to have weights that are highly bimodal, with some citizens extreme on one side, others extreme on the other side of the distribution, and fewer in the middle. In actual surveys of weights, Hagerty and Land calculate that all distributions they examined for citizens in 40 countries are unimodal rather than bimodal distributions, increasing the likelihood of agreement by an index. (In fact, if an indicator is as

⁵ A uniform distribution of citizen’s importance weights ensures that any proportion between 0 and 1 has equal likelihood of being chosen and assigned to an index component for a simulated individual. In fact, as noted in the text, empirical distributions of importance weights show that some values are more likely to occur than others. Therefore, the assumption of a uniform distribution represents an extreme that is used to ascertain whether or not the resulting index is highly correlated with the QOL index.

highly polarized as abortion, we recommend that it *not* be included in the index because it decreases the chance of agreement, though it should be included in the social report).

The third property that increases agreement is whether the distribution of citizens' weights is negatively correlated for many indicators. In such a case, people who highly value one indicator would always place a very low value on another indicator. Interestingly, Hagerty and Land have found no such negative correlations in the WVS or the EIU surveys, increasing the likelihood of agreement.

The last property that increases agreement is whether every citizen weights an indicator with a positive number. For instance, no one prefers lower life expectancy over higher life expectancy. This property seems quite reasonable for most social indicators (health, income, housing, job satisfaction), and in fact, most surveys do not allow negative weights (Inglehart 2000; Campbell et al. 1976). In contrast, including an indicator such as the number of abortions is likely to create this condition. Such a condition generates more radical differences among individual citizens, and results in lower agreement for any QOL index. Hence we recommend not including any indicators where some citizens hold positive weights but others hold negative weights (though of course all indicators should be included in the larger social report).

Optimal weights for a QOL index. Analyzing a weighted average model of QOL judgment of the form of Eq. 9.1, Hagerty and Land (2007) show mathematically that: (1) *if a survey is available to measure the distribution of citizens' importance weights for each indicator, then agreement is maximized when the index is constructed using the mean weights of citizens.* But, since such surveys are often not available, they also prove that (2) *constructing an index with equal weights produces what in statistics is termed a minimax estimator (that is, equal weighting will minimize maximum possible disagreements).* We note that many of the indices reviewed in this chapter already use equal weighting, but the reasoning behind equal weighting was never well justified. In the context of the weighted average model of QOL judgments of Eq. 9.1, the proofs of Hagerty and Land (2007) now place current practice on a sound theoretical footing, and show how it is possible to further increase acceptance through surveys.

Review of Existing QOL Indices

Having articulated several principles for QOL index construction, we can now review and evaluate a number of existing QOL indices. Composite indicators of QOL have historical roots in economics, where Bentham's social welfare function simply added the individual happiness of each person to get total social welfare. Sen (1993) continues this research stream, provides a set of minimal requirements for a summary utility index to exist, and helped develop the Human Development Index. Kahneman et al. (2004) propose a formal set of National Well-Being Accounts that adds results from psychology to the economic framework, which we review below. In the area of sociology, Land (2000: 2687) documents the rapid growth of QOL indices:

With the tremendous increase in the richness of social data available for many societies today as compared to two or three decades ago, a new generation of social indicators researchers has returned to the task of summary index construction. Some examples: (1) at the level of the broadest possible comparisons of nations with respect to the overall quality of life, the *Human Development Index* (United Nations Development Programme 1993), *Diener's* (1995) *Value-Based Index of National Quality-of-Life*, and *Estes' (1988; 1998) Index of Social Progress*; and (2) at the level of comparisons at the national level over time in the United States, the *American Demographics Index of Well-Being* (Kacapyr 1996), the *Fordham Index of Social Health* (Miringoff and Miringoff 1996), and the *Genuine Progress Indicator* (Redefining Progress 1995).

The QOL indices he cites vary on number of indicators, whether they incorporate only "objective" indicators such as crime rate or "subjective indicators" such as social surveys, whether they are cross-sectional (multiple countries at one point in time) or time series (one country at multiple points in time), and the weights they assign to social indicators. Each will be briefly described here. A summary of each index is given in Table 9.2. Further detail on many of these indices is provided by Hagerty et al. (2001).

1. *The Human Development Index (HDI).* The HDI is a combination of three indicators measured cross-sectionally for each of a set of countries: longevity, knowledge (literacy, weighted 2/3, and years of schooling, weighted 1/3), and income. Sen's capability approach to QOL is used, described as "a process of enlarging people's choices" (United Nations Development Program 1990: 10).

Table 9.2 Fourteen current QOL indices and some of their properties

Index name	Disaggregates into subseries (domains)?	Disaggregates into subpopulations?	Standardizes each indicator?	Construction of index	Reflects citizens' importance weights?
Human Development Index UNDP (2001)	Yes, income, education, health	Yes, in later indices for the poor and women	Yes: Max-Min. Income is log not linear.	Additive with equal weights. Log (income) + education + health	No survey, but experts preview
Genuine Progress Indicator Redefining Progress (1995)	Yes	No	Yes: to dollars	Additive with equal weighting for all money units	No
Index of Economic Well-Being Osberg and Sharpe (2000).	Yes, income, investment, inequality, insecurity	No	Yes	Additive with unequal weights	No, though does sensitivity analysis on alternate weights
National Well-Being Index Kahnemann et al. (2004)	Yes, 19 domains considered	Possible, but not reported	Yes: to affect on 6 point scale	Additive with weights equal to time spent on that activity	Yes, time-use survey
Money Magazine Guterbock (1997)	Yes, 40 indicators, heavily weighted to economic conditions	No	Proprietary and unknown	Proprietary and unknown, but includes cost of housing, amenities, etc.	Yes, through a reader-convenience sample of 250. Also allows readers to input their own weights
Index of Social Progress Estes (1997)	Yes	No, though some indicators include only at-risk population	Yes	Additive with equal weights	Yes, panel of expert citizens
Index of Social Health Miringoff and Miringoff (1999)	Yes, 16 indicators. But not calculated over entire population. Includes vulnerable subgroups only	No, though some indicators describe only vulnerable subgroups	Yes: Max-Min	Additive with equal weights	No, fails to include GDP, average life expectancy
Happy Life-Expectancy Veenhoven (1996)	Yes, subjective happiness, life expectancy	No, though possible	No: cardinal measurement of both domains	Multiplicative with equal weights	No

American Demographics Index of Well-Being Kacapyr (1997)	Yes: reports individual indicators	No	Yes	Additive with equal weights	No
Netherlands' Living Conditions Index Boelhouwer and Stoop (1999)	Yes, reports individual indicators but not sub-groups	Yes, reports by province	Yes	Additive with equal weights	No
Economist Intelligence Unit Economist Intelligence Unit (2005)	Yes	Not reported: subgroup data from some countries not available	Yes	Additive with unequal weights	Yes, but convenience sample of readers
Australian Quality of Life Index	Yes, standard of living, health, relationships, what they are achieving in life, safety, community connection, future security	Yes	Yes	Additive with equal weights	No
Child and Youth Well-Being Index Land et al. (2001, 2007)	Yes, seven domains: material well-being, social relationships, health and safety, educational attainments, place in community, and emotional/spiritual well-being	Yes, by race and ethnic group	Yes	Additive with equal weights	No, but reviewed by experts
Kids Count Index (Annie E. Casey Foundation)	Yes	Yes	Yes	Additive with equal weights	No, but reviewed by experts

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 percapita gross domestic product. Subsequently, the GDP/capita was modified by using an Atkinson formulation that “the higher the income relative to the poverty level, the more sharply the diminishing returns affect the contribution of income to human development” (United Nations Development Program 1993: 91).

Each HDI indicator is standardized in the sense that it is assigned a value between 0 and 100, where 0 represents the lowest-ranking country and 100 the highest-ranking country. The use of minimum and maximum values is faulted when standardization is performed each year. 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: 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 (Trabold-Nubler 1991: 241).

2. *The Genuine Progress Indicator (GPI)*. The GPI (Redefining Progress 1995) was developed from an economic background, and attempts to value all of its indicators in dollar terms from 1950 to present. It broadens the conventional gross domestic product 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 (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 depreciation 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.

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. Also, the economic statistics are difficult to disaggregate to subgroups such as the poor, disabled, etc.

3. *The Index of Economic Well-Being (IEWB)*. The IEWB was developed by Osberg and Sharpe (2000) and is posted at www.csls.ca. Though it is derived from strictly economic theory, it does not attempt to measure QOL in dollars, and integrates four major QOL domains: average consumption flows (including 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, cost of commuting, household pollution abatement, auto accidents, crime, changes in working time, government services, and the value of unpaid work), aggregate accumulation of productive stocks (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 CO₂ emissions), inequality in the distribution of individual incomes (measured by 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), and insecurity in the anticipation of future incomes (change over time in the economic risks associated with unemployment, illness, “widowhood,” and old age). The weights attached to each of these four components of economic well-being can vary, depending on the values of different observers, though for most of their publications, the weights assigned are [.4, .1, .25, .25].

The IEWB has been estimated at the national and international level and can be disaggregated to the province level, so it can help policy makers at these levels in program and policy development. However, it is difficult to disaggregate it to special populations, such as the elderly or immigrants because government statistics do not break out these groups.

4. *National Well-Being Accounts (NWBA)*. An attempt to concatenate economic with psychological theory was made by Kahneman et al. (2004) with their proposed National Well-Being Accounts; see also the chapter by Diener and Tov in this *Handbook*. It is proposed to use time diaries to track citizens' positive and negative affect (pleasant and unpleasant emotions) during each of 19 activities (intimate relations, socializing after work, dinner, lunch, relaxing, exercise, praying, socializing at work, watching TV, phone at home, napping, cooking, shopping, computer at home, housework, childcare, evening commute, working, and morning commute). It is likely that some of these activities can be combined, since they are similar and contain similar affect.

The NWBA approach assumes that well-being is separable over time, so that average affect can be weighted by time and added to get overall well-being for one person, and averaged to get average well-being for the population. The resulting index is standardized because affect is measured on a seven-point scale. It can be computed for any sub-population because it is survey-based. The model is a weighted additive, where the weights are time spent in each activity. The Bureau of Labor Statistics now publishes the monthly American Time Use Survey, though it does not currently collect affect ratings for each activity. Kahneman et al. argue that this index is consistent with economic theory and should be acceptable to economic experts. It remains to be fully developed, implemented, and reported on a continuing basis.

5. *Money Magazine's "Best Places" (MBP)*. 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. *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 are 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 twenty-first century Real Estate brokers, and (7) housing appreciation rate over the past 12 months from twenty-first century. In the third stage, the individual indicators are aggregated into nine broad categories matching categories previously derived in the first stage.

Guterbock (1997) does a masterful job of "retro-engineering" the skimpy data provided by *Money* over 10 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 USA that are ranked in this index (Guterbock 1997). We applaud the use of surveys to assess citizens' weights for this QOL index. However, we show later that the inclusion of indicators, such as "housing prices," is likely to increase distortion and reduce public acceptance of this index as a QOL index.

6. *Estes' Index of Social Progress (ISP)*. In a series of publications dating back to 1984, Richard J. Estes (1984, 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 (1) identify significant changes in the "adequacy of social provision" occurring throughout the world and (2) assess national and international progress in providing more adequately for the basic social and material needs of the world's growing population.

The ISP consists of 46 social indicators that have been subdivided into ten subindexes: Education, Health Status, Women Status, Defense Effort, Economics, Demographic, Geography, Political

Participation, Cultural Diversity, and Welfare Effort. 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.” Estes has computed the ISP on 10 and 5-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 subindex was analyzed for its relative contribution toward explaining the variance associated with changes in social progress over time. Standardized scores of the component indicators then were multiplied by the factor loadings to create weighted subindex scores which then were summed to obtain the “Weighted Index of Social Progress” (WISP).

7. *The Index of Social Health (ISH)*. The Index of Social Health was developed by the Fordham Institute for Innovation in Social Policy (Miringoff and Miringoff 1996, 1999). 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 and Miringoff (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.

However, the authors fail to address the question of whether these measures are valid. 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 Miringoff and Miringoff (1999), only one page is devoted to discussing why they chose the 16 components of their index, and no validation studies are cited.

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: It is vulnerable to outliers, and will vary with the number of years in the sample (Hagerty 1999). On the other hand, explaining their index to lay people is easier than explaining standardized scores.

8. *Veenhoven’s Happy Life-Expectancy Scale (HLE)*. 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 0–10 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: 29)

Veenhoven validates the HLE by showing positive correlations (controlling for a country’s affluence) for HLE and many social indicators (e.g., purchasing power, state expenditures as a percent of GDP, percent literate).

A potential problem for HLE is that it changes very slowly, so that country rankings will not change much each year. It may be considered a very useful “output” or “outcome” measure, but it is missing the “throughput” measures of a country’s performance on the other domains (freedom, family and job satisfaction, etc.).

9. *American Demographics’ Index of Well-Being (AD-IWB)*. The *American Demographics* magazine

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 was a monthly composite of five indicators and was unique in that it was updated every month, with subseries updated monthly by government sources. The five areas were monitored with 11 monthly time series: consumer attitudes (Consumer Confidence Index and Consumer Expectations Index), income and employment opportunity (real disposable income per capita and employment rate), social and physical environment (number of endangered species, crime rate, and divorce rate), leisure (168 less weekly hours worked and real spending on recreation per capita), and productivity/technology (industrial production per unit of labor and industrial production per unit of energy). Each component was “benchmarked” to an April 1990 level of 100. The separate reporting of each component and the socioeconomic forces undergirding the change were an important, informative feature of the Index. The weights for each element of each component were determined “by fitting a trend line to the series from 1983 to 1997. Then 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 was inversely proportional to the variance from its own trend, which means that data series with relatively smaller fluctuations around their trends were given more importance in the index. The weights were 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. After applying the weights, both moves are seen to be equivalent” (e-mail communication 4/9/99). Thus, by the above-described device, change in the Index was influenced equally by each of the five components.

The AD-IWB employed a weighting scheme unique among QOL studies. The purpose was to equalize the influence on change, rather than the influence of the item upon the output of QOL.

10. *The Netherlands' Living Conditions Index (LCI)*. The Netherlands Social and Cultural Planning Office (Boelhouwer and Stoop 1999) has developed the Living Conditions Index (LCI). Its base year is 1974, with annual updates 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. The specific indicators have changed over the years to address new public policy problems. The authors argue strongly that only objective indicators 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 than the separate components, because the correlation of LCI with happiness is higher than the correlation of any of the separate components. Hence, the separate domains are not redundant, but provide some additional predictive validity. However, a multiple regression should be reported in order to sort out which domains add significant explanation to LCI and happiness. Unequal weights are assigned in computing the LCI by factor-analyzing 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.
11. *The Economist Intelligence Unit's Quality of Life Index (EIU-QOLI)*. The Economist Intelligence Unit (2005) published their first QOL index, composed of ten publicly available series. The domains are: material well-being (in GDP PPP \$), health (in life expectancy), political stability and security (Economist ratings), family life (in divorce/1000), community life (church or trade union attendance), climate and geography (in latitude), job security (unemployment rate), political freedom (ratings by Freedom House), and gender equality (ratio of

average male to female earnings). The index is a weighted average model, with weights derived from a multiple regression predicting life satisfaction in 74 countries where data were available. These scores are then related in a multivariate regression to various factors that have been shown to be associated with life satisfaction in many studies. As many as nine factors survive in the final estimated equation (all except one are statistically significant; the weakest, gender equality, falls just below). Together these variables explain more than 80% of the intercountry variation in life-satisfaction scores. Using beta coefficients from the regression to derive the weights of the various factors, the most important were health, material well-being, and political stability and security. These were followed by family relations and community life. Next in order of importance were climate, job security, political freedom, and finally gender equality. No subgroups within countries were calculated. Data are not available for subgroups from many of those countries.

12. *The Australian Unity Wellbeing Index (AUWBI)*. Cummins et al. (2005) have developed a continuing survey sampling 2,000 Australian citizens semiannually and have created two indices called the Personal Wellbeing Index and the National Wellbeing Index. The Personal Wellbeing Index is the average level of satisfaction across seven aspects of personal life—health, personal relationships, safety, standard of living, achieving, community connectedness, and future security. The National Wellbeing Index is the average satisfaction score across six aspects of national life—the economy, the environment, social conditions, governance, business, and national security. Both indices are based on subjective indicators measuring satisfaction with each domain. Each indicator is a single item in the survey, rating satisfaction on a 0–10. The scores are then combined across the seven domains to yield an overall Index score, which is adjusted to have a range of 0–100. Hence each series is already standardized.

The indices are embedded in an extensive social report that disaggregates the indices by domain and by subpopulations, examines trends over time, and relates changes to changes in current events and changes in demographics. The index currently extends from April 2001 and is in its 16th wave.

As more waves are collected, time-series analysis correlating the subjective measures with official objective statistical series can be done.

13. *The Child and Youth Well-Being Index (CWI)*. The Foundation for Child Development Child and Youth Well-Being Index Project (Land et al. 2001; Land et al. 2007) calculates changes in the QOL of children and youth in the USA. A general description of the Index, annual reports, charts and tables, and scientific papers are posted at www.soc.duke.edu/~cwi/. The CWI is composed of 28 national-level Key Indicator time series since grouped into seven domains that are based on Cummins' (1996) review of subjective well-being studies: family economic well-being, social relationships (with family and peers), health, safety/behavioral concerns, educational attainments, community connectedness (participation in schooling or work institutions), and emotional/spiritual well-being. It uses equal weights of Key Indicators within domains and equal weights of the seven domains to calculate a composite QOL index for children and youth. Annual changes are indexed from two base years, 1975 and 1985. Trends for children's QOL are plotted from the base years. The trends are broken down by race and ethnicity, by infancy, childhood and adolescence, and by each of the seven domains. In Land et al. (2007), some evidence of the external validity of the CWI is provided in the form of a high correlation with trends in overall life satisfaction of high school seniors from 1975 to 2003. Annual reports on the CWI are broadly disseminated to the American public by the Foundation for Child Development and have resulted in much print and electronic media coverage.
14. *Kids Count Index (KCI)*. In collaboration with the CWI Project, the Annie E. Casey Foundation has developed a Kids Count Index to estimate changes in the QOL of children and youth in each of the 50 states of the USA. The index includes ten indicators, which have not been subdivided into domains (the authors say that ten indicators are few enough in number to make domains unnecessary). The indicators are: percent of low-birth-weight babies, infant mortality rate, child death rate (ages 1–14), teen death rate (ages 15–19), teen birth rate, high-school dropout rate, idle teens, parental underemployment, child poverty rate, and children in

single-parent households. The indicators are each standardized and equally weighted to create the index. They calculate change over time as a change in indicator from baseline, relative to baseline year of 1990. One of their stated goals is to generate publicity for the plight of children using scientifically generated data. They are indeed achieving their goal of publicity, citing over 1,160 newspaper articles referencing the index, 509 television interviews, and hundreds of radio interviews, and over 750,000 internet visits per year. Interestingly, they report that the state's rank is listed in the headline of the newspaper article 36% of the time, and is mentioned in the body of the article in 62% of articles.

Common Criticisms of Indices and Recommended Solutions

An index of QOL is a relatively novel concept to journalists and laymen, and they will have questions to assess whether the index is credible, unbiased, and informative. Below are some typical criticisms and solutions that are commonly posed.

1. *"A composite index can obscure whether some indicators have moved in opposite directions."* We agree that this is a danger, and remind critics that every summary statistic suffers this drawback. This problem can easily be remedied by including discussion in a companion social report on which indicators are improving, and which are declining, both of which are important information for citizens and policy analysts. A QOL index is not intended to stand alone, but must be accompanied by a social report that examines trends in each subseries.
2. *"A composite index could obscure sub-group comparisons, such that disadvantaged populations may be worse off even when the average QOL index improves."* Again, we agree that this is a danger, and our principles recommend that the social report disaggregates measures of conditions for disadvantaged groups. Such breakdowns for the elderly, children, and minorities in the social report already are standard practice in the Swedish and German social reports. In summary, composite indices are quite useful to *begin* a report, but should not *end* the reporting.
3. *"Composite indices may be appropriate for uni-dimensional phenomena such as the CPI, but they cannot capture multi-dimensional concepts such as Quality of Life."* We agree that developing indices for multidimensional phenomena is more difficult than for unidimensional concepts. But citizens and decision makers are already making these judgments without the help of science to make political decisions and to draft laws. The words "quality of life" are invoked more than 20 times per week on the floor of the US Congress (GPO 1999), with varying definitions and no measurements. Citizens and decision makers would certainly benefit from scientific attempts to capture QOL, by improving the reliability and validity of subseries, by reducing perceptual biases to which humans are prone, and by providing a common language to discuss which indicators should be included and how they should be weighted for each application. This chapter provides seven principles for achieving this.
4. *"A composite index could be dominated by a single indicator. If the index assigns very high weight to one domain, then the index will be driven by that domain only, and the index would be distorted."* This is a potential danger, and a section of the social report (1) must show how each subseries is standardized to prevent one subseries from dominating, and (2) must justify what weights are applied. In the absence of surveys of citizens or decision makers to assess their weights, an easy way to avoid this problem is simply to apply equal weights to all indicators, which Hagerty and Land (2007) mathematically proved to be the minimax solution that minimizes maximum distortion of the index.
5. *"A composite index may not reflect the 'true' weights that citizens actually apply to social indicators."* Johansson (2002) warns that even surveys of citizens' weights may not be correct because citizens' weights may *change* as they discuss the issues and listen to candidates. Such dynamically changing weights are likely to occur for some indicators and instances, and, as surveys become better at measuring weights, it would be informative to track any changes in weights during an electoral cycle. Such a development parallels the history of the CPI, which was initiated with static weights but was modified to dynamic weights as research progressed.
6. *"A composite index provides an 'easy way out' for citizens and policy makers to avoid reading the entire report."* We have no doubt that many citizens

will only hear the “headlines” of any report because they are “satisficers” with limited time, memory, and cognitive skills. To serve them best, we should develop a QOL index that as closely as possible mimics their own judgments if they were to read the entire report. And of course we encourage them to read the report for themselves to understand the movement of subseries and their causes.

7. “A composite QOL index raises the specter that the government begin ‘social planning’ where bureaucrats push citizens into programs they have not helped design.” We strongly reject this type of social planning, and instead suggest that a QOL index should be used to hold agencies *accountable* for improving QOL for citizens in their purview.
8. “If composite QOL indices are so valuable, why doesn’t the government officially adopt a QOL index?” Federal/national governments will probably be the last organizations to adopt QOL indices, because they require acceptance by the largest number of people. But smaller government units have already adopted QOL indicators. (Miringoff and Miringoff (1999) count 11 states and 28 communities). One federal government has already adopted a QOL index (Netherlands LCI), and another country is evaluating a candidate QOL index (Canadian Index of Wellbeing 2009). As experience and credibility with QOL systems grow among local governments and nongovernmental organizations, we expect federal governments to eventually adopt not just one, but a “family” of QOL indices similar to those for the CPI, each appropriate for different subpopulations or situations. This is part of the movement toward evidence-based measures of QOL.

Conclusions

Seven principles for constructing QOL indices have been stated and described above. Based on these principles, several recommendations can be made:

1. We recommend that the social indicators be integrated into a QOL index using the weighted average model, since it well captures the QOL judgments made by real citizens. The model also is robust to errors in measurement.
2. We recommend that the weights used be proportional to surveys of citizens’ own weights for the various indicators, some of which are given in

Table 9.1. This procedure maximizes agreement between citizens and the index, and has the further advantage of protecting the index from political manipulation of the weights and indicators. If surveys of citizens’ weights are not available, then equal weighting minimizes the worst disagreements.

3. We recommend that the set of indicators span the major domains of QOL shown in Table 9.1 or at least as many thereof as possible (the exact name of each domain has not been standardized, nor is this essential). Again, this assures that domains that citizens designate as important are included in the index.
4. We recommend that an indicator be rejected for use in the QOL index (though should be kept in the larger social report) when some citizens place negative weights but other citizens place positive weight on it. As discussed above, “number of abortions” per year may be positively weighted by some as “freedom from government interference,” but negatively by others as murder. Inclusion of such an indicator would decrease agreement by all citizens and lead to lower acceptance. We stress that not all social indicators should be included in a QOL index. A more subtle example of an indicator that should not be included is “average price of a 3-bedroom home” in *Money* magazine’s index. Some people (homeowners) would place a high positive weight on this, but others (homebuyers) would place a high negative weight. In fact, this is an example of a zero-sum negotiation game where every gain for a buyer is a loss for the seller, and joint gains are always zero regardless of the price. Negotiation researchers (Pruitt and Kim 2004; Carnevale and Pruitt 1992) recommend instead including indicators that allow positive joint gains to enhance the framing of shared interests. Much research has shown that this increases the likelihood of agreement and increases joint gains in negotiations. Applying these principles to the *Money* magazine example, a simple “laddering” procedure (“what deeper goals are you trying to achieve with lower housing prices/higher housing prices?”) could replace the single zero-sum attribute (price) with two shared goals: lower cost per square foot of new construction and higher personal income. Both of these new indicators would conform to our assumptions and would result in higher likelihood of agreement.

5. We recommend that an indicator be rejected for use in a QOL index (though should be kept in the larger social report) when the indicator is a “policy indicator” rather than a “goal” or “outcome” indicator. An example of a “policy indicator” is tax policy, where conservatives place a negative weight on average tax burden, and liberals tend to place a positive weight. Tax policy is better viewed as a means to an end, and a successful QOL index would again apply laddering to include the end-state variables (e.g., better health care, education, pollution control, and economic growth). These examples clarify that a QOL index would *not* remove the need for policy analysis and political discussion, but would better *focus* policy analysis and politics by forcing proponents to estimate each policy’s results on the QOL index.

Using these recommendations and the seven principles for constructing a QOL index, our review suggests that it is quite feasible to create QOL indices that are reliable and valid, robust to errors, and well accepted by the public because they capture the QOL judgments that a citizen would make if she were to read the entire report. Such “evidence-based” principles would help prevent the political manipulation of weights and indicators and would strengthen the democratic process.

References

- Boelhouwer, J., & Stoop, I. (1999). Measuring well-being in the Netherlands: The SCP index from 1974 to 1997. *Social Indicators Research*, 48, 51–75.
- Campbell, A., Converse, P. E., & Rodgers, W. L. (1976). *The quality of American life*. New York: Russell Sage.
- Canadian Institute of Well Being. (2009). *How are Canadians really doing? A closer look at select groups* (Special Report, December 2009). http://ciw.ca/Libraries/Documents/ACloserLookAtSelectGroups_FullReport.sfb.ashx.
- Carnevale, P. J., & Pruitt, D. G. (1992). Negotiation and mediation. *Annual Review of Psychology*, 43, 531–582.
- Charnes, A., Cooper, W. W., & Rhodes, E. L. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2, 429–444.
- Charnes, A., Cooper, W. W., Lewin, A. Y., & Seiford, L. M. (Eds.). (1994). *Data envelopment analysis: Theory, methodology, and application*. Boston: Kluwer Academic.
- Cummins, R. A. (1996). The domains of life satisfaction: An attempt to order chaos. *Social Indicators Research*, 38, 303–328.
- Cummins, R. A., Woerner J., Tomy, A., Gibson, A., & Knapp T. (2005). *The wellbeing of Australians – Personal relationships* (Australian Unity Wellbeing, Index: Report 14). Melbourne: Australian Centre on Quality of Life, School of Psychology, Deakin University. ISBN 1 7415 6024 1, http://www.deakin.edu.au/research/acqol/index_wellbeing/index.htm.
- Dantzig, G. B. (1963). *Linear programming and extensions*. Princeton: Princeton University Press.
- Diener, E. (1995). A value based index for measuring national quality of life. *Social Indicators Research*, 36, 107–127.
- Economist Intelligence Unit. (2005). The economist intelligence unit’s quality of life index. *The world in 2005*, pp. 1–5. http://www.economist.com/media/pdf/QUALITY_OF_LIFE.pdf.
- Erickson, R. (1993). Descriptions of inequality: The Swedish approach to welfare research. In M. Nussbaum & A. Sen (Eds.), *The quality of life*. Oxford: Clarendon.
- 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. (1997). Social development trends in Europe, 1970–1994: Development prospects for the new Europe. *Social Indicators Research*, 42, 1–19.
- Estes, R. J. (1998). Social development trends in transitional economies, 1970–1995. In R. H. Kempe, Sr. (Ed.), *Challenges of transformation and transition from centrally planned to market economies* (pp. 13–30). United Nations: United Nations Centre for Regional Development.
- Government Printing Office (GPO). (1999). *Congressional register searchable website*. <http://www.access.gpo.gov>.
- Guterbock, T. M. (1997). Why money magazine’s ‘Best places’ keep changing. *Public Opinion Quarterly*, 61, 339–355.
- Hagerty, M. R. (1999, February) Comment on the construction of the forham index of social health. social indicators network news. *Working group 6 of international Sociological association*, 57, p 10.
- Hagerty, M. R., & Land, K. C. (2007). Constructing summary indices of quality of life: A model for the effect of heterogeneous importance weights. *Sociological Methods and Research*, 35, 455–496.
- Hagerty, M. R., Cummins, R. A., Ferriss, A. L., Land, K., Michalos, A. C., Peterson, M., Sharpe, A., Sirgy, J., & Vogel, J. (2001). Quality of life indexes for national policy: Review and agenda for research. *Social Indicators Research*, 55, 1–96.
- Inglehart, R., et al. (2000). *World values surveys and european values surveys, 1981–1984, 1990–1993, and 1995–1997*. [Computer file]. ICPSR version. Ann Arbor: Institute for Social Research [producer]. Ann Arbor: Inter-university Consortium for Political and Social Research [distributor].
- Johansson, S. (2002). Conceptualizing and measuring quality of life for national policy. *Social Indicators Research*, 58, 13–32.
- Kacapyr, E. (1996). *Economic forecasting*. M. E. Sharpe.
- Kacapyr, E. (1997, October). Are we having fun yet? *American Demographics*, 19, 28–29.
- Kahnemann, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. (2004). Toward national well-being accounts. *American Economic Association Papers and Proceedings*, 94(2), 429–434.
- Land, K. C. (2000). Social indicators. In E. F. Borgatta & R. V. Montgomery (Eds.), *Encyclopedia of Sociology* (Revth ed., pp. 2682–2690). New York: Macmillan.

- Land, K. C. (2004). An evidence-based approach to the construction of summary quality-of-life indices. In W. Glatzer, M. Stoffregen, & S. von Below (Eds.), *Challenges for quality of life in the contemporary world* (pp. 107–124). New York: Kluwer.
- Land, K. C., Lovell, C. A. K., & Thore, S. (1993). Chance-constrained data envelopment analysis. *Managerial and Decision Economics*, 14(November–December), 541–554.
- Land, K. C., Lamb, V. L., & Mustillo, S. K. (2001). Child and youth well-being in the United States, 1975–1998: Some findings from a new index. *Social Indicators Research*, 56, 241–320.
- Land, K. C., Lamb, V. L., Meadows, S. O., & Taylor, A. (2007). Measuring trends in child well-being: An evidence-based approach. *Social Indicators Research*, 80, 105–132.
- Lucas, R. E., Clark, A. E., Georgellis, Y., & Diener, E. (2003). Reexamining adaptation and the set point model of happiness: Reactions to changes in marital status. *Journal of Personality and Social Psychology*, 84, 527–539.
- Miringoff, M. L., Miringoff, M. L., & Opdycke, S. (1996). Monitoring the nation's social performance: The index of social health. In E. F. Zigler, S. L. Kagan, & N. W. Hall (Eds.), *Children, families, and government* (pp. 10–30). New York: Cambridge University Press.
- Miringoff, M. L., & Miringoff, M. L. (1999). *The social health of the nation: How American is really doing*. New York: Oxford University Press.
- Moller, V., & Dickow, H. (2002). The role of quality of life surveys in managing change in democratic transitions: The South African case. *Social Indicators Research*, 58(1–3), 267–292.
- Munda, G., & Nardo, M. (2003). *On the methodological foundations of composite indicators used for ranking countries*. Paper presented at the First OECD/JRC workshop on composite indicators of country performance. JRC, Ispra
- Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., & Giovannini, E. (2005). *Handbook on constructing composite indicators: Methodology and user guide*. Paris: Organization for Economic Co-operation and Development.
- Osberg, L., & Sharpe, A. (2000, November). *International comparisons of trends in economic well-being*. Paper presented at the annual meeting of the American Economic Association, Boston.
- Pruitt, D. G., & Kim, S. H. (2004). *Social conflict: Escalation, stalemate, and settlement* (3rd ed.). New York: McGraw Hill.
- Redefining Progress. (1995). *The genuine progress indicator: Summary of data and methodology*. San Francisco: Redefining Progress.
- Sastre, M. T. M. (1999). Lay conceptions of well-being and rules used in well-being judgments among young, middle-aged, and elderly adults. *Social Indicators Research*, 47, 203–231.
- Sen, A. (1993). Capability and well-being. In A. Sen & M. C. Nussbaum (Eds.), *The quality of life* (pp. 30–53). Oxford: Clarendon.
- Trabold-Nubler, H. (1991). The human development index: A new development indicator? *Intereconomics*, 26, 236–243.
- United Nations Human Development Program. (1990). *Human development report*. New York: Oxford University Press.
- United Nations Human Development Program. (1993). Human development index: Survey of recent reviews. *Human development report*. New York: Oxford University Press.
- United Nations Development Program. (2001). *Human development report*. New York: Oxford University Press.
- Veenhoven, R. (1993) Happiness in Nations, Rotterdam: Risbo. Data available from web-site: <http://www.eur.nl/fsw/research/happiness>.
- Veenhoven, R. (1996). Happy life-expectancy: A comprehensive measure of quality of life in nations. *Social Indicators Research*, 39, 1–58.
- Zhou, P., Ang, B. W., & Zhou, D. Q. (2010). Weighting and aggregation in composite index construction: A multiplicative optimization approach. *Social Indicators Research*, 96, 169–181.