FREDERIK BOOYSEN

AN OVERVIEW AND EVALUATION OF COMPOSITE INDICES OF DEVELOPMENT

ABSTRACT. The search for alternative indicators of development has witnessed the development of a variety of composite indices of development. These indices integrate various social, political and economic aspects of development in measurement. This paper presents an overview and evaluation of composite indices of development in terms of certain broad dimensions of measurement. In terms of method and technique, composite indices are generally additive ones with equally weighted components consisting of variables selected in an ad hoc manner. Numerous criticisms have been leveled at these methods employed in composite indexing. Composite indices are mainly quantitative insofar as the indices are all presented in numerical format. Composite indices are subject to subjectivity despite the relative objectivity of the methods employed in composite indexing. Composite indices are of a cardinal nature, but remain ordinal insofar as differences in index values cannot be interpreted meaningfully. The multi-dimensionality of composite indices represents one of their main advantages. Indices represent aggregate measures of a combination of complex development phenomena. The comparative application of composite indices of development over space and time remains problematic. Composite indices generally combine measures of ends and means. In respect of method and technique, composite indexing is relatively complex. Composite indices are relatively flexible, because changes in selection, scaling, weighting and aggregation can be effected readily, albeit at the cost of comparability. Composite indices perform relatively well in terms of cross-national availability, but few indices perform well in terms of inter-temporal availability. On the strength of the systematic positive association between income and popular composite indices such as the HDI, many have claimed that these indices represent no real contribution to the literature on indicators research. Composite indices, furthermore, are often considered to be ideological statements rather than practically functional indicators. Yet, composite indices represent useful supplements to income-based development indicators. These indices remain invaluable in terms of their ability to simplify complex measurement constructs, to focus attention and to catch the eye, thus enhancing their political appeal.

Social Indicators Research **59:** 115–151, 2002.

1. INTRODUCTION

The search for alternative indicators of development has witnessed the development of a variety of composite indices. These indices represent macro-level valuations of development insofar as they integrate various social, political and economic elements of development in measurement (Todaro, 1989: pp. 108–113). These efforts are founded on the belief that 'no single yardstick exists to measure development just as no single set of objectives can describe adequately the diversity of development conditions in the world' (Wilson and Woods, 1982: p. 11). Sainz (1989: pp. 156–160), furthermore, argues that the increasing social and economic heterogeneity of nations necessitates the combining of both synthetic and specialised indicators in measurement. The motivation behind efforts at composite indexing differs. While some analysts intended their indices to independently evaluate the social situation, others saw them as indicators to contrast with economic indicators, or indicators that combined different aspects of development in measurement (Sainz, 1989: p. 156).

2. A FRAMEWORK FOR EVALUATION OF DEVELOPMENT INDICATORS

One can classify and evaluate development indicators according to a number of general dimensions of measurement (Table I). Although these dimensions do not fully address all the methodological and conceptual issues involved in measurement, they do present a useful framework for distinguishing between different types of development indicators. The focus here is on a general comparison of composite indices of development in terms of certain broad parameters rather than a discussion of issues relating to specific types of measurement.

The call for such a framework for classifying and evaluating development indicators is not new. Drewnowski (1972: p. 77) claimed that one requires some 'ordering principles for the selection of useful indicators and rejection of ill-conceived and inapplicable ones'. Wish (1986: pp. 97–98) similarly argued that 'indicators require a systematic rationale for categorisation'. These dimensions,

TABLE I

General dimensions for classifying and evaluating development indicators

Dimension	Description
1. Content	What aspects or facets of development does the indicator measure?
2. Technique and method	Does the indicator measure development in a quantitative (qualitative), objective (subjective), cardinal (ordinal), or uni-dimensional (multi-dimensional) manner?[1]
3. Comparative application	Does the indicator compare the level of development (a) across space ('cross-section') or time ('time-series'), and (b) in an absolute or relative manner?
4. Focus	Does the indicator measure development in terms of input ('means') or output ('ends')?
5. Clarity and simplicity	How clear and simple is the indicator in its content, purpose, method, comparative application and focus?
6. Availability	How readily available are data on the particular indicator across time and space?
7. Flexibility	How relatively flexible is the indicator in allowing for changes in content, purpose, method, comparative application and focus?

however, are not always mutually exclusive, but often overlap and/or are interdependent. In each case the nature of this overlap or interdependence will be related in the subsequent discussion. Finally, the sequence in which the dimensions are presented here should not be seen as denoting their relative importance.

3. AN OVERVIEW AND EVALUATION OF COMPOSITE INDICES OF DEVELOPMENT

In the subsequent pages composite indices of development are discussed and evaluated at the hand of the seven dimensions of measurement described in Table I. There is no necessarily 'ideal' or

'correct' answer to each of the questions posed in Table I. Particular development indices, rather, will each yield a unique set of answers.

3.1. *Content of Composite Indices of Development*

The first question one needs to answer when wishing to classify any measure of development is what aspect of development the indicator measures. The conceptual framework stands central in this regard. This relates to the meaning of development in the particular analysis. Development indicators tend to differ according to the different views on the meaning that individual researchers ascribe to development (Baster, 1985: pp. 38–43). The discussion on content is here confined to explaining what are understood to be composite indices of development. Arguments as to which specific elements of development are to be included in composite indices are related elsewhere. Composite indices represent measures arrived at via some empirical aggregation of a number of economic, social and political variables (Babbie, 1995: pp. 161–175). In Sainz's (1989: p. 156) words, composite indices are those measurement efforts that require a 'synthesis of numerous factors into one given factor'. As with other indicators of development, these indices propose to 'facilitate international comparisons or the study of a single country's development over time' (Thiessen, 1997: p. 13).

3.2. *Technique and Method of Composite Indexing*

Composite indexing entails the aggregation of any number of economic, social and political indicators. Composite indexing involves four steps, i.e. selection; scaling; weighting and aggregation; and validation (McGranahan et al., 1972). It is important to note that the steps of composite indexing do not necessarily follow in this sequence. It is a concurrent effort during which selection can be altered, weights adjusted and variables rescaled in order to arrive at final index estimates.

3.2.1. *Selection of variables and components*

Choices regarding two issues are required (McGranahan et al., 1972: pp. 8–10; Drewnowski, 1974: pp. 19–33; Ginsberg et al., 1986). In the first instance, the number and nature of the components that will make up part of the composite index need to be determined.

Secondly, the specific variables employed in estimating each of the component indices need to be selected. Such selection is generally based on theory, empirical analysis, pragmatism or intuitive appeal, or some combination thereof (Adelman and Morris, 1972: pp. 117–119; Diener and Suh, 1997: pp. 192–200). Political and policy considerations also figure in selection insofar as composite indices are in some cases developed with a view to informing particular audiences regarding certain issues (Stewart, 1985: pp. 1–2).

Both bivariate and multivariate statistical techniques are employed where selection is based on empirical analysis (Babbie, 1995: pp. 161–175). Whereas bivariate analysis measures the strength of the association between all pairs of variables, multivariate analysis assesses the overall power of any collection of variables to measure any other variable. Bivariate analyses traditionally employ correlation matrices in selection. The most strongly correlated variables are selected from a larger sample of applicable variables. The rationale behind this is that development is a highly interdependent process and that characteristics will tend to be grouped together in successful economies (McGranahan et al., 1972: pp. 14–18). Most notable of the multivariate techniques employed in composite indexing are discriminant, principal component and factor analyses (Felipe and Resende, 1996; Kallmann, 1997: p. 18). The purposes of these analytical techniques are to determine the number of latent variables underlying the data, to condense the data and to define the content and meaning of the factors or latent variables accounting for the variation in the data (Child, 1970; Everitt and Dun, 1991). These techniques require a proxy for development. Per capita income and life expectancy are the two variables most often employed as dependent variables in discriminant and principal component analyses for the purpose of indicator selection.

Another central consideration during selection is the purpose of measurement. Since the goal is the international comparison of well-being, the components will have to be of universal significance and cross-cultural applicability. Cross-national empirical analyses of the association between development phenomena are useful in identifying these shared components. National and group-specific analyses can also be dealt with in composite indexing. One

can select different components and variables to devise composite indices for countries, regions or groups. The UNDP, for example, reports human poverty indices for developed as opposed to developing countries. Arguments in favour of the latter approach to selection are often based on the fact that industrialised countries have reached the upper end of the so-called plateau curve of basic needs achievement, thus compromising the discriminant ability of the resulting indices (Diener, 1995: pp. 107–125).

Equally dependent on purpose is the distinction between variables focused on ends as opposed to means. Adelman and Morris (1972: pp. 117–119) claim that variables included in composite indices should focus on either the one or the other. McGranahan et al. (1972: p. 12), however, argue that only intended ends should be included in composite indices. Thus, unintended negative outcomes such as crime and pollution should be excluded. But some indices, such as the UNDP's Human Development Index (HDI), were specifically designed to include both means and ends. Some variables, furthermore, represent measures of both ends and means (Morris, 1979: pp. 20–40). Literacy, for example, is both an end (being able to read and write) and a means (increasing one's future earnings potential).

Other important selection criteria include validity, reliability, comparability, simplicity, and data availability (Morris, 1979: pp. 20–40; Estes, 1984: pp. 30–31). In terms of validity, it is crucial that variables measure the component they are supposed to measure. Differences in indicators must reflect differences in the particular measurement construct (Morgan, 1968: pp. 37–39; Babbie, 1995: pp. 161–175). Some variables relate to needs that can be met in a variety of ways. The need for transportation, for example, can be satisfied via a variety of modes of transport, depending on geographic constraints and personal preferences. Where modes of satisfaction differ depending on circumstances, tastes and culture, variables need either to be adjusted so as to allow for these differences, or to be excluded, given their resulting lack of validity (McGranahan, 1995: pp. 54–55). Other variables are problematic in that they are both positively and negatively associated with certain aspects of development (Drewnowski, 1972: pp. 85–86; Rose, 1995: 114; Streeten, 1995b: p. 28; Veenhoven, 1996: pp. 2–5). Divorce,

and suicide among the terminally ill, are cases in point. While both signify an expansion of choice, they are also considered indicative of a deterioration of the social fabric and of family life. Urbanisation, furthermore, may cause both improved access to health services and increased criminal activity. A good but difficult test for validity is that improvements in the selected variables must imply a gain for all members of society (Morris, 1979: pp. 20–40; Estes, 1984: pp. 30–31; Streeten, 1995a: pp. xii–xiv).

When it comes to comparability, meaningful comparisons of indices require a standardisation of the concepts and methodologies employed in data collection (Morris, 1979: pp. 20–40; Estes, 1984: pp. 30—31). As will become evident in these pages, some variables have come a long way towards complete standardisation. Others, though, remain suspect. According to Ginsberg et al. (1986: pp. 101–120), the sample of countries to which composite indices are applied also needs to be relatively comparable in terms of the nature of political and economic systems, the level of development, population size, and geographic proximity. Different systems of economic organisation often imply different assumptions as to the prescribed patterns of development, this in turn implying a need for differential selection (Morris, 1979: pp. 20–40). Thus, some analyses include only countries with populations in excess of one million and exclude countries with centrally-planned economies and authoritarian regimes. The UNDP, for example, excludes industrial countries from their estimation of the Human Poverty Index (HPI) because the particular indicators cannot adequately quantify deprivation in advanced economies (UNDP, 1997). In general, though, most efforts at composite indexing have been relatively indiscriminate in terms of country coverage.

Selection requires a balance between simplification and complication (McGranahan et al., 1972: pp. 8–10; Drewnowski, 1974: pp. 19–33). On the one hand, the greatest threat to simplicity is the tendency to keep on adding variables and components to indices as the scope of development issues expands (Thais, 1989: p. 174). According to Weigel (1986: pp. 1423–1428), indices concerned with preferences rather than needs are particularly susceptible to this tendency. The resulting indices are anything but simple and manageable (an important criterion for a good index) (Ul Haq, 1995:

pp. 46–48). This may also divert attention from crucial components of development (Porwit, as quoted in Rao et al., 1978: pp. 31–32). Oversimplification, on the other hand, runs the danger of omitting significant components of the development problem. Unfortunately there are no hard and fast rules for identifying indices that are simplistic yet substantive and informative.

In terms of data availability, a 'constant awareness of the sources and interpretation of data' is required (McGranahan et al., 1972: p. 3). Data availability needs to be assessed with regard to both the timeliness and the width of the database. Timeliness requires that analyses employ the most recently available and similarly dated estimates of data. According to Estes (1984: pp. 30–31), this condition may be waived where analyses are provisional, the particular variable is of the utmost importance, or acceptable proxies are not available. With regard to the width of the database, data samples need to be large enough to ensure statistically significant results. Where more than one variable represents a proxy of the particular component and variables are equally qualitative, one should select the variable(s) with the widest base (McGranahan et al., 1972: pp. 14–18). Where representativeness and statistical significance are in jeopardy, estimates of missing observations can be derived from existing data using regression techniques. Alternative options include the use of group averages and mean, median or random values (Ginsberg et al., 1986: pp. 17–47; Babbie, 1995: pp. 161–175).

Despite claims that value-judgements and cultural issues should be avoided as far as possible in selection (Morris, 1979: pp. 20–40), composite indexing remains an inherently value-laden and subjective exercise (Drewnowski, 1972: pp. 85–86; Rose, 1995: p. 114). Even when empirical techniques are employed in data reduction, analysts are required to make choices regarding the variables included in the original sample of variables. These choices are further compounded by the fact that (i) some constructs of economic, social and political development are more difficult to measure than others, (ii) opinions regarding the relative desirability of components and variables differ, (iii) development is an interdependent and multidimensional process within which many grey areas remain, and (iv) underdevelopment differs in terms of

its regional characteristics (Felipe and Resende, 1996: p. 184). Thus, there exists no general framework for selection (Wish, 1986: pp. 93–97).

3.2.2. *Scaling of variables*

Scaling entails the 'ordering (of) things in some meaningful way', e.g. labelling a thermometer as Fahrenheit did. The aim is to point out the relation among certain objects, how far apart they are and in what direction they lie relative to each other. It is not a prerequisite for scaling that variables are cardinal in nature (Dunn-Rankin, 1983: pp. 1–8). Ordinally scaled variables, like political ratings and subjective well-being, can either be rescaled or employed in their original form. Nor are variables required to be expressed in similar units for scaling purposes (Morgan, 1968: pp. 35–37). These are two of the main advantages of composite indexing.

Scaling for composite indexing purposes can be performed in one of four ways. Firstly, there is the option of not scaling variables. This is an especially viable option where variables are already scaled, e.g. reported in percentage terms or some ordinal response scale. Some of the early efforts at composite indexing, however, employ unscaled variables in their original, unadjusted form. This is not desirable. Unscaled variables cannot during aggregation be meaningfully added or subtracted, although multiplication and division are feasible (Morgan, 1968: pp. 35–37).

In the *second* instance, the use of standard scores (z and t values) is also popular in composite indexing. Raw scores are first adjusted for directionality by multiplying each with either +1 or –1. Standardisation then involves transforming raw scores on each indicator into standard scores, e.g. z = (actual score – mean)/standard deviation. Normalised t values are ideal for indexing purposes as they have a mean of zero and standard deviation of one (World Economic Forum, 1996: pp. 36–39). Standard scores can be further adjusted if calculations yield awkward values. Options include the multiplication of all scores by 10 to obtain more visually manageable scores, adding the proportionate share of each component in the composite index to each component score, rounding each score to eliminate decimals, and adding 100 to each score to obtain better indexed scores (Estes, 1984: pp. 169–179). In some cases, standard scores

are expressed relative to that achieved by an individual country, e.g. USA = 100. The selected country may be the one that obtained the highest standard score, a world leader in respect of development, or the country employed as case study in the particular analysis. Standard score scaling is not without its shortcomings. Development data are not always centrally distributed around a mean and the inclusion of outliers substantially influences results (McGranahan et al., 1972: pp. 68–136).

Thirdly, there is the option of transforming variables into ordinal response scales. This may be done either during the survey itself or at a later stage using available data. In the case of Johnston and Sheehy's (1995) Index of Economic Freedom (IEF), for example, experts grade items on a scale of 1 (free) to 5 (least free). Although most ratings are semantically defined, a number of scales are linked to specific ranges of indicator values. Examples include the use of government consumption as percentage of GDP, top tax rates, and average tariff rates to index the degree of government intervention and protectionism (Johnston and Sheehy, 1995: pp. ix–21). Gwartney et al. (1996) employs similar scaling techniques in constructing their Economic Freedom Indices (EFIs).

Finally, there is the conventional linear scaling transformation (LST) method. Variables are scaled from 0 to 100 with the aid of this technique. This requires points of reference relative to which indicators can be scaled (Drewnowski, 1972: pp. 83–85). A minimum and a maximum value are usually identified for each of the variables (Weigel, 1986: pp. 1429–1432). The HDI, for example, scales life expectancy relative to a minimum of twenty-five and maximum of eighty-five years (UNDP, 1996: p. 106). Index values are determined by subtracting the minimum value of the particular variable from its actual value and dividing it by the difference between the selected maximum and minimum values (Thiessen, 1997: p. 142). Some analysts, though, employ only maximum values in scaling. Index values are then calculated by dividing the actual value of the particular variable by its maximum value. This may be advantageous where the minimum and maximum values are so wide apart that they distort index scores. This distortion involves the degree to which scales act as implicit weights in addition to the explicit weights often introduced during the aggregation of index values. The wider the

minimum and maximum values are apart, the higher the implicit weighting and vice versa. This is because it becomes more difficult to achieve the same relative increase in that particular variable. An increasing magnitude of variance in the variable is explained by one percentage point in the particular index (Morris, 1979: pp. 41–56).

The reference points employed in scaling can be selected relative to the observed minimum and maximum values of the particular variable, be it for a specific year or over an extended period of time. Alternatively, reference points can be determined relative to expert expectations of observed minimum and maximum values (Drewnowski, 1974: pp. 19–33; Morris, 1979: pp. 41–56). So, for example, the goalposts employed in estimating the HDI's life expectancy component were set with reference to the highest and lowest observed and expected values during the previous and next thirty years (Ul Haq, 1995: pp. 48–50). In the case of variables where there exists no readily identifiable maximum value (e.g. number of motor vehicles), variables have to be adjusted so as to allow scaling. Variables, for example, can be expressed in population ratios (e.g. motor vehicles per 1000 population) or in terms of terms of percentages of the total population with access to the particular end or means (e.g. percentage of population with access to safe water) (Weigel, 1986: pp. 1429–1432).

The fact that reference points may differ across societies has often been highlighted (Diener and Suh, 1997: pp. 192–200). Thus, index scores for nations at different levels of development may be calculated using different sets of reference points. Rao (1991: pp. 1453–1459), for example, argues for differential scaling to be employed in scaling GDP per capita in the HDI. Countries at different levels of development may aim for targets other than the average of world income formerly employed in scaling. This, however, precludes the comparison of index scores across nations at different levels of development.

Drewnowski (1972: p. 83) argues that scaled values need to take account of inequality. This is to be achieved by multiplying each of the index values with the particular variable's distribution coefficient or its inverse (Drewnowski, 1974: pp. 19–33). It is, however, not always possible to derive such coefficients for all variables because distributional data are not always readily available. An

alternative option is to multiply the composite index value by the Gini coefficient or other general indicator of distribution (Baster, 1985: p. 36). Adjustments for structural differences such as population and country size can also be affected at either the component or the index level (McGranahan et al., 1972: pp. 13–20).

The most important criterion in selecting reference points for scaling is that a balance be found between the width of the range and the spread of index scores (Babbie, 1995: pp. 161–175). Index scores should not be so closely approximated as to prevent nations from being meaningfully distinguished from one another. Nor should the index scores be so widely spread as to keep relatively similar nations from achieving similar scores. So, for example, Streeten (1994: pp. 235–236; 1995a: pp. xii–xiv) argues that index scores based on literacy rates and life expectancy are better balanced than those based on per capita income.

3.2.3. *Weighting and aggregation of component indices*

One also needs to decide on the weighting system and method employed in aggregating component scores into one composite index. In addition to the implicit weights introduced during scaling, explicit weights may be introduced during aggregation. The aim with explicit weighting is that weights should reflect the relative importance of each of the variables and/or components (Drewnowski, 1974: pp. 19–33). The first option, though, is not to employ explicit weights. Here component and index scores are simply averages of the corresponding variable and component scores. Slottje (1991: pp. 686–688) calls this an 'attributes-based' weighting system. Thus, no weighting is introduced other than that implicitly introduced during the scaling of variables.

Where explicit weights are employed, the conventional practice has been one of selecting weights following consultation with experts. Examples of ad hoc, expert-based weighting systems are those employed by Gwartney et al. (1996: pp. 37–41) in estimating their Economic Freedom Indices. So, for example, experts were asked to assign a total of 100 points between index components, after which each component was weighted with the average percentage point it scored in this expert survey in calculating the index value. Weights may also be based on the analyst's percep-

tions of the attitude of policy makers (Drewnowski, 1974: pp. 19–33). Harbison and Myers (1964: pp. 23–24) employed such an ad hoc weighting system in their Human Resource Development Index (HRDI). Tertiary education, in their opinion, was much more important than secondary education in explaining differences in the quality of human capital. Thus, they multiplied the tertiary enrolment rate by five before adding it to the secondary enrolment rate to obtain the composite index value (which counts out of a maximum of 600). Alternatively, weights can be based on explicit agreements between policymakers as to the desirability of different development goals, derived from policy statements of governments and development agencies, or derived from collective utility functions based on individual preference functions of the particular population (Drewnowski, 1974: pp. 19–33). These are all relatively subjective methods of weighting. Each is based on subjective perceptions regarding the relative desirability of certain developmental goals. Indices employing such weighting systems are often singled out for their arbitrariness in weighting (Diener and Suh, 1997: pp. 192–200).

Multivariate techniques present an empirical and relatively more objective option for weight selection. In the case of principal component analysis, components are weighted with the proportion of variance in the original set of variables explained by the first principal component of that particular component. This technique has the advantage of determining that set of weights which explains the largest variation in the original variables (Ram, 1982: pp. 230–232; Slottje, 1991: pp. 686–688). Component scores can also be weighted by their coefficients of correlation or regression with some selected variable not included in the index. Slottje (1991: pp. 686–688) calls this hedonic weighting. The drawback, though, is that multivariate techniques allow one no control over the selection and weighting of components (Ginsberg et al., 1986). Thus, although methodologically sound, these techniques introduce conceptual rigidity in composite indexing.[2]

Since different weighting systems imply different results and, given the subjectivity inherent in many of these weighting systems, no weighting system is above criticism (Morris, 1979: pp. 41–56; Gwartney et al., 1996). It is for this reason that Babbie (1995:

pp. 161–175) argues that equal weighting should be the norm and the burden of proof should fall on differential weighting. According to Wish (1986: pp. 93–97), many of the early composite indices lack an adequate justification for the particular weighting system adopted. Analysts today tend to experiment with a variety of weighting techniques and compare results across these techniques before selecting either one or a combination of techniques in deriving index estimates (Ginsberg et al., 1986: pp. 101–120; Slottje, 1991: pp. 686–688; McGranahan, 1995: p. 4550). Morris (1979: pp. 41–56), however, after experimenting with different weighting systems, found that the different indices remained fairly well correlated. This represents a move towards recognising Wish's (1986: pp. 93–97) call for a balance to be struck between the objective assignment of weights and the role of subjective choice.

As in the case of selection, it often makes sense that different types of societies employ different weighting systems. Nations do not advance equally or uniformly on the path to development (Townsend, 1971: pp. 2–12). Differential weighting may be desirable where countries have already achieved a high success rate in respect of certain variables or where particular variables are culturally less prominent in particular societies (Veenhoven, 1996: pp. 2–5). Yet, differential weighting prohibits meaningful comparisons of index values. Thus, analysts almost invariably employ a uniform weighting system in compiling index values for large samples of countries.

After weights have been assigned to each component index and the component scores weighted accordingly, these scores are aggregated into a composite score. The aggregation of indices tends to be of either an additive or a functional nature. Whereas the former entails the mere addition of component scores to arrive at index values, the latter is based on the estimated functional relationship between certain variables (Adelman and Morris, 1972: pp. 111–112). Beckerman and Bacon (1966) employed the latter type of technique in estimating their Index of Real Consumption. Seven independent economic variables were combined in five different types of equations in order to explain differences in private consumption per capita (the dependent variable). The equation explaining the largest proportion of variance in private consumption

was then employed to arrive at an Index of Real Consumption by substituting the actual values of the seven variables into the equation for each country (Beckerman and Bacon, 1966: p. 524). Stoikov (1967) was very critical of these functional indices. He argued that the determinants employed in selecting the best equation, i.e. multiple regression coefficients and standard errors, are wholely dependent on the number of observations and particular variables selected. Thus, functional indices are empirically biased.

Ideally, composite indices should remain relatively simple in terms of their construction and interpretation (Morris, 1979: pp. 20–40). According to Ginsberg et al. (1986: pp. 101–120), the choice of method employed in weighting and aggregation is ultimately dependent on the nature and scope of the particular study. So, for example, studies aimed at exploring theoretical aspects of composite indexing usually employ rather complex techniques. Studies aimed at presenting a simple and informative view of general well-being or at informing officials regarding particular issues tend to opt for relatively simpler methods, thus allowing for indices to be easily comprehensible and readily calculable.

3.2.4. *Validation of composite indices of development*

Composite indices also need to be validated. Only through continued validation and adjustment resulting from constructive debate can indices be improved. During validation adjustments are effected in selection, scaling, weighting and aggregation in order to improve the quality of the final estimates (Ul Haq, 1995: pp. 54–57). Validation is normally performed by using either item analysis or external validation (Adelman and Morris, 1972: pp. 125–128; Babbie, 1995: pp. 161–175).

Adelman and Morris (1972: pp. 125–128) describe item analysis as an evaluation of the discriminant ability and correlation of component and index scores. Lind (1992) claims that the HDI lacks discriminant validity, because it is too sensitive to small variations in variables. Consider, for example, Denmark, whose GDP per capita considerably exceeds that of Australia. Although the two countries achieved relatively similar scores on the other two components, Denmark finds itself ranked below Australia (Lind, 1992: pp. 96–97). This supports McGranahan's (1995) argument

regarding the undesirability of combining arithmetic (life expectancy and literacy) and geometric (GDP per capita) variables in composite indices employing conventional scaling. This is because the relative rankings of countries at similar levels of development will not correspond (McGranahan, 1995: p. 55). Countries will achieve relatively higher ratings on life expectancy and literacy than on per capita income. This compromises the meaningful comparison of index values. Improvements may be effected by either reselection or rescaling. Correlation analysis is equally useful during validation. Where the correlation between components and index scores or variable and component scores is especially weak, the specific components or variables may be excluded from the index. Profiling can also be employed in validating composite indices. It entails the plotting of component scores for different countries (or the same country over time) and enables one to identify those components of an index that require particular attention and to detect shared patterns of development (McGranahan, 1995: pp. 43–45).

External validation entails an analysis of the relationship between component and index scores and items (validators) not included in the index, e.g. correlating index values with per capita income. External validation is used to distinguish between 'bad' as opposed to 'good' indices. An index is considered 'good' if both the index and the component scores correlate well with the validator. External validation can also be used in distinguishing between 'good' as opposed to 'bad' validators. A good validator is one that correlates well with both the index and component scores (Babbie, 1995: pp. 161–175). Hopkins (1991), for example, identifies economic growth with development success and concludes that the HDI lacks validity since countries such as Sri Lanka boast high HDIs but lack high growth rates. This follows from his argument that improved levels of education are meant to instill entrepreneurship and other desired skills required for improved economic growth (Hopkins, 1991). Yet, insofar as the HDI measures the extent to which the citizens of countries are able to enjoy relatively 'long, educated, comfortable and peaceful lives' (Rao, 1991: p. 1451), it may be considered a valid measurement construct. According to McGranahan et al. (1972: 3) and Srinivasan (1994b: pp. 240–242), external validation is hampered by the fact that there is no

objective and universal validator against which composite indices can be validated.

Ultimately, good composite indices need to be sufficiently sensitive to differences in the particular measurement constructs, reliable in terms of repeated measures yielding stable results, and accurate in being free from systematic error resulting from the omission of influences whose effects are presumed to be incorporated in the index (Adelman and Morris, 1972: pp. 125–128).

The variety of composite indices presented in academic articles and statistical publications is summarised in Table II, indicating in each case the particular methods employed in selection, scaling, weighting and aggregation. The number of variables and component indices included in each of the composite indices is also noted, as is the number of countries for which index values were reported. Only composite indices for which details on each of these elements were available are listed here. This, therefore, is not an exhaustive but nonetheless relatively comprehensive list of composite indices of development. The twenty indices are ordered more or less chronologically. In the light of Table II, composite indices generally seem to be additive ones with equally weighted components consisting of variables selected in an ad hoc manner.

The multidimensionality of the composite indices presented in Table II is illustrated in Table III. Where a variable or component included in the particular index applied directly to a specific dimension, it was accordingly noted, highlighting the multidimensional nature of composite indices of development. Only a limited number of composite indices listed in Table II cover less than three of the twelve different dimensions of development identified here.

Several criticisms are leveled at composite indices for not representing improved measures of well-being.[18] Not one single element of the methodology of composite indexing is above criticism. Ram (1982) reiterates this and puts it down to various reasons. Apart from the index number problem inherent in any construction of composite indices, the underlying data often remain deficient, results remain sensitive to the methods employed in scaling, weighting and aggregation, as well as the order in which the steps in composite indexing are completed, and results remain sample-specific (Ram, 1982: pp. 237–239). Accordingly, Diewert (1986, as quoted in Slottje,

TABLE II

Chronological summary of twenty composite indices of development

Description	Number of variables (components)[3]	Selection method[4]	Scaling method[5]	Weighting method[6]	Aggregate format[7]	Country coverage
1. Combined Consumption Level Index (Bennett, 1951)	19 (6)	ad hoc	0–100 indexed scores	equal average	additive	31
2. Human Resources Development Index (HRDI) (Harbison and Myers, 1964)	2	ad hoc	none	ad hoc	additive	75
3. Real Index of Consumption (Beckerman and Bacon, 1966)	7	ad hoc	none	correlation coefficients	functional	80

TABLE II

Continued

Description	Number of variables (components)[3]	Selection method[4]	Scaling method[5]	Weighting method[6]	Aggregate format[7]	Country coverage
4. UNRISD General Index of Development (GID) (McGranahan et al., 1972)[8]	18 (2)	correlation matrix (73)	0–100 indexed socres	correlation coefficients	additive	58
5. Physical Quality of Life Index (PQLI) (Morris, 1979)	3	ad hoc	0–100 indexed socres	equal average	additive	150
6. Composite Basic Needs Indices (Ram, 1982)	5	ad hoc	none and 0–100 indexed scores	principal component analysis	additive	82

TABLE II

Continued

Description	Number of variables (components)[3]	Selection method[4]	Scaling method[5]	Weighting method[6]	Aggregate format[7]	Country coverage
7. Index of Social Progress (ISP) (Estes, 1984)[9]	44 (11)	ad hoc	standard scores	equal average	additive	107
8. World Standard Distance Scales (Ginsberg et al., 1986)[10]	42/22 (5)	ad hoc	standard deviation multiples	equal average	additive	143
9. Human Suffering Index (HSI) (Camp and Speidel, as related in Hess, 1989)	10	ad hoc	0–10 metric scales	equal average	additive	60 plus
10. Quality of Life Rankings (Slottje, 1991)	20	ad hoc	none	mixed	functional	126

TABLE II

Continued

Description	Number of variables (components)[3]	Selection method[4]	Scaling method[5]	Weighting method[6]	Aggregate format[7]	Country coverage
11. Combined Quality of Life Indices (CQLI) (Diener, 1995)[11]	14 (7)	ad hoc	standard scores	equal average	additive	77
12. Index of Economic Freedom (Johnston and Sheehy, 1995)	10	ad hoc	1–5 metric scales	equal average	additive	100 plus
13. Economic Freedom Indices (EFIs) (Gwartney et al., 1996)	17 (4)	ad hoc	0–10 metric scales	equal impact and ad hoc	additive	103
14. Human Development Index (HDI) (UNDP, 1996)[12]	4 (3)	ad hoc	0–100 indexed scores[13]	equal average and ad hoc[14]	additive	174

TABLE II

Continued

Description	Number of variables (components)[3]	Selection method[4]	Scaling method[5]	Weighting method[6]	Aggregate format[7]	Country coverage
15. Capability Poverty Measure (CPM) (UNDP, 1996)	3	ad hoc	none	equal average	additive	101
16. Gender-related Development Index (GDI) (UNDP, 1996)	4 (3)	ad hoc	adjusted 0–100 indexed scores[15]	equal average	additive	163
17. Gender Empowerment Measure (GEM) (UNDP, 1996)	4 (3)	ad hoc	adjusted 0–100 indexed scores	equal average	additive	100

TABLE II

Continued

Description	Number of variables (components)[3]	Selection method[4]	Scaling method[5]	Weighting method[6]	Aggregate format[7]	Country coverage
18. Global Competitiveness Indices (World Economic Forum, 1996)[16]	155 (8)	data coverage (300+)	standard scores	ad hoc and correlation coefficients	additive	49
19. Human Poverty Index (HPI-1) for developing countries (UNDP, 1999)	5 (3)	ad hoc	none	equal average	functional	92
20. Human Poverty Index (HPI-2) for developed countries (UNDP, 1999)	4	ad hoc	none	equal average	functional	17

TABLE III

Dimensions of twenty-two composite indices of development

Description	A	B	C	D	E	F	G	H	I	J	K	L
1. Combined Consumption Level Index	x	x	x	x		x	x				x	
2. Human Resources Development Index		x										
3. Real Index of Consumption				x							x	
4. General Index of Development	x	x	x	x					x	x	x	
5. Physical Quality of Life Index	x	x	x									
6. Composite Basic Needs Indices	x	x	x	x					x			
7. Index of Social Progress	x	x	x	x	x	x		x	x		x	
8. World Standard Distance Scales	x			x					x			
9. Human Suffering Index	x	x	x	x				x	x	x		
10. Quality of Life Rankings	x	x	x	x	x			x	x	x		
11. Combined Quality of Life Indices		x	x	x	x	x	x	x	x		x	
12. Index of Economic Freedom												x
13. Economic Freedom Indices								x	x			x
14. Human Development Index	x	x							x			
15. Capability Poverty Measure		x	x									
16. Gender-related Development Index	x	x							x		x	
17. Gender Empowerment Measure								x	x	x		
18. Global Competitiveness Indices				x				x	x			
19. Human Poverty Index (developing)	x	x	x	x								
20. Human Poverty Index (developed)	x	x								x	x	

Note: The dimensions of development distinguished here are: A – demographic dynamics, B – education, training and knowledge, C – health, food and nutrition, D – human settlement, infrastructure and communication, E – political and social stability, F – culture, social fabric and family values, G – environmental resources and pressures, H – political and civil institutions, I – income and economic growth, J – unemployment and labour utilisation, K – poverty and inequality, L – economic freedom.[17]

1991: pp. 684–685) claims that the best one can do is to continue searching for a composite index that balances the need for conceptual clarity and methodological simplicity. The following is a brief discussion of the main lines of critique levelled at composite indices.

Firstly, it is often argued that particular indices exclude one or more essential components of development. Perthel (1981: pp. 6–7), for example, criticises the UNRISD's General Index of Develop-

ment (GID) for its failure to adjust index values for differences in population structure and inequality and its exclusion of development domains such as justice and violence. An index equally criticised for its limited dimensionality is the Physical Quality of Life Index (PQLI). Two of its three components are health-related (Todaro, 1989: pp. 109–112; Felipe and Resende, 1996: pp. 187–190). Morris (1979: pp. 41–56) in defense argues that adults die for different reasons than infants and that the two health variables were justifiably included. It may also be argued that the limited availability of data at the time gave analysts a limited choice as to variables suitable for composite indexing (Todaro, 1989: pp. 109–112). One of the most common criticisms levelled at the HDI is that it excludes other social achievements crucial to the quality of life (Lind, 1992: pp. 96–97), most notably political freedom and human rights (Srinivasan, 1994b: pp. 240–242). The needs that the HDI does include, can be equally well met in prisons and under authoritarian rule (Hopkins, 1991: pp. 1469–1473). Streeten (1994: p. 236; 1995a: p. xiv), though, argues against the addition of political and human rights variables to the HDI. He cites four reasons for this. In the first instance, this construct is too important to simply trade off against the other components of the HDI. In the second instance, these variables are much more volatile than the three other components. This may severely affect the inter-temporal stability of index values. In the third instance, these variables are less objectively measured than the others included in the index, thus compromising the overall objectivity of the index. In the final instance, the nature of the relation between political freedom and human development is not that clear-cut.

Secondly, it may be argued that particular components of indices be quantified with the aid of different variables (Streeten, 1995a: p. viii). Anand and Sen (1993, as quoted in Streeten, 1995b: p. 26), for example, call for the HDI to be differentiated for developed and developing countries. They argue that it should be left unchanged for low-income economies but could benefit from being supplemented with tertiary enrolment ratios and infant mortality rates, respectively, in the case of high and middle-income economies. The same applies to environmental indicators. As a result of differences in the level of economic development and geophysical character-

istics, countries have to cope with different environmental problems (McGranahan et al., 1972: pp. 4–6). So, for example, industrialised countries are more concerned with pollution emissions and LDCs are more concerned with the destruction of arable land (Pomfret, 1997: p. 278).

Thirdly, indices are often criticised for being unable to reveal anything that a single variable (and particularly per capita income) alone cannot reveal. This argument often is employed by proponents of income-based indicators. Stewart (1985), McGillivray (1991) and Srinivasan (1994a) point to the high correlation between the HDI's three components. Felipe and Resende (1996: pp. 187–190) highlight the high degree of multicollinearity to which the Physical Quality of Life Index (PQLI) is subject. Social indicators, in fact, are often highly correlated with economic indicators of development (Diener and Suh, 1997: pp. 192–200). Ogwang (1994: pp. 2011–2014), furthermore, goes so far as to argue that indices such as the HDI reveal nothing that per capita income and life expectancy alone would not have revealed, except perhaps in the case of a few unique countries where rankings differ substantially.

Fourthly, ad hoc selection is subject to severe criticism. The HDI, for example, has been made out to be a politically motivated index intended to boost the development rankings of countries making concerted efforts at addressing health and education backlogs (Todaro, 1994: pp. 63–66; Elkan, 1995: p. 12). Selection may also be biased insofar as technical criteria such as the availability and accuracy of data alone drive the selection process. Ideological considerations may also introduce bias in selection. Examples include accepted ideas regarding the desirability of political democracy and gender equality (Townsend, 1971: pp. 2–12; Anand and Ravallion, 1993: pp. 136–138; Greeley, 1994: p. 57).

In the fifth instance, the accuracy and comparability of the data employed in composite indices are often criticised. Lind (1992: p. 96), Ogwang (1994: p. 2011) and Srinivasan (1994b: pp. 240–242) consider the HDI empirically unsound and conceptually weak given the measurement errors, biases and incomparability inherent in the underlying data. Some of the underlying variables are estimates based on mathematical extrapolations rather than actual observations. Others are based on different definitions of similar concepts

(Srinivasan, 1994b: pp. 240–242). Perthel (1981: pp. 6–7) comes to a similar conclusion regarding the UNRISD's General Index of Development (GDI). Most of the indicators related in these pages, all of which are or have the potential of being employed in composite indexing, are subject to these shortcomings.

In the *sixth* instance, there is criticism against the weighting and aggregation techniques employed in composite indexing. According to Todaro (1989: pp. 109–112) and Rao (1991: pp. 1453–1459), equal average weighting systems such as those employed in the majority of indices related in Table II are not adequately justified. Often, no clear rationale is presented for preferring them to alternative weighting systems. Hopkins (1991), furthermore, argues that additive aggregation implies that one can measure apples and oranges individually and then aggregate them into some meaningful index.

Finally, there are those criticisms aimed at the supposed lack of practical value of composite indices. Rao (1991: pp. 1453–1459) claims that composite indices such as the HDI give no specific and focused policy advice. Srinivasan (1994b: pp. 240–242) adds that the HDI has failed to bring about any significant changes in development policy. It is important, though, to note that many composite indices were never developed with this purpose in mind. Some were intended as tools for theoretical analysis whilst others were presented as alternative, simple and more visible ways of reporting on the development situation.

In the light of these criticisms of composite indices, one can reflect as follows on the four aspects of technique and method identified in Table I. The composite indices presented in Table II are quantitative insofar as the indices are all presented in numerical format. Indices are subject to *subjectivity* despite the *objectivity* of the methods employed in composite indexing. Principal component analysis, for example, presents a relatively objective means of selecting and weighting variables during composite indexing, but subjectivity may be introduced during the selection, scaling, weighting and/or aggregation phases of composite indexing. Given the general characteristics of the indices presented in Table II, the inherent subjectivity of most indices is evident. The often-employed ad hoc selection of components and variables is the most prominent

of these sources of subjectivity. Weighting and aggregation methods also, in general, remain relatively subjective. Diener and Suh (1997: pp. 206–210) call for the use of a combination of methods during selection, scaling, weighting and aggregation in order to improve the objective quality of composite indices.

The composite indices presented in Table II are of a *cardinal* nature in the sense that these indices reflect the magnitude of differences between certain entities in terms of index values. Each ranks nations relative to each other in terms of some combination of variables. Yet, the indices remain ordinal insofar as the magnitude of these differences cannot be interpreted meaningfully. So, for example, it is not possible to interpret directly the meaning of country A achieving a rating 0.06 higher than country B on the HDI. The ordinal nature of composite indices also represents a drawback insofar as index values in themselves have no clear meaning, thus limiting their interpretative value (Hopkins, 1991; Veenhoven, 1996: pp. 2–5). Others, though, have argued the converse. They claim that composite indices enhance research by aggregating variances and reducing data to new and more efficient units for comparison and analysis (Babbie, 1995: pp. 161–175; Diener and Suh, 1997: pp. 206–210).

The *multidimensionality* of composite indices represents one of their main advantages. Indices represent aggregate and relatively simple measures of a combination of complex development phenomena (Diener and Suh, 1997: pp. 206–210). Yet, Khan (1991, as quoted in Kallmann, 1997: p. 13) maintains that single index values are not feasible. He argues that single index values contradict the complex and interdependent nature of the development process and remain arbitrary due to the continued debate as to the best methods of scaling and weighting.

3.3. *Composite Indices of Development in Comparative Application*

Comparative application of composite indices of development remains problematic. To a large extent, this results from the various shortcomings of composite indices related above. Comparison bias is inherent in composite indices insofar as they are but summary measures of a number of individual survey-based indica-

tors. Comparisons of composite indices are also problematic due to other reasons. On the one hand, differences in the same index cannot be meaningfully compared across time and space. *Inter-temporal* bias results from (i) differences in the relative importance of the variables and components included in the index (Veenhoven, 1996: pp. 2–5), (ii) changes in the variables and components included in indices, and (iii) changes in the methods employed in weighting and aggregating index values. *Interspatial* comparisons of the same index are equally problematic. Stewart (1985: pp. 54–66) argues that composite indices conceal inequalities in the particular components and variables, thus restricting their value in comparative analyses. Townsend (1971: pp. 14–24) and Streeten (1995a: pp. x–xii) come to a similar conclusion. Ul Haq (1995: 50–54), though, maintains that composite indices remain useful in comparative analyses since index values are distributed less skew than the un-indexed variables. In the case of the HDI, for example, per capita income ranges between $100 and $40,000, whereas the income component index varies only between 0 and 100. He also points out that a more meaningful comparison and application of composite indexing can be achieved by estimating disaggregated index values (e.g. by gender, ethnicity or regional location) where the specific variables are available in disaggregate form.

On the other hand, different indices cannot be meaningfully compared, neither across space nor across time, given differences in content and methodology. Slottje (1991) illustrates the substantial variance in indices employing different weighting and aggregation methods. As a result of the ordinal nature of composite indices, indices are also problematic in terms of their interpretative value during comparative analyses. The standard practice in comparing composite indices, either with other indices or with individual indicators, is one of comparing the differences in the rankings nations achieve on different indices and/or indicators. Elkan (1995: p. 13), for example, compares rankings on the HDI and GNP per capita for a sample of 132 countries. At the extremes of this comparison the UAE's ranking on GNP per capita exceeded its HDI ranking by forty-three positions (highest downward ranking), while Sri Lanka achieved a ranking on the HDI thirty places higher than on GNP per capita (highest upward ranking). Apart from conceptual and

methodological differences and disparities in the quality of the underlying data, differences in rankings may also be attributable to other factors. These include differences in culture and population size. Smaller countries, for example, often achieve substantial upward rankings. Arabic nations tend to be ranked downward on the HDI, given the extent to which the gender inequalities endemic in their cultural milieu push their scores on the literacy and life expectancy component indices down. According to Beckerman and Bacon (1966: pp. 521–529), furthermore, comparisons may be less meaningful where the sample of nations composite indices are compared for is not representative of the total income range. Thus, comparisons of index values conceal differences originating from a variety of sources and are not useful in detailed micro-level comparisons.

In terms of the distinction between *absolute* as opposed to *relative* indicators, composite indices are all relative in nature. The primary aim, in fact, with composite indexing is to rate and compare nations relative to some standard of achievement. This standard is set relative to the limits of achievement employed during the scaling, weighting and aggregation of composite indices.

3.4. *Focus of Composite Indices of Development*

Composite indices generally combine measures of ends and means (Ul Haq, 1995: pp. 50–54). There has, however, been some debate as to whether indices should be based on ends, means or a combination of the two. Morris (1979: pp. 94–96) argued that indices based on ends alone lack policy relevance insofar as policies and programmes are often defined in terms of means rather than ends. Morris (1979) consequently combined variables of ends and means in his Physical Quality of Life Index (PQLI) so as to highlight the important synergy between ends and means. Another reason for employing indicators of both means and ends in composite indices is that certain ends are means in themselves, thus complicating the distinction between means and ends. Veenhoven (1996: p. 5), however, claims that such combination of ends and means lacks theoretical justification and political relevance. Diener and Suh (1997: pp. 192–200) expand on this criticism of ends and means-based indices. According to them such practice complicates the value of indices during causal analyses. Accordingly, Veenhoven (1996: p. 5) argues

in favour of ends-based measures since these can be directly related to policy goals. In fact, indicators are often developed expressly with the purpose of monitoring the impact of interventions targeting specific issues, e.g. primary health care (infant mortality rates), transportation (kilometres of paved roads per square kilometre), and crime (homicides and rapes per 100,000 of the population).

3.5. *Clarity, Simplicity and Flexibility of Composite Indices of Development*

In respect of method and technique, composite indexing is relatively complex. Yet, indices such as the PQLI and HDI are often applauded for their relative simplicity (Todaro, 1989: pp. 109–112). In general terms, indices based on ad hoc selection, traditional 0–100 scaling, equal weighting and additive aggregation are simpler than those employing multivariate techniques in selection and weighting, standard scores in scaling, and functional aggregation. Contrary to the claim of proponents of composite indexing that index values are relatively simpler to interpret than the underlying variables (Ul Haq, 1995: pp. 50–54), the interpretative value of composite indices remains a contested issue. In fact, the multiplicity and multidimensionality of composite indices are often impediments rather than aids in interpretation (Ram, 1982: p. 228). Composite indices, insofar as they represent an empirical transformation and collation of data, are relatively flexible (Ul Haq, 1995: pp. 46–48). Changes in selection, scaling, weighting and aggregation can be effected readily, albeit at the cost of comparability.

3.6. *Availability of Composite Indices of Development*

The composite indices listed here perform relatively well in terms of cross-national availability (Table II). The smallest and largest samples, respectively, report index estimates for thirty-one and 174 nations. Cross-national availability has improved over time as the availability of the underlying data has improved. The only indices that perform relatively well in terms of inter-temporal availability are the UNDP's Human Development Index (HDI), the Economic Freedom Indices reported by the Liberty Fund, and the World Economic Forum's Global Competitiveness Indices. HDI estimates are available for 1990 to 1997, while the other two indices have

been published annually since the 1980s. Subsequent estimates, though, are often based on a revised index content and/or measurement methodology. The other indices are primarily confined to the specific time frames to which the particular analyses apply. Ul Haq (1995: p. 46) puts this down to the fact that other composite indices have either been methodologically unsound or were never applied for a sufficient period of time to allow them to be further developed, refined and tested so as to remain prominent. Rao (1991: p. 1459) suggests greater collaboration and coordination between agencies producing reports on composite indices. He calls for reports that are coordinated in terms of timing, content, frequency and focus.

4. CONCLUSION

On the strength of the systematic positive association between income and popular composite indices such as the HDI, many have claimed that these indices represent no real contribution to the literature on indicators research. Composite indices, furthermore, are often considered to be ideological statements rather than practically functional indicators. Yet, composite indices represent useful supplements to income-based development indicators. These indices remain invaluable in terms of their ability to simplify complex measurement constructs, to focus attention and to catch the eye, thus enhancing their political appeal.

NOTES

[1] The distinction between uni-dimensional as opposed to multi-dimensional measures refers to the difference between indicators individually measuring one particular component of development (e.g. the unemployment rate) as opposed to indicators combining a number of individual indicators into one composite (e.g. the Human Development Index).

[2] In a departure from conventional multivariate weighting techniques, countries may be treated as variables (or components), and variables (or components) as observations. Matrices can then be transposed to determine the orthogonal factors to be used as weights. Factor loadings here display the extent to which each country is associated with each variable (or component) (Ginsberg et al., 1986: pp. 49-65). Thus, countries rather than index components are weighted.

[3] Where only variables are listed, it means that either each of the component indices of the composite index is represented by a single indicator or that the composite index is not made up of individual component indices. In some cases, though, certain of the component indices of the composite index include more than one indicator. Here the number of component indices is noted in parentheses. Due to lack of space, details as to the specific variables included in each of these indices are not related here.

[4] *Ad hoc* selection is based on theory, expert opinions and/or intuitive appeal. Where bivariate and multivariate techniques were employed in selection, the original number of variables included in the particular analysis is noted in parenthesis.

[5] *None* means that the variables were employed in their original form in calculating index values and were not rescaled using any of the techniques normally employed in rescaling variables.

[6] *Equal average* weighting refers to the use of no explicit weighting system in aggregation. *Ad hoc* weighting is based on expert opinions and/or intuitive appeal.

[7] During *additive aggregation* composite scores are calculated by adding up each of the component scores. In the case of *functional aggregation*, composite scores are calculated by combining the component scores in some functional equation.

[8] Two earlier composite indices developed by the UNRISD include the Level of Living (LoL) and Social Development (SDI) Indices. These two indices respectively included twenty and sixteen variables and were combined into six and two component indices each (UNRISD, as quoted in Todaro, 1989: p. 108 and Felipe and Resende, 1996: p. 187).

[9] In the field study preceding its wider application this index was called the Index of Social Vulnerability (NSV) (Estes, 1984: pp. 2–19). A variant of ISP, the Index of Net Social Progress (ISNP), is based on the exclusion of the geographic component index (Estes, 1984: pp. 169–179).

[10] These scales consist of a Composite (RICHDEX) and Growth Potential (GROTPOT) Index that respectively employ forty-four and twenty-two variables (Ginsberg et al., 1986).

[11] CQLI consists of a Basic Value (BQLI) and Advanced Value (AQLI) Quality of Life Index. Each employs seven variables selected in an attempt to discriminate between developing and industrialised countries in terms of the same general domains of development (Diener, 1995: pp. 107–125).

[12] The first refined version of the HDI was published in 1992, although the groundwork was laid in the late 1980s (Ul Haq, 1995: pp. 46–48).

[13] With the exception of GDP per capita, the HDI employs conventional scaling. Up to 1998 it employed Atkinson's formula in transforming each country's per capita GDP relative to the threshold level of the world average income (UNDP, 1996: p. 106). Since 1999 the conventional LST technique is used but with the difference that each value is transformed into logs (UNDP, 1999: p. 159).

[14] The component index for educational attainment weights adult literacy at two thirds and the combined enrolment ratio at one third (UNDP, 1996: p. 106).

[15] The achievement on each of the three HDI component indices is adjusted for the observed disparity between men and women (UNDP, 1996: pp. 107–108).
[16] Some weights are assigned in ad hoc fashion, given assumptions regarding the relative reliability of data. Two variants of the Competitiveness Index (CI) are estimated. The Growth Index (GI) combines the CI with per capita income. The Market Growth Index (MGI) combines the GI with the size of the economy in terms of each country's contribution to overall world economic growth (World Economic Forum, 1996: pp. 5–13, 36–39).
[17] Economic freedom relates to the 'extent to which rightly acquired property is protected and individuals are free to engage in voluntary transactions' (Gwartney et al., 1996: p. 12). It is measured in terms of a variety of elements of economic systems. These include trade policy, taxation, government consumption, monetary policy and inflation, capital flows and foreign investment policy, banking, wage and price controls, property rights, government regulation, and black marketeering (Johnston and Sheehy, 1995: pp. ix–21; Gwartney et al., 1996: pp. 1–46).
[18] This discussion is often informed by discussions on the Human Development Index (HDI). This is not intended to represent either a wholesale condemnation or approval of the HDI. It is simply because no other composite index, except perhaps the Physical Quality of Life Index (PQLI), has provoked such extensive debate as to its relative merit.

REFERENCES

Adelman, I. and C.T. Morris: 1972, 'The measurement of institutional characteristics of nations: Methodological considerations', Journal of Development Studies 8, pp. 111–135.

Anand, S. and M. Ravallion: 1993, 'Human development in poor countries: On the role of private incomes and public services', Journal of Economic Perspectives 7, pp. 133–150.

Babbie, E.: 1995, The Practice of Social Research (Wadsworth, Washington, pp. 161–175).

Baster, N.: 1985, 'Social indicator research: Some issues and debates', in J.G.M. Hilhorst and M. Klatter (eds.), Social Development in the Third World: Level of Living Indicators and Social Planning (Croom Helm, London, pp. 36–43).

Beckerman, W. and R. Bacon: 1966, 'International comparisons of income levels: A suggested new measure', Economic Journal 76, pp. 519–536.

Bennett, M.K.: 1951, 'International disparities in consumption levels', American Economic Review 41, pp. 632–649.

Child, D.: 1970, Essentials of Factor Analysis (Holt, Reinhart and Winston, London, pp. 1–10).

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

Diener, E. and E. Suh: 1997, 'Measuring quality of life: Economic, social, and subjective indicators', Social Indicators Research 40, pp. 189–216.

Drewnowski, J.: 1972, 'Social indicators and welfare measurement: remarks on methodology', Journal of Development Studies 8, pp. 77–90.
Drewnowski, J.: 1974, On Measuring and Planning the Quality of Life (Mouton, Paris, pp. 19–33).
Dunn-Rankin, P.: 1983, Scaling Methods (Lawrence Erlbaum, London, pp. 1–27).
Elkan, W.: 1995, An Introduction to Development Economics (Harvester Wheatsheaf, London, p. 12).
Elster, J. and J.E. Roemer (eds.): 1991, Interpersonal Comparisons of Well-Being (Oxford University Press, New York, pp. 5–15).
Estes, R.J.: 1984, The Social Progress of Nations (Praeger, New York, pp. 1–179).
Everitt, B.S. and G. Dun: 1991, Applied Multivariate Data Analysis (Edward Arnold, London, pp. 97–126).
Felipe, J. and M. Resende: 1996, 'A multivariate approach to the measurement of development: Asia and Latin America', Journal of Developing Areas 30, pp. 183–210.
Ginsberg, N., J. Osborn and G. Blank: 1986, Geographic Perspectives on the Wealth of Nations, Department of Geography Research Paper No. 220 (University of Chicago, Chicago, pp. 17–120).
Greeley, M.: 1994, [1] Measurement of poverty and poverty of measurement', Institute of Development Studies Bulletin 25, pp. 50–58.
Gwartney, J., R. Lawson and W. Block: 1996, Economic Freedom of the World (1975–1995) (Free Market Foundation, Johannesburg, pp. 1–46).
Harbison, F. and C.A. Myers: 1964, Education, Manpower, and Economic Growth (McGraw-Hill, New York, pp. 23–24).
Hess, P.: 1989, 'The military burden, economic growth, and the human suffering index: evidence from the LDCs', Cambridge Journal of Economics 13, pp. 497–515.
Hopkins, M.: 1991, 'Human development revisited: A new UNDP report', World Development 19, pp. 1469–1473.
Johnston, B.T. and T.P. Sheehy: 1995, The Index of Economic Freedom (Heritage Foundation, Washington, pp. ix–21).
Kallmann, K. (ed.): 1997, An Introduction to Social Indicators and Social Reporting: A Digest of the Literature, Centre for Social and Development Studies Research Report No. 9, (University of Natal, Durban, pp. 13–18).
Lind, N.C.: 1992, 'Some thoughts on the Human Development Index', Social Indicators Research 27, pp. 89–101.
McGillivray, M.: 1991, 'The Human Development Index: Yet another redundant Composite Development Indicator?', World Development 19, pp. 1461–1468.
McGranahan, D.V., Richard-Proust, C., Sovani, N.V. and Subramanian, M.: 1972, Contents and Measurement of Socioeconomic Development, A Staff Study of the United Nations Research Institute for Social Development (UNRISD) (Praeger, New York, pp. 3–136).
McGranahan, D.: 1995, 'Measurement of development: Research at the UN's Research Institute for social development', International Social Science Journal 143, pp. 39–57.

Morgan, J.N.: 1968, 'The supply of effort, the measurement of well-being, and the dynamics of improvement', American Economic Review 58, pp. 31–39.
Morris, M.D.: 1979, Measuring the Condition of the World's Poor: The Physical Quality of Life Index, Pergamon Policy Studies No. 42, (Pergamon Press, New York, pp. 20–56).
Ogwang, T.: 1994, 'The choice of principal variables for computing the Human Development Index', World Development 22, pp. 2011–2014.
Perthel, D.: 1981, Laws of Development, Indicators of Development, and their Effects on Economic Growth, Doctoral Thesis (University of Amsterdam, Amsterdam, pp. 6–7).
Pomfret, R.: 1997, Development Economics (Prentice Hall, London, p. 278).
Ram, R.: 1982, 'Composite indices of physical quality of life, basic needs fulfilment, and income: a principal component representation', Journal of Development Economics 11, pp. 227–247.
Rao, V.V.B.: 1991, 'Human Development Report 1990: Review and assessment', World Development 19, pp. 1451–1460.
Rao, M.V.S., K. Porwit and N. Baster: 1978, Indicators of Human and Social Development: Report on State of the Art (United Nations University, New York, pp. 31–117).
Rose, R.: 1995, 'Making progress and catching up: Comparative analysis for social policy-making', International Social Science Journal 143, pp. 113–125.
Sainz, P.: 1989, 'An index of social welfare', in F. Bracho (ed.), Towards a New Way to Measure Development, Report on the International Meeting on More Effective Development Indicators, July 31 to August 3 (Caracas, Venezuela, pp. 156–160).
Slottje, D.J.: 1991, 'Measuring the quality of life across countries', Review of Economics and Statistics 73, pp. 684–693.
Srinivasan, T.N.: 1994a, 'Data base for development analysis: An overview', Journal of Development Economics 44, pp. 3–27.
Srinivasan, T.N.: 1994b, 'Human development: A new paradigm or reinvention of the wheel?', American Economic Review (Papers and Proceedings) 84, pp. 238–243.
Stewart, F.: 1985, Planning to Meet Basic Needs (Macmillan, London, pp. 1–66).
Stoikov, V.: 1967, 'International comparisons of income levels', Economic Journal 77, pp. 939–944.
Streeten, P.: 1994, 'Human development: Means and ends', American Economic Review (Papers and Proceedings) 84, pp. 232–237.
Streeten, P.: 1995a, 'Forward', in M. Ul Haq (ed.), Reflections on Human Development (Oxford University Press, New York, pp. viii–xiv).
Streeten, P.: 1995b, 'Human development: The debate about the index', International Social Science Journal 143, pp. 25–37.
Thais, L.: 1989, 'Reflections on an Alternative Development Indicator', in F. Bracho (ed.), Towards a New Way to Measure Development, Report on the International Meeting on More Effective Development Indicators, July 31 to August 3 (Caracas, Venezuela, p. 174).

Thiessen, H.: 1997, Measuring the Real World: A Textbook of Applied Statistical Methods (John Wiley & Sons, Chichester, pp. 13, 142).
Todaro, M.P.: 1989, Economic Development in the Third World (Longman, New York, pp. 108–113.
Townsend, P. (ed.): 1971, The Concept of Poverty: Working Papers on Methods of Investigation and Life-Styles of the Poor in Different Countries (Heinemann, London, pp. 2–24).
Ul Haq, M.: 1995, Reflections on Human Development (Oxford University Press, New York, pp. 46–57).
United Nations Development Programme: 1996, Human Development Reports 1996 (Oxford University Press, New York, pp. 106–108).
United Nations Development Programme: 1999, Human Development Reports 1999 (Oxford University Press, New York, pp. 159–163).
United Nations Development Programme: 1997, Glossary of Poverty and Human Development, Available: http://www.undp.org/undp/news/HDR97/pr7-eng.html (14 October 1998).
Veenhoven, R.: 1996, 'Happy life-expectancy: A comprehensive measure of quality-of-life in nations', Social Indicators Research 39, pp. 1–58.
Weigel, Van B.: 1986, 'The basic needs approach: Overcoming the poverty of Homo Oeconomicus', World Development 14, pp. 1423–1434.
Wilson, R.K. and C.S. Woods: 1982, Patterns of World Economic Development (Longman Sorrett, Melbourne, p. 11).
Wish, N.B.: 1986, 'Are we really measuring the quality of life?', American Journal of Economics and Sociology 45, pp. 93–99.
World Economic Forum: 1996, Global Competitiveness Report 1996 (WEF, Geneva, pp. 5–39).

Department of Economics
University of the Free State
Bloemfontein
South Africa