ABSTRACT / The widespread interest in the concept of sustainable environment and development has been accompanied by the need to develop useful systems of measurement. We discuss the use of indicators which might be used to assess such conditions. Our characteristics, or criteria, for desirable global sustainability indicators are:

- sensitivity to change in time
- sensitivity to change across space or within groups
- predictive ability
- availability of reference or threshold values
- ability to measure reversibility or controllability
- appropriate data transformation
- integrative ability
- relative ease of collection and use

We discuss the basis of these characteristics, and examine two categories of indicators (soil erosion and population) and two specific indicators (physical quality of life index and energy imports as a percentage of consumption) for their value as sustainability measures.
to assess global and regional conditions and trends (see Table 1).

In the most recent State of the World report, Brown (1987) uses two key indicators, per capita gross world product and per capita fossil fuel consumption, to measure trends in economic growth. He suggests that society becomes nonsustainable when the costs of this growth in terms of forest damage, climate change, and soil erosion, for example, exceed the benefits provided. Brown believes that there are important threshold values for such variables as tree tolerance to pollution damage, tropical forest soil moisture, and atmospheric carbon dioxide, where the system becomes unsustainable in terms of forest damage, forest fires, soil formation, and the functioning of the carbon cycle.

He also uses the measures of carrying capacity developed by the UN Food and Agriculture Organization and the World Bank to examine the prospects for sustainable agriculture and fuel use in developing countries. The studies he cites calculate the potential population that can be supported by agricultural land and forests in different ecological zones. If the current population exceeds the supportable population, then this is taken as an indicator of a nonsustainable society.

Clark (1986) introduces the topic of the “sustainable development of the biosphere” with a series of density measures. He uses the indicators of population density (people per unit area), agricultural production density (value added per hectare per annum), and energy consumption density (oil equivalents per hectare per year) to compare the sustainability of regions and nations.

Heilbroner (1974) sees the priority indicators for assessing future trends as oil prices, economic growth rates, and environmental pollution. Brubaker (1972) provides a classification system for measuring environmental threats which includes pollution, erosion, climate change, and overpopulation. And the World Conservation Strategy (IUCN 1980) uses topsoil and farmland loss, deforestation, water supply disruption, and fishery pollution and depletion, as indicators of nonsustainable resource utilization.

Although these recent contributions imply that progress is being made in the development and critical analysis of sustainability indicators, in many cases the existing or proposed indicators are not the most sensitive or useful measures. The developers of these indicators are relying on easily available data, and often fail to use fairly simple transformations or ratios which would provide clearer insights into questions of change, thresholds, and other nonsustainable trends.

Critics of Malthusian thinking (for example, George and Paige 1982) complain about using population, population growth, and even population-resource ratios as measures of sustainability. These critics claim that income, land ownership, and access to technology are more significant determinants of an individual’s or a society’s pressure on resources than are absolute population levels.

Ecologists have also criticized existing criteria and have suggested alternative measures. Westman (1977) points out that biological organisms are often much more efficient and integrative indicators of water pollution than are measurements of pollutant concentration in water. Smith and Theberge (1987), in discussing criteria for evaluating natural areas, point out the problems in using diversity and fragility as indicators of the stability or vulnerability of different regions. Some ecologists have been critical of using carrying capacity or sustained yield to assess sustainability because of difficulties in establishing the appropriate threshold measures (for instance, Holling 1977).

Likewise, some economic measures of development, such as Heilbroner’s (1974), have been criticized as sustainability measures. For example, Pearson (1985) points out that most economic accounts focus on flows rather than stocks of resources and do not incorporate

| Table 1. Suggested measures of sustainability (refer to text for details). |
|-----------------------------|-----------------------------|-----------------------------|-----------------------------|
| Brown (1987)                | Per capita gross world product |
|                             | Per capita fossil fuel consumption |
|                             | Ecological deflator for soil erosion |
|                             | Carrying capacity |
| Clark (1986)                | Population density |
|                             | Agricultural production density |
|                             | Energy production density |
| Brown (1981)                | Renewable resource use |
|                             | Population growth |
|                             | Raw material recycling |
|                             | Soil erosion |
|                             | Rates of forest and fish harvest in relation to sustained yield |
| Heilbroner (1974)           | Oil prices |
|                             | Economic growth rates |
|                             | Environmental pollution |
| Brubaker (1972)             | Pollution |
|                             | Erosion |
|                             | Climate change |
|                             | Overpopulation |
| IUCN (1980)                 | Topsoil and farmland loss |
|                             | Malnourishment |
|                             | Deforestation |
|                             | Water supply disruption |
|                             | Fishery pollution and depletion |
nonmarket environmental services or impacts. Economic indicators, particularly those of national income like Gross National Product (GNP), have also been criticized for ignoring intranational distribution of income and hence overlooking possibilities for considerable inequity.

Chandler (1987) suggests that sustainable economics requires a net present value criterion, a conservation criterion and, perhaps, an equity criterion. Net present value is an efficiency measure where maximum output is obtained relative to input; the conservation criterion is met only if productivity is maintained for the next generation; and the equity criterion minimizes the ratio of the income share of the richest 20% of a population compared to that of the poorest 20%.

In the environment literature, there are general principles established for critically evaluating existing measures or for establishing new indicators. Many authors emphasize the need for comparable measurement units, time frames, and geographic scales in any system of global indicators (Clark, 1985). The National Academy of Sciences (1977), for example, recommends that indicators be cost effective, carefully collected, anticipatory, able to provide insight into causes and trends, accurate, able to relate physical environmental quality and condition of biological organisms, and well coordinated in multiple monitoring efforts.

In the development literature, the evaluation of indicators has also been an important topic. Seers (1972) suggests that, in order to devise meaningful indicators, measurement systems must begin with rigorous definitions of what is being measured. He goes on to demonstrate how many development indicators are value laden, biased, and poorly collected, particularly those related to income. Several criteria for quality-of-life measures are proposed by Morris (1979), including lack of cultural bias, sensitivity to the social distribution of conditions, and ease of collection. He provides an extensive critique, based on his criteria, of GNP as a development indicator.

Desirable Characteristics of Sustainability Measures

In our own work, we have established a set of guidelines to assess the value of sustainability indicators and, where appropriate, to aid in the development of more sensitive measures. We do not expect, however, that any one indicator will meet all of our criteria. Indeed, the objectives of these criteria sometimes conflict, and tradeoffs might be necessary to select or reject any particular indicator. Nevertheless, our set of guidelines is designed to provide an initial framework to evaluate and construct sustainability measures at global, regional, and local levels.

To measure sustainability, we think an indicator needs to have several fundamental characteristics. It should be sensitive to changes in time and to distributions over space and among social groups, predictive or anticipatory, referenced to some sort of range or threshold levels, free from bias, and able to tell something about the likelihood (and costs) of reversing or controlling changes.

In addition, since it is often the case that raw data alone cannot describe explicitly the conditions of sustainability, transformation and integration of the data should be considered where appropriate to produce more usable and meaningful indicators. Finally, from an institutional and policy-making perspective, an indicator should be relatively easy to collect and interpret.

Sensitive to Change in Time

Discussions of sustainability often focus on rates of change over time, toward or from conditions identified as sustainable. An indicator must, therefore, be collected at a frequent enough time interval to detect significant trends and variations, and ideally should be part of a historic time series that can illustrate long-term trends. The very long-term changes and cycles, spanning centuries or longer in the case of physical and chemical phenomena, are more difficult to measure, but can sometimes be reconstructed, as in the case of biological proxies for climatic change. Of course, change alone does not necessarily imply movement to or from sustainability, but can be part of the healthy functioning of a system (for example, seasonal cycles). A good indicator should be able to separate such normal cycles from trends away from a sustainable state.

Unfortunately, many current indicators are not available on a timely, up-to-date basis. For example, demographic data is often limited by the 10- or even 20-year time period between national censuses, whereas important changes can occur on a shorter time scale. For some types of measures, such as for national drinking-water quality, there may be only a single survey ever undertaken for a country or region.

Sensitive to Change Across Space or Within Groups

A major issue in social and development indicator research is the quest for measures which are sensitive to the distribution of conditions within a population or over a geographic region. Gross National Product (GNP) has often been criticized as a measure of develop-
opment because it may show an increase even when conditions are deteriorating for the majority of people in a country. It has sometimes been replaced by, or contrasted with, measures of income distribution, such as the ratio of the income of the wealthiest in a population divided by that of the poorest.

The distributional sensitivity issue is not only a matter of income. National average measures of land ownership, food availability, and energy use can hide vast differences in the access to land, food, or energy within a country. These differences are not only between rich and poor, but between men and women, and between different regions and ethnic populations of a country. If sustainability is defined as the provision of basic life supports, as well as quality of life, for everyone, then measures of sustainability must try to take into account the distribution of conditions. One way to focus the distributional issue is to measure conditions for those people or places at greatest risk, for example infants or the poor, or the ecological margins of forests or deserts.

Predictive or Anticipatory

Indicators which can predict or anticipate nonsustainable conditions have tremendous value for managers. Time series can be used in predictive extrapolation or simulation modeling, and global modelers have used a combination of empirical estimates and theoretical assumptions to warn of potentially "nonsustainable" futures (Meadows and others 1972).

An indicator can be anticipatory without being derived from a time series, though, if the indicator is known to preface serious threats to survival. Early warning systems have been proposed for anticipating famine based on information about the ability of a society to obtain food (USAID 1986). In such systems, increases in urban food prices, cattle slaughter rates, and population movements are seen as precursors of the food shortages and economic deprivation that create famine.

Reference or Threshold Values Available

Trends in natural and social systems are sometimes meaningless unless they can be evaluated within the context of reference values. Changes in life expectancy, for example, are easier to evaluate when the range of possible values is known. Increases in crop yields must be seen in terms of the potential yield defined by photosynthetic limits. Calorie availability per capita is only relevant if reported in terms of the caloric need for survival. Threshold values can be critical, for example, where a change in pollution levels or resource exploitation brings a species to a level where its extinction becomes likely. Any indicator purporting to measure sustainability should be reported with reference to any known threshold value of that indicator.

Unbiased

Many indicators of development and environmental quality are ethnocentric and based on the values of western civilization. Literacy, western housing standards (in regards to space, privacy, and appliances), and industrialization are valued differently by some societies. Drewnowski (1972) points out that it is impossible to establish a totally unbiased set of indicators. Value judgments enter into every stage of indicator construction—what, where, how, when to measure, how to weight, and how to present the results. Morris (1979) suggests that indicators such as infant mortality and life expectancy, contain little bias and could be considered universal measures of living conditions.

Reversible or Controllable

From a management perspective, it is critical to identify indicators which reveal whether changes are reversible and controllable. Perhaps the most critical changes to life support systems are those which involve a permanent and irreversible shift in conditions, at least on the time scale of human civilization. Such changes might include the total removal of topsoil, destruction of tropical forests, desertification, and the release of nondegradable toxic materials. It can also be useful to know the cost of controlling the degradation of life support systems.

Appropriate Data Transformation

Data transformations are often useful in designing sustainability indicators. Raw data, such as the area of forest in a country, do not indicate very directly whether the system is sustainable. Rates or throughput values, such as area deforested per year, are better, but still require a reference level, such as the remaining stock of a resource, to provide a direct measure. Hence, the ratio of forest loss per annum in relation to original or existing forest area may be a better indicator of sustainability, since it provides a measure of throughput in relation to remaining stocks.

Population-resource ratios can also provide more insight than pure demographic data and can be compared to estimates of land capability for food production to determine whether a threshold is being exceeded. The resource risk indicator, for example, developed by Berwick and reported in World Resources 1986 (WRI/IIED 1986) describes risks to coastal ecosystems by classifying each country on the basis of two
indices. One index measures the relative abundance of a particular type of coastal ecosystem as a function of area of ecosystem and length of coastline. The other index measures relative vulnerability and is derived from data on GNP, population density, and population growth rate. The lower the value of the resource abundance index and the higher the value of the vulnerability index, the greater is the risk to the resource. By ranking countries according to these indices, geographic regions needing critical attention can be identified.

**Integrative**

Composite indicators, such as the resource risk indicator just described, which integrate various measures into an index can be useful tools for measuring sustainability. Such indices are common in the environment and development literature, and have been used as measures of such conditions as the quality of life, soil erosion potential, and air and water quality. Great care needs to be taken, however, in the scaling and weighting of components in such indicators, especially in combining incommensurables. The composites may also be difficult to communicate and explain to the public and to policy makers.

As an alternative to composite indicators, techniques may be used to identify representative indicators—those which tend to mirror the behavior of a wider group of measures, or countries which can be used as representatives of larger regions. Such indicators could be chosen through expert judgment or through multivariate statistical techniques.

**Relative Ease of Collection and Use**

Despite progress in developing environmental monitoring systems and data collection networks, there exist problems with information availability, overlaps between monitoring systems or collection networks, and the appropriateness of the scale and nature of measured variables.

The OECD, for example, states that “environmental information is still characterized within countries ... and at the international level by major data gaps for some remaining environmental problems and most emerging ones” (1985:268). To add to this difficulty, *World Resources 1986* states that “there is no single theory or approach that can be used to identify the most important indicators of world resource use and environmental quality” (WRI/IIED 1986:xiiii). Within either the planning or research context, multiple conflicting objectives for establishing a monitoring system seem inevitable. For example, one important objective might be to minimize the cost of collecting, storing, and accessing data, without impairing the usability of the data and the ability to understand what is being measured.

Unfortunately, some of the most insightful indicators of social and environmental conditions are difficult and expensive to collect. Data on nutritional levels and ecological changes in remote areas are notably difficult to obtain and require careful sampling, the use of surrogate measures, or remote sensing. Often such data are only reported for small areas or case studies and can not be used for time series analysis or comparative work.

While the allocation of greater programmatic resources, including the emergence of new data gathering, storage, and dissemination technologies, will help alleviate some of the limitations, care is essential in selecting indicators so as to provide the greatest possible knowledge of global sustainability trends within the budgetary constraints of the data collection organizations. Some researchers indicate that, in many cases, existing monitoring systems are likely to be used, with modifications made, to meet the needs of new programs (Izreal and Munn 1986). Indeed, a significant constraint in designing such systems might be the necessity of following long-term trends in environmental conditions, making it unwise to routinely abandon existing historical time series.

Once acquired, the communication of indicators to policy makers can be greatly improved with the appropriate use of graphics (Cleveland and McGill 1985), computer demonstration programs, and simulation games (Clark 1986). Indicators are often better understood, particularly by policy makers, if they have some immediate social or political meaning. For example, pollution may be measured at the source (waste outfall or emission stack), in the environment (water or air quality), or in the receptor (concentration in biota or humans). The latter is most socially and politically immediate and relevant to a politician, but the former may be more meaningful to a scientist or engineer.

**How Well Do Current Indicators Measure Sustainability?**

In the following section, we examine two categories of indicators (soil erosion and population) and two specific indicators (physical quality of life index, and energy imports as a percentage of energy consumption) with respect to the guidelines we have just presented above. These examples are intended to demonstrate the process we would use to review whether an indicator tells us what we really want to know about sustainability.
Soil Erosion

Soil conservation is recognized as an important consideration in the long-term maintenance of the land resource base and measurements of the amount of soil erosion have been used to describe sustainability, particularly concerning the productivity of agricultural systems. Unfortunately, soil erosion data are difficult to collect, are limited in their usefulness, and are often not reported in the most appropriate form.

For example, the World Resources reports contain compilations of soil erosion estimates which are available for only a handful of countries (see Table 5.5 of WRI/IIED 1986). The reported rates were derived from numerous sources, which utilized different methods of estimation by the collecting governments or institutions, and are not comparable spatially or temporally between countries. Several estimates are based on single samples, and the lack of time series data limits predictive ability; some rates are reported for watersheds or regions, while other rates are reported for entire countries. And most of the field measurements are based on sheet and rill erosion estimates and do not account for wind erosion, which for some areas, is the more significant agent of erosion (WRI/IIED 1986:272).

In addition, as Blaikie (1985:16) points out, measurements of transported soil may not be that meaningful since they "do not identify the spatial or temporal source of sediment." He also cites scale as an important factor affecting soil erosion measurement. For example, if the measurement area is sufficiently large, there could occur considerable soil transport and deposition within the measurement area, but net soil erosion from the area could be estimated as being considerably small. Stocking (1987) discusses several studies that show many errors when estimates are aggregated from one scale to another. It is thus difficult to ascertain reliable estimates about changes over space and time or to make predictions based on so few and incompatible measurements.

Additionally, soil erosion rates alone reveal little about sustainability because they measure rates without reference to the stock that is being displaced. A rate-to-stock indicator, such as the loss of topsoil in relation to the amount remaining, could be reported to give a sense of this depletion. Reporting that a region has lost half its topsoil in a given time period is more indicative of unsustainable use than just reporting a loss of x tons per hectare per year.

An example of a threshold used to indicate the seriousness of soil erosion is the T value, used by the US Department of Agriculture. T values are defined as the maximum annual soil loss that can be sustained without adversely affecting productivity of the land. When erosion exceeds these values, it is in a sense, nonsustainable. T values, at first glance seem likely candidates for sustainability measures. However, as Batie (1983) and Stocking (1987) point out, they require much field work and research to estimate, and they do not incorporate the impact of the technology and economic conditions that are part of day-to-day farm management.

The T values are frequently related to measures obtained from the universal soil loss equation (USLE), an integrated index which combines several factors (rainfall/runoff, soil erodibility, slope length, slope steepness, crop cover and management, and farming practice) to estimate water-induced sheet and rill erosion (Batie 1983). The USLE was developed and calibrated using soil, rainfall, and land use conditions prevalent in the eastern and midwestern United States. Although it is widely used by the US Soil Conservation Service, there are a number of limitations in its widespread application, particularly in areas of the world with differing physiographic and climatic conditions or where the are rapid changes in agricultural practices (Stocking 1987).

The issue of loss of soil fertility is ignored by most measurements of soil erosion and, as Stocking (1987) points out, erosion rates are poor indicators of losses in soil fertility or productivity. Brown (1987) proposes an ecological deflator which decreases production estimates by the amount of production which is associated with unsustainable uses of land—specifically, production on highly erodible land (as well as production which uses unrechargeable irrigation). Based on this criterion, he estimates sustainable US grain production to be five-sixths of current levels.

The World Resources Institute stresses the importance of trying "to fill the gaping holes in current knowledge" about soil erosion. However, since existing information about global soil erosion rates is not only limited but, in general, fails to meet most of our criteria, as much attention should be given to developing more appropriate measures as is given to the actual collection efforts.

Population

Population data are routinely collected for almost every country and reported in most publications on global environment and development conditions. Some demographic indicators are more oriented toward measuring sustainability than others, especially those which measure the threat of population to the life support capacity of the earth through the use of appropriate data transformations and ratios.
The most widely reported and easily collected demographic measures are population level and population growth rate. These indicators, by themselves, provide very little insight into sustainability, although the transformation of population growth rate into doubling time (the number of years it will take for a population to double at a given growth rate) provides a clear measure of the potential increase in population pressure on resources.

Population data can be used as the basis for sophisticated demographic projections and can be used to anticipate future problems, particularly when projections consider the number of females under age fifteen and their future fertility.

Sustainability is measured more directly when population measures are combined with resource estimates. Population density (persons per unit of land area) is a common indicator and is used by Clark (1986) in discussions of sustainability. In terms of the life support systems, however, this transformation can be misleading when there are vast differences in the quality of resources, such as cultivable land, between regions. For example, Kenya and Mexico, with similar population densities for the total land area (Table 2), have very different densities of population in terms of cropped land.

Differences in the quality of resources may be obscured by differences in technology and income. Although the FRG has the same population density per cropland area as Kenya, West Germany’s population lives at a higher standard (measured by calories/capita) than Kenya’s because of its more intensive agriculture and higher income. Although Mexico and the FRG both seem to have adequate diets (126% of calories required), Mexico has more problems sustaining population because calories are not equally distributed within the national population. In this case, the better indicator may be one which takes into account the distribution of social conditions.

In terms of income distribution, the wealthiest 10% of the population of the FRG, Mexico, and Kenya earn 24%, 41%, and 46% of their national incomes, respectively. The poorest 20% in each country earn 7.9%, 2.9%, and 2.6%, respectively, of the national income (World Bank 1986).

Table 2. Population indicators 1985 (WRI, 1986).

<table>
<thead>
<tr>
<th>Indicators</th>
<th>Mexico</th>
<th>Kenya</th>
<th>FRG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>78 m</td>
<td>21 m</td>
<td>61 m</td>
</tr>
<tr>
<td>Population growth rate</td>
<td>2.6%</td>
<td>4.1%</td>
<td>0.0%</td>
</tr>
<tr>
<td>Doubling time</td>
<td>27 years</td>
<td>17 years</td>
<td>Infinity</td>
</tr>
<tr>
<td>Population/land (density)</td>
<td>41/km²</td>
<td>35/km²</td>
<td>245/km²</td>
</tr>
<tr>
<td>Population/cropland</td>
<td>340/km²</td>
<td>875/km²</td>
<td>816/km²</td>
</tr>
<tr>
<td>Calories/capita</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>as % need</td>
<td>88%</td>
<td>126%</td>
<td>126%</td>
</tr>
</tbody>
</table>

capacity of land for food production is, and will be, exceeded. There are, however, many problems with such population carrying capacity studies. For example, many of them fail to assess the social, economic, and political constraints to increased food production (Hekstra and Liverman 1986).

This discussion illustrates just a few of the ways in which population indicators can be used to assess sustainability and some of the disadvantages of certain measures.

Physical Quality of Life Index

The physical quality of life index (PQLI) has been proposed as a sensitive and appropriate measure of development (Morris 1979). It is used by the Overseas Development Council and several other agencies to assess general trends in basic social conditions. Some books (for example, Myers 1986) use PQLI in discussions of sustainable development. The PQLI is a composite indicator which combines literacy, infant mortality, and life expectancy. It was designed to meet several criteria including ease of collection, freedom from bias, sensitivity to distribution, and integrative potential.

Infant mortality, life expectancy, and literacy are transformed by scaling and weighting in the PQLI. Each variable is scaled from 0 to 100 using reference values of 0 for the lowest measured level and 100 for the highest in the year the index is calculated. In 1970, for example, life expectancy ranges from 41 years for Mali to 75 years for Iceland, denoted as 0 and 100, respectively, on the scale. Although this scale does provide comparable reference values, it creates considerable difficulty when constructing time series. In 1980, the reference values change to 46 for Nigeria (0) and 78.6 for Sweden (100). Once the three variables are scaled, they are added together and averaged to provide a PQLI on a 0 to 100 scale which in 1970 ranges from 12 for Upper Volta to 97.5 for Iceland. The equal weighting of the three variables makes the data transformation simple, but does not necessarily reflect the relative importance of each.
Changes in infant mortality and life expectancy are generally linked to the condition of life support systems and, hence, do indicate changes in sustainability. Literacy, however, is a less relevant measure, since cultures value it differently and it is not clear how literacy relates to sustainable development. The inclusion of literacy in PQLI also creates general problems with the index. For example, Brodsky and Rodrick (1981) demonstrate that literacy data are often very outdated and difficult to obtain, which means that PQLI is not very sensitive to change and that time series of the index are extremely difficult to construct for many countries.

Although Morris (1979) claims that PQLI is a relatively unbiased, unethnocentric measure because infant survival, long lives, and education are valued in all societies, it is not at all clear that the value of literacy, in particular, is universal. When PQLI is compared with other indicators, like GNP and calorie availability, the inclusion and equal weighting of literacy in the PQLI results in higher rankings for countries in Asia and lower rankings for countries in the Middle East region.

The value of the PQLI in prediction is limited by the lack of good time series data. It also seems that PQLI is likely to lag behind rather than anticipate social and environmental changes in life expectancy, for example, will decline only after several years of poor life support conditions, such as famine or contaminated water supply.

Morris also claims that PQLI is sensitive to distribution because improvements in infant mortality and life expectancy tend to reflect better conditions for the poor. He claims, therefore, that PQLI is a much more sensitive measure than GNP per capita, improvements in which often reflect increases in the incomes of a wealthy few. However, our research has shown wide variations in PQLI by gender, region, and income class within a country and that improvements in one group or area are often not seen in the aggregate national PQLI.

For example, in 1980, 68 countries had differences of more than 6 points between male and female PQLIs (for instance, in Zaire: women 32 and men 70) strongly influenced by literacy differences. Morris himself reports that, in Ghana, urban PQLIs are twice as high as rural PQLIs, and, in Mexico, we found that improvements in national PQLI over time are influenced by dramatic improvements in just a few areas, while some regions have declining quality of life.

We also found significant correlations between PQLI and GNP per capita (0.65 in 1970, 0.6 in 1980) and a relatively consistent relationship between them over time which casts doubts on the relative unbiasedness and distributional sensitivity of PQLI. Finally, such a composite indicator as PQLI, despite its integrative potential, may be difficult to interpret. The scaling of PQLI, and the inclusion of out-of-date, gender- and culturally-biased literacy, makes the index a confusing measure of sustainability. Perhaps a composite of life expectancy and infant mortality with a standard scale over time, would be a more appropriate measure.

In summary, this preliminary review of the PQLI suggests that the index fails to meet several of our desired characteristics for sustainability indicators. Although it seeks to integrate, to be sensitive to distributions, and to be unbiased, it has a number of problems in meeting these objectives. The data required to construct the index are sometimes difficult to collect or are outdated, and it is difficult to use in the index in assessing change over time. The weighting and scaling of the components within the index are somewhat arbitrary, the inclusion of literacy may be culturally biased, and, and we have shown, the index is not particularly sensitive to distributions.

Energy Imports as a Percentage of Energy Consumption

An adequate supply of energy is essential to sustaining all natural and human systems. It is the lifeblood of any economy. As an indicator, energy imports as a percentage of energy consumption, measured in petajoules of commercial energy, addresses the issue of national self-sufficiency in commercial energy and is one of ten energy indicators reported in the World Resources reports.

The indicator is integrative in the sense that it relates supply-side and demand-side information (Table 5). In this regard, it is more useful than annual production levels of solid, liquid, or gaseous fuels, which are three other widely reported indicators. These latter values, while certainly of some interest are difficult to interpret in that, by themselves, they provide no indication of how energy production is related to need, or whether the level of annual energy production is significant in relation to the resource available. While energy imports as a percentage of consumption appears to be a useful data transformation, it nevertheless has some significant drawbacks.

It is sensitive to change in time, as annual time intervals are appropriate for most national energy management purposes. Between 1970 and 1983, the value for the USA changed from 7% to 13%, indicating that the USA was gradually becoming an energy importer. Given the massive capital structure involved in com-
**Table 3. Comparison of three energy indicators.**

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Commercial energy consumption (petajoules)</th>
<th>Imports as a percentage of consumption (%)</th>
<th>Imports as a percentage of mercantile exports (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitive to change in time</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Sensitive to change across space</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Predictive ability</td>
<td>Yes, although per capita comparisons of some value</td>
<td>Yes</td>
<td>Yes, imports limited to export earnings</td>
</tr>
<tr>
<td>Reference values available</td>
<td>No</td>
<td>No</td>
<td>Yes, via demand-side actions or for countries with large resources</td>
</tr>
<tr>
<td>Unbiased</td>
<td>Biased toward commercial fuels</td>
<td>Biased toward commercial fuels, but relevant to commercial energy situation</td>
<td>Biased toward commercial fuels, but relevant to export requirement</td>
</tr>
<tr>
<td>Reversible or controllable</td>
<td>Yes, to a considerable degree</td>
<td>Yes, for countries with large resources relative to use</td>
<td>Yes, via demand-side actions or for countries with large resources</td>
</tr>
<tr>
<td>Appropriate data transformation</td>
<td>No</td>
<td>Somewhat Measures throughput</td>
<td>Somewhat, relates energy demand and supply flows</td>
</tr>
<tr>
<td>Integrative</td>
<td>No</td>
<td>Somewhat, relates energy demand and supply flows</td>
<td>Yes, relates commercial energy flows to export ability</td>
</tr>
<tr>
<td>Ease of collection and use</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Commercial energy resource exploitation, this measure will have considerable stability and, therefore, may have predictive ability.

The indicator seems especially weak in that it has no meaningful reference values or thresholds. The range of 1983 values for this indicator (as reported in *World Resources 1987*) is from −3518% for Oman, a large exporter of oil relative to its population and economy, to +100% or more for 39 of the 146 countries listed (IIED/WRI 1987). (Imports can be more than 100% of consumption due to electricity accounting procedures, as well as data errors or inconsistencies.) While these values convey some meaning—Oman is better off than the USA, which is better off than Bangladesh—there are no known thresholds established. Indeed, some countries which appear to be in a disadvantaged position, such as Japan with a 94% value in 1983, are easily capable of paying for their energy imports. This picture if further complicated by the fact that renewable energy sources, such as firewood, on which roughly half the world’s population is heavily reliant, are not included in this measure.

This indicator does not necessarily reveal whether changes (or levels) are controllable. In countries where there are significant commercial energy resources, there appears to be some choice as to the level of imports relative to consumption. This choice requires significant capital and time. For those that do not have the resource, there is little choice, regardless of efficiency or use of renewables.

This indicator is not sensitive to change across space or within the population of a country, as it represents a national average. Because of the large economic importance of the flows involved, these data are routinely and easily collected. Also, the indicator is significantly biased toward activities that use commercial sources of energy. Finally, while the indicator is somewhat integrative, as stated above, it does not provide a means for relating energy flows to the ability of the economy to afford the flows. It also ignores energy use and activities that are dependent on renewable sources of energy. Thus, while energy imports as a percentage of energy consumption has some advantages relative to a simpler measure, such as oil production or commercial energy consumption per capita, it is rather limited in usefulness. This comparison is summarized in Table 3.

In response to this evaluation, we used the eight characteristics, or criteria, in an attempt to develop an improved indicator of energy sustainability—one that would respond to at least some of the criticisms raised above. One indicator that we have evaluated is energy imports as a percentage of mercantile exports, which we used the *World Resources 1986* data tables and World Bank information (WRI/IIED, 1986; World Bank, 1986, respectively). The evaluation of this indicator is also summarized in Table 3.

Its greatest advantage relative to energy imports as a percentage of energy consumption is that it directly relates the flow of commercial energy imports to the ability to pay for those imports, which must in the long run come from exports. Thus, it is clearly a superior...
data transformation with regard to integrating information on energy flows with the economic capacity to pay for those flows. The indicator also has some meaningful threshold values. Energy imports cannot exceed mercantile exports on a long-term basis and, in light of other import needs, energy imports must, in the long run, remain at some fraction of exports. Exceeding a certain fraction of exports on a long-term basis to pay for energy imports likely would be a useful predictor of impending economic difficulties.

For other criteria, this improved indicator offers little or no advantage. It is insensitive to change across space or among population groups and it is biased toward economic activities that are dependent on commercial energy sources. On balance, however, it offers significant advantages over energy imports as a percentage of energy consumption.

Conclusion

Global sustainability has become an important, yet ill-defined, concept in environmental policy and research. In order to study currently used indicators, as well as prospective indicators, we have proposed a set of desired characteristics for indicators of sustainability. While it is unlikely that any indicator can meet all of these criteria, many currently used indicators fail to meet many of these criteria. There is a need to evaluate carefully the existing sets of indicators which are incorporated into current and planned global monitoring systems and studies of ecologically sustainable development. In many cases, the failures of the indicators relate to lack of historic time series or to inadequacies of existing data and will be difficult to remedy. New monitoring technology, such as remote sensing satellites and computerized geographic information systems, and new programs such as those described in the introduction, may be able to address some, but certainly not all, of these problems. In view of the considerable resources being committed to collecting data on the currently used indicators, there is a need for a critical analysis to evaluate, modify, and, where necessary, replace these measures. We hope that the criteria proposed in this article will provide a basis for further research and analysis on this urgent task.

Acknowledgments

We acknowledge graduate students, Altha Cravey, Bruce Herrick, Dan York, and Kevin Carroll, for their invaluable assistance in preparation of this report. We thank Dr. Leong Yueh-Kwong of the Universiti Sains Malaysia for his helpful review and comments on earlier drafts of this manuscript. Also, we recognize our colleagues at the University of Wisconsin—Madison who contributed their support and advice. This project was supported by funds granted by the William and Flora Hewlett Foundation to the University of Wisconsin—Madison Institute for Environmental Studies.

Literature Cited


Pearson, C. 1985. Down to business: multinationals, the environment, and development. World Resources Institute, Washington, DC.


