WP-2010-013

Selection Criteria for Sustainable Development Indicators

Hippu Salk Kristle Nathan and B. Sudhakara Reddy

Indira Gandhi Institute of Development Research, Mumbai
August 2010
http://www.igidr.ac.in/pdf/publication/WP-2010-013.pdf

Selection Criteria for Sustainable Development Indicators

Hippu Salk Kristle Nathan and B. Sudhakara Reddy
Indira Gandhi Institute of Development Research (IGIDR)
General Arun Kumar Vaidya Marg
Goregaon (E), Mumbai- 400065, INDIA
Email (corresponding author): hnathan@igidr.ac.in

Abstract

The selection of optimum number of indicators is the key to any Sustainable Development Indicator (SDI) research. Indicators, too less in number may be inadequate to convey the message and; too many may dilute the purpose. To arrive at a limited number of indicators from a larger set of potential ones, a set of criteria is applied. In the past SDI-initiatives, emphasis was not laid on structuring criteria; these were invariably considered in a linear way. This paper proposes a framework to organize criteria in a tree fashion at successive levels. For each level, the weights of a criterion relative to others are determined by following Analytical Hierarchy Process where experts' opinions are considered. For aggregation of scores, 'displaced ideal' method is proposed over linear additive model.

Keywords:

What-How-Whom (WHW) framework, Value tree, Analytical Hierarchy Process, Displaced ideal

JEL Code:

Q01, Q56, D70

Acknowledgements:

Selection Criteria for Sustainable Development Indicators

Hippu Salk Kristle Nathan* and B. Sudhakara Reddy

Indira Gandhi Institute of Development Research (IGIDR)
General Arun Kumar Vaidya Marg
Goregaon (E), Mumbai-400065, INDIA
*Corresponding Author Email: hnathan@igidr.ac.in
Phone: +91-9867582994, Fax - +91-22-28416399

Abstract

The selection of optimum number of indicators is the key to any Sustainable Development Indicator (SDI) research. Indicators, too less in number may be inadequate to convey the message and; too many may dilute the purpose. To arrive at a limited number of indicators from a larger set of potential ones, a set of criteria is applied. In the past SDI-initiatives, emphasis was not laid on structuring criteria; these were invariably considered in a linear way. This paper proposes a framework to organize criteria in a tree fashion at successive levels. For each level, the weights of a criterion relative to others are determined by following Analytical Hierarchy Process where experts' opinions are considered. For aggregation of scores, 'displaced ideal' method is proposed over linear additive model.

Keywords: What-How-Whom (WHW) framework, Value tree, Analytical Hierarchy Process, Displaced ideal

JEL Codes: Q01, Q56, D70

1 Introduction

The notion of sustainability is gaining momentum and construction of sustainability development indicators (SDI) has become a popular trend. Efforts have been made to assess sustainability at different levels and domains and to keep track of the progress towards achieving sustainable development (SD). SDI initiatives have been undertaken by international, national authorities as also by regional and local bodies in industrialized countries as well as in developing ones encompassing both public and private sectors (Parris and Kates, 2003; Boulanger, 2007).

One of the trickiest concerns in indicator research is to arrive at a final list of 'limited' number of SDIs. If the number of indicators is too large, being unwieldy, it defeats the purpose all together; as by definition, indicators reduce the number of measurements and parameters to provide the exact representation of a given situation.[1] Recent literature has argued to optimize the number of indicators which then becomes usable and give a synoptic and representative view of the actual situation (Boulanger, 2007; Laloë, 2007; Rey-Valette et al., 2007a). Any indicator to find a place in the final list has to meet several criteria such as relevance to objective, simplicity in understanding, analytical soundness, policy responsiveness, flexibility, etc. The set of criteria may change with purpose, but in general, criteria are multiple, multi-dimensional and multi-leveled.[2] Moreover, all criteria may not be of equal importance. In such a scenario, choice and management of criteria needs both care and logic.

In the light of the above, this study bears importance as it looks into the selection aspect of the SDIs. The relevance of the study can be further linked to the growing concern in the literature about the fact that SDIs have emerged simply as measurement indicators and their real use in terms of influencing policies is still at a stage of infancy (Bell and Morse, 2003; Pinter et al., 2005). In this paper, we propose a set of criteria and develop a framework to organize them. The exhaustiveness of the proposed framework is tested by revisiting the past initiatives. We explore Multi Criteria Analysis (MCA), which handles several options contributing differently to different criteria, to check its applicability in SDI research. We use Analytical Hierarchy Process to determine importance of each criterion relative to the other. Also, we propose an aggregation technique to add scores of indicators obtained under different criteria.

2 Selection Framework

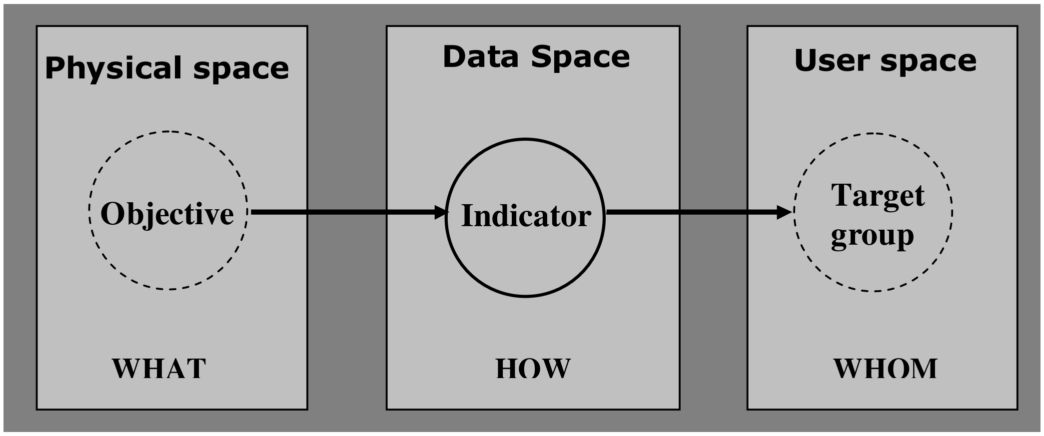
A review on the past initiatives shows despite criteria being applied for indicator selection, the literature is silent on the application methodology. Among all the past initiates, only Australia (1998), South Africa (DEAT, 2001), and OECD (2003) have grouped the criteria in some sense.[3] However, the groupings appear confusing in their nomenclature as there are conceptual overlaps. Moreover, South Africa (DEAT, 2001) and OECD (2003) initiatives are focused only on environment, whereas the present study is on three dimensions— economy, society and ecology (Nathan and Reddy, 2008).[4] The common feature in all the past initiatives is that the criteria are considered linearly without any structure and all criteria are given equal weight relative to each other.

We differ from this trend of linear handling of criteria with equal weight to each criterion. Linearity assumes perfect substitutability which means that a differential improvement (or increment) in one criterion at any value can be substituted or neutralized by an equal differential decline (or decrement) in another indicator at any other value (Nathan et al., 2008). This assumption does not fit the present context. For instance, let us assume a two-criterion scenario of *data quality* and *data availability* with equal weights to both. For a linear additive model, the overall scores of an indicator with score of 50% in both criteria equals to that of another indicator which scores nil in *data quality* and full in *data availability*. However, as common sense suggest the former indicator has a better prospect than the later. Also, simplicity might have been the only motive behind giving each criterion equal weight relative to the other. So, lack of organization for criteria in one hand and the simplistic assumption of giving all criteria equal weights on the other, are the prime motivation behind the development of a selection framework for SDI research.

2.1 WHW framework

We conceptualize a What–How–Whom (WHW) framework to organize the criteria. This three-dimensional framework follows from the basic meaning of the indicator i.e. an indicator provides means of communicating information from physical space to users via a communicable data. The information, which is meant to assess the health of any system towards any specific objective, is transferred through indicators in a simplified, yet effective manner to the target group for knowledge and response. In WHW framework, *what* component addresses the physical or science space, *how* is about the data whereas *whom* component deals with the users of indicators (see Figure 1).

Figure 1 What–How–Whom (WHW) framework for Indicators



The set of questions the WHW criteria framework will generate answers are the following.

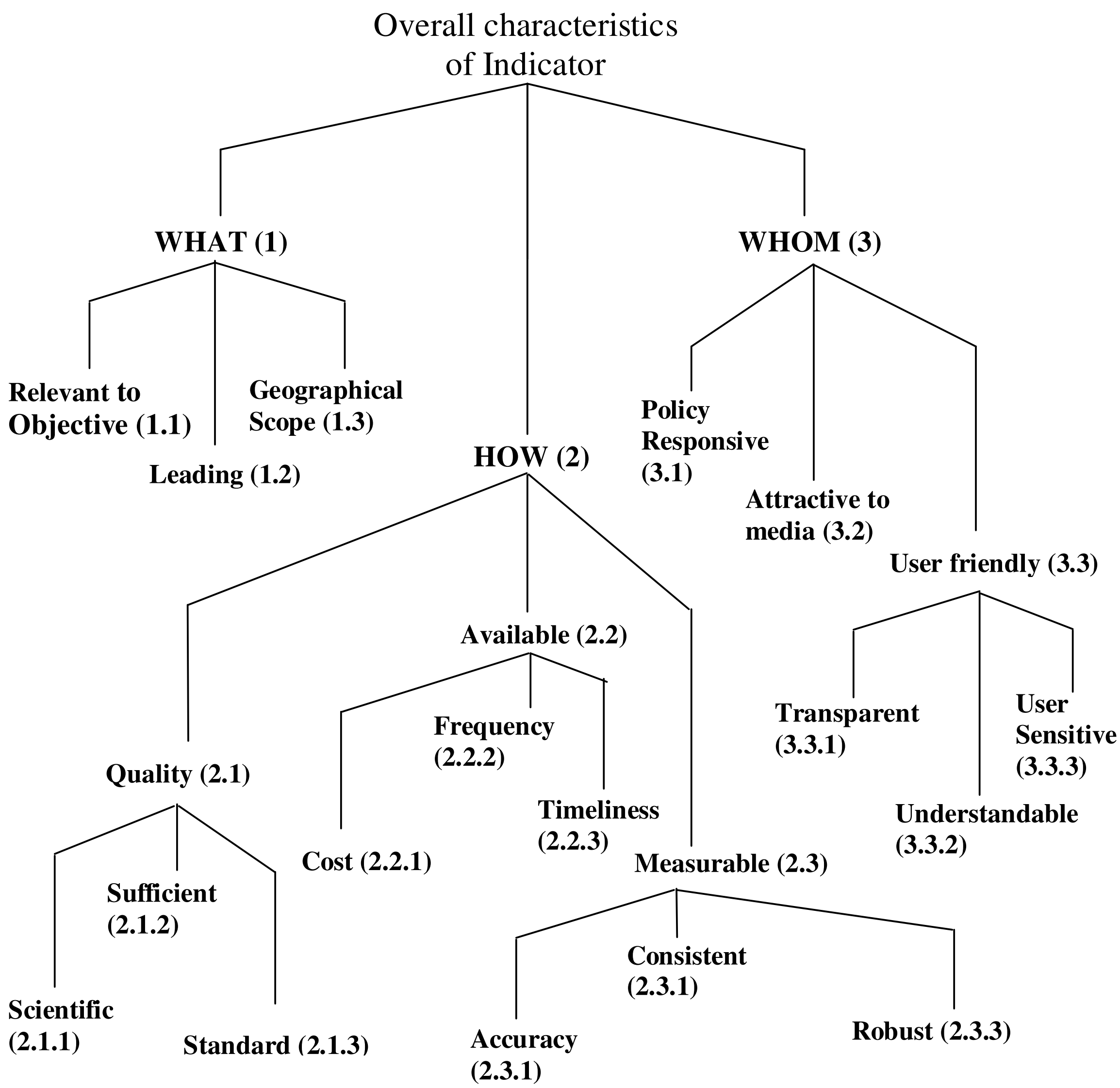
What— What does this indicator communicate? Is it relevant to the objective it claims to represent? What is its scope? Is it leading, i.e., does it possess inherent characteristics of futuristic role?

How— How does this indicator communicate? Is the data readily available or can be made available at a reasonable cost? Can the data be compiled regularly and without long delays? What is the quality of data? Is the data sufficient— not too much information, but just adequate to provide suitable picture of the situation? Is the data logically and scientifically defensible? Is it reliable and of international standard? Is the data accurate and consistent? Is it robust enough?

Whom— To whom does this indicator communicate? Is it being best used by the target group? Is it transparent and accessible to the citizens? Is it user-friendly— clear, simple to understand, and motivating? Is it responsive to policy interventions? Is the indicator attractive to media?

While answering the above questions different criteria of indicators have been identified and they are organized in a hierarchical fashion through a 'value tree' as shown in Figure 2. The *what*, *how* and *whom* dimensions are broken down into different levels to individual criterion. The exhaustiveness and robustness of the value tree is tested by examining the criteria mentioned in the past initiatives.[5] Table 1 shows the correspondence of criteria mentioned in the past eight initiatives to the proposed value tree scheme of things. The criteria are detailed in Appendix 1 with examples drawn mostly from energy domain.

Figure 2 Value tree of criteria-set developed over What–How–Whom (WHW) framework



Note: The numbers in the bracket show the numerical code for individual criterion.

Once the criteria are decided, the next steps are to determine the weights of each criterion relative to each other, score the indicators for each criterion, aggregate the scores, and then, rank or shortlist indicators based on the scores. For this, multi-criteria analysis techniques are explored next.

Table 1 Mapping of Criteria mentioned in Past Initiatives with WHW framework

Initiative	*Objective/Purpose*	*Criteria and its correspondence with WHW framework*
1. Sustainable Seattle, (2004)	To measure the sustainability of Seattle community	No of Criteria: Eight i. Relevant – Relevant to objective (1.1) ii. Reflect community value (3.3.3) iii. Attractive to local media (3.2) iv. Statistically measurable (2.2.1) v. Logically and scientifically defensible (2.1.1) vi. Reliable (2.3.1, 2.3.2) vii. Leading (1.2) viii. Policy Relevant (3.1)
2. United Nations Commission for Sustainable Development (UNCSD, 1996)	To understand the dimensions of sustainability and their interactions. To train and build capacity for countries to develop own set of SDIs. To monitor execution of Agenda 21 and its further development.	No of Criteria: Nine i. National in scope (1.3) ii. Relevant – to assess sustainable development (1.1) iii. Understandable, clear, unambiguous (3.3.1, 3.3.2) iv. Within the capabilities of national government (1.3) v. Conceptually sound (2.1.1) vi. Limited in number, open ended, adaptable to future needs* vii. Broad in coverage of Agenda 21 (3.1) viii. Representative of international consensus (2.1.3) ix. Cost effective data with known quality (2.1, 2.2.1)
3. Environmental Indicators for human settlements in Australia (1998)	To measure the impact of urban system on the environment, and to measure their success in providing an adequate environment for their inhabitants.	No of Criteria: 15 <u>Important</u> i. Reflect values aspect of the environment (1.1) ii. national in scope (1.2) iii. facilitate community involvement (3.3) <u>Feasible</u> iv. Be monitored regularly with relative ease (2.2) v. Be cost-effective (2.2.1) vi. Comply with international agreements (2.1.3, 3.1) vii. Consistent and comparable with other countries (2.1.3) <u>Credible</u> viii. statistically verifiable and reproducible with time trends and apply to range of environmental regions (2.3, 2.2.2, 1.3) ix. Be scientifically credible (2.1.1) x. use existing commercial and managerial indicators (2.2) <u>Understandable</u> xi. Have relevance to policy and management (3.1) xii. Be easy to understand (3.3.2) <u>Useful</u> xiii. Robust indicator of environmental change (2.3.3) xiv. Early warning of potential problems needs (1.2) xv. Monitoring of progress of environmental policies (3.1)
4. EU Local Sustainability Indicator (Ambiente Italia, 2003)	To evaluate the EU sustainable development strategy and implementation of policy measures	No of Criteria: 10 i. Capture the essence of the problem (1.1) ii. Clear and accepted normative interpretation (2.1.1) iii. Robust and statistically validated (2.3.3) iv. Responsive to policy intervention (3.1) v. Measurable across, and internationally comparable (2.1.3) vi. Timely and susceptible to revision (2.2.3 and 2.2.1) vii. Not a burden disproportionate to its benefits (2.2.1) viii. Balanced across different dimensions (1.1)

Initiative	*Objective/Purpose*	*Criteria and its correspondence with WHW framework*
		ix. Mutually consistent within a theme (1.1) x. Transparent and accessible to the citizens (3.3.1)
5. Winnipeg (1997)	To measure the progress towards 'Plan Winnipeg' vision, and understand the impacts of decision and actions over time	No of Criteria: Six i. Policy relevance (3.1) ii. Simplicity (3.3.2) iii. Validity (2.1.1 and 2.3) iv. Data Availability (2.2) v. Representativeness (1.1) vi. Sensitivity – to change in science and to user (1.1, 3.3.3)
6. OECD (2003)	To monitor sustainability in the OECD member countries.	No of Criteria: 11 <u>Policy relevance and utility for users</u> i. Simple, easy to interpret and show time trends (3.3.2, 2.2.2) ii. Responsive to changes (1.1) iii. Provide a basis for international comparisons (2.1.3) iv. National in scope (1.3) v. Have a threshold or reference value (2.1.3) <u>Analytical Soundness</u> vi. Theoretically well founded and scientific basis (2.1.1) vii. Based on international standards (2.1.3) viii. Linked to economic/forecasting models (2.1.1, 3.1) <u>Measurability</u> ix. cost/benefit ratio (2.2.1) x. Adequately documented and of known quality (2.1) xi. Updated at regularly with reliable procedures (2.2.2, 2.3)
7. Ireland (2003)	To measure the national progress of Ireland	No of Criteria: Seven i. Easy to read and understand (3.3.2) ii. Policy relevant (3.1) iii. Mutually consistent* iv. Timely availability (2.2.3) v. Comparable across member states and (2.1.3) vi. Selected from the reliable sources (2.1.1) vii. Not impose too large a burden (2.2.1)
8. South Africa (DEAT, 2001)	To do environmental reporting	No of Criteria: 13 <u>Scientific</u> i. Clear in value - direction is clear (1.1, 2.1.1) ii. Clear in content – easily understandable (3.3.2) iii. Appropriate in scale – optimally aggregated (2.1.2) iv. Hierarchical – user can delve down into the details (2.1.2) <u>Functional</u> v. Policy relevant- for all stakeholder (3.1) vi. Compelling and suggestive of effective action (3.3.3) vii. Sufficient (2.1.2) viii. Leading (1.2) <u>Pragmatic</u> ix. Feasible: measurable at reasonable cost (2.2.1) x. Tentative: up for discussion, learning and change^ xi. Timely: compliable without long delays (2.2.3) xii. Democratic: peoples' participation and access (3.3.1) xiii. Participatory : People can measure themselves (3.3.1)

*this criterion is a set property, not an individual property
^Could not be mapped.

3 Multi-Criteria Analysis (MCA) and its applicability in SDI research

MCA techniques are used when multiple options are to be evaluated against multiple criteria. The key feature of MCA is its emphasis on the judgment of the decision-making team in establishing criteria, estimating their relative importance (weights) and judging the score of each option (indicator) for every criterion. The scores and weights are entered in a performance matrix, X_{nm} of n options and m criteria, where x_{ij} is the performance score assigned to option i against criterion j. The relative importance of criteria is measured with a weight vector W where w_j represents the importance of the j^{th} criterion. MCA offers a number of ways of aggregating the score on individual criterion to get the overall performance of each option.

There is an increasing evidence of the use of MCA in indicator research (Rey-Valette *et al,* 2007; Roussel *et al*, 2007). The advantages of MCA can be outlined as following. First, it gets rid of the difficulties that human decision-makers face in handling large amounts of complex information in a consistent way. Second, it is applicable to both quantitative and qualitative criteria together where performance cannot be deduced to common monetary terms. Thirdly, it is flexible as criteria, scores and weights, once given, can be amended if necessary. Last, but not the least, the unique feature of MCA is the interactive nature of the technique, which provides means of communication within the decision-making body and sometimes later between that body and the wider community.

Among different MCA techniques, the simplest is elimination of dominated options by direct inspection of performance matrix.[6] This method has limited applicability in SDI research as here one is bothered about short-listing a set of indicators rather than just choosing one indicator among many. However, this method finds relevance when indicators are chosen for constructing a composite index, where usually one or two qualify to represent a dimension or component of the index.[7] It is worth noting here that when dominance occurs, before elimination, it is helpful to examine if there is some advantage of the dominated indicators that are not represented by the criteria; this may reveal new criteria that have been overlooked. Outranking is an advanced method of dominated elimination where weights are assigned to exert greater influence on some criteria than others.

MCA technique based on conjunctive (disjunctive) model eliminates (allows) options those miss (meet) the set levels of performance on one or more criteria. Both conjunctive and a disjunctive filters can be used in SDI research to quickly filter out indicators when thresholds are posited for one or more criteria. For instance, from budgeting point of view, indicators which exceed certain level of cost can be eliminated.

Lexicographic ordering is an MCA technique, where criteria are placed in the order of importance and all options are first compared in terms of the criterion deemed most important. If there is a unique best-performing option in terms of this criterion, then that option is selected as the most preferred. If there is a tie, then the selection process moves on to the second-ranked criterion and the process continues until a unique option is identified or all the criteria have been considered. For SDI research, lexicographic ordering can be used in combination with conjunctive/disjunctive models (such combined technique is known as 'elimination by aspects' in MCA literature) to arrange the criteria in the order of importance and set the threshold levels for each. For instance a researcher may choose to put criteria like *data availability*, *data quality* and *data measurability* in a strict order of preference and introduce thresholds to be crossed at each level for further processing to happen.

Analytical Hierarchy Process (AHP), which was originally devised by Satty (1980), is an MCA technique which uses pair-wise comparison for deriving weights and scores. In assessing weights, the decision-makers are asked questions on how important one particular criterion is relative to another. For scores, the importance of one option over the other for a given criterion is asked. In SDI research, AHP is useful because pair-wise comparison is straightforward and convenient where judgment-based decisions are intended. However, AHP method allows intransitivity of decision rules, suffers from limitation of rank reversal,[8] and is questioned for lack of theoretical foundation (Belton and Gear, 1983; French, 1988; Goodwin and Wright, 1998). Nevertheless, it is widely used in a variety of applications involving multi stakeholder analysis (Zahedi, 1986; Golden *et al*, 1989; Shim, 1989).

The MCA method, which comes closest to universal acceptance, is based on multi attribute utility theory derived from the work of von Neumann and Morgenstern (1947) and Savage (1954) and developed for use as an MCA technique by Keeney and Raiffa (1976). This is an aggregation technique where the overall performance is expressed as a single number index, *U,* in terms of a mathematical function, which allows mutual interaction among criteria and accounts for uncertainty. In spite of its distinct advantages, the method has limited use in SDI research because of its mathematical complexity, non-participatory character, and non-applicability to problem types where performance cannot be expressed in a mathematical term. A special case of multi-attribute theory, which has more applicability in public-sector decisions, is linear additive model, in which criteria are mutually preference-independent. The overall score for each option is obtained by multiplying the score of each criterion by the weight of that criterion, and then adding all those weighted scores together.

$$S_i = \sum_{j=1}^{m} w_j x_{ij} = w_1 s_{i1} + w_2 s_{i2} + \ldots + w_m s_{im} \qquad (1)$$

where, S_i is the overall score of option i and x_{ij} the score of option i for criterion j of which w_j is the weight. The inputs in the model are both weights and scores, which may be judged differently by different people. A simultaneous variation of weights and scores makes liner model complicated. In SDI research, the model can be employed only to find scores after the weights are determined; and it is so applicable wherever the criteria can be adjudged to be mutually preference independent.

There are other methods under MCA techniques, which are based on fuzzy sets, rough sets, or methods heavily dependent on interactive development using specially constructed computer packages. These methods will not be useful for SDI research as they tend to be complex and difficult for non-specialists to understand. Also, for a variety of reasons, none of these is likely to find widespread application to mainstream public sector decision-making.

3.1 Determination of weights for criteria

A review of different MCA techniques shows that, AHP is the most suitable methodology for deciding the weights of criteria relative to each other.[9] AHP technique is based on pair-wise comparison and such comparisons will be used for criteria under the same level to get their relative weights. Supposing there are n criteria, which are arranged in an $n \times n$ matrix where each element shows the relative importance of a row criterion over column criterion. Since criteria are considered in the same order in rows and column, the diagonal of the matrix turns out to be unity, i.e., the relative importance of any criterion over itself. To scale the scores of relative importance, a nine-point intensity scale has been proposed in the literature, with 1, 3, 5, 7, 9 indicating equally, moderately, strongly, very strongly and extremely important; and the intermediate scores 2, 4, 6 and 8 are used for expressing intermediate importance values (NERA, 2000).

For the current exercise, the scale of relative importance scores is reduced to 1 to 5. Table 2 gives the verbal definition of each score. Multiplicative inverse is used for scoring inverse relationships. For instance, if criterion A is three times more important than B, then automatically criterion B becomes one-third important than criterion A.

Table 2 The Fundamental Scale for Pair wise comparison

Definition	*Score*
Equal importance	1
Moderate importance	2
Strong importance	3
Very strongly more important	4
Overwhelmingly more important	5

Table 3 gives an example of pair-wise comparisons among the three sub-criteria under *data availability* criterion (criterion code 2.2. in Fig 2). For instance, *cost* is considered two times more important than *frequency* is and half as important as *timeliness*; and *frequency* is one-third important compared to timeliness, the matrix will look like Table 3. It needs to be noted, AHP pair-wise scores are symmetric, but need not be transitive. Since the inverse scores are automatically determined; in the n x n matrix only $\frac{n(n-1)}{2}$ entries need to be filled. For the above example, only three entries need to be filed.

Table 3 Pair-wise comparison scores for sub criteria under *data availability*

Data Availability	*Cost*	*Frequency*	*Timeliness*
Cost	1	2	1/2
Frequency	1/2	1	1/3
Timeliness	2	3	1

Note: It is obvious that diagonal will be unity. Also, note how AHP pair wise scores are symmetric, but need not be transitive. Boxes are made only for entries for which inputs are required. Others are generated.

For the proposed WHW framework and the structured set of criteria, eight matrices are required to be filled as given in Appendix 2. In total, 24 input data points are needed to get the full information on the eight matrices..

Once we obtain the pair-wise values, the weights are determined by finding the elements in the eigenvector associated with the maximum eigen value of the matrix. Instead of going to relatively advanced matrix algebra, a simpler alternative proposed in the literature, where weight of each criterion is equal to normalized geometric mean of the values in corresponding row.

$$w_j = \frac{\left(\prod_{k=1}^{p} y_{jk}\right)^{1/p}}{\sum_{j=1}^{m}\left(\prod_{k=1}^{p} y_{jk}\right)^{1/p}} \qquad (2)$$

For the sample scores given in Table 3, the resulting weights of *cost*, *frequency* and *timeliness* are, 0.297, 0.163 and 0.540.[10]

Typically, the criteria are weighted by decision-makers to reflect their relative importance (CIFOR, 1999). In order to find weights of each criterion for the proposed WHW framework, 12 experts are chosen from six higher learning institutes of the city in the domain of social science, technical and management, population and demography and economics and development. The experts have worked in the area of indicator research; five of them are faculty members and seven are research scholars.[11] They are provided with the criteria matrices with an instruction sheet explaining the meaning of each criterion (Appendix 1) and a background on the initiative.

Based on the responses, weights for criteria relative to each other are calculated (see Table 4). Among the criteria under *whom*, the criterion *relevance to policy* dominates and takes approximately half of the weight, whereas *user friendliness* takes a weight of one-third and *attractiveness to media* gets a weight of one-sixth. Similarly, among the criteria under *what*, the criterion *relevance to objective* gets a weight close to half, whereas *leading* and *scope* criteria get weights close to one-forth. Also, the *sufficiency* and *accuracy* criteria among the criteria under *data quality* and *data measurability*, dominates over other two criteria in their respective branches. In rest of the branches, weights of the criteria are close to having equal share.

3.2 Scoring of indicators

Once the criteria weights are obtained, the potential list of indicators to be scored against each criteria and the certain number of indicators from top can be shortlisted as the final list.[12] To obtain the scores of indicators against each criterion, AHP will not be preferred for the following reasons. Since the number of potential indicators can be high, the dimension of pair-wise matrices will be large, which would necessitate a huge number of inputs as every indicator has to be weighed against every other indicator and the process has to repeat for each criterion.[13] So, like most of multi-criteria decision analysis, we will score the indicators on a five point intensity scale; 1–poor, 2–average, 3–good, 4–very good and 5–excellent and apply a model for aggregation to obtain the overall score.

Table 4 Weights of Criteria under WHW framework

Criteria codes		*Partcipants*												*AVG*
		P1	*P2*	*P3*	*P4*	*P5*	*P6*	*P7*	*P8*	*P9*	*P10*	*P11*	*P12*	
WHAT	1	0.143	0.238	0.333	0.333	0.400	0.163	0.333	0.333	0.333	0.333	0.333	0.333	**0.301**
	1.1	0.333	0.387	0.540	0.400	0.429	0.286	0.627	0.500	0.500	0.540	0.413	0.400	0.446
	1.2	0.333	0.169	0.297	0.400	0.429	0.571	0.254	0.250	0.250	0.163	0.260	0.200	0.298
	1.3	0.333	0.443	0.163	0.200	0.143	0.143	0.118	0.250	0.250	0.297	0.327	0.400	0.256
HOW	2	0.286	0.625	0.333	0.333	0.400	0.540	0.333	0.333	0.333	0.333	0.333	0.333	**0.376**
	2.1	0.200	0.226	0.400	0.333	0.240	0.309	0.309	0.327	0.400	0.260	0.200	0.400	0.300
	2.1.1	0.210	0.582	0.540	0.200	0.196	0.196	0.143	0.413	0.333	0.400	0.196	0.571	0.332
	2.1.2	0.550	0.309	0.163	0.400	0.493	0.493	0.714	0.327	0.333	0.400	0.311	0.286	0.398
	2.1.3	0.240	0.109	0.297	0.400	0.311	0.311	0.143	0.260	0.333	0.200	0.493	0.143	0.270
	2.2	0.400	0.101	0.400	0.333	0.550	0.582	0.582	0.260	0.200	0.327	0.400	0.200	0.361
	2.2.1	0.149	0.122	0.540	0.311	0.143	0.540	0.498	0.260	0.250	0.250	0.500	0.200	0.314
	2.2.2	0.474	0.320	0.163	0.196	0.429	0.297	0.135	0.413	0.250	0.500	0.250	0.400	0.319
	2.2.3	0.376	0.558	0.297	0.493	0.429	0.163	0.367	0.327	0.500	0.250	0.250	0.400	0.368
	2.3	0.400	0.674	0.200	0.333	0.210	0.109	0.109	0.413	0.400	0.413	0.400	0.400	0.338
	2.3.1	0.200	0.584	0.527	0.400	0.400	0.540	0.455	0.333	0.333	0.333	0.333	0.333	0.398
	2.3.2	0.400	0.184	0.290	0.200	0.400	0.297	0.091	0.333	0.333	0.333	0.333	0.333	0.294
	2.3.3	0.400	0.232	0.183	0.400	0.200	0.163	0.455	0.333	0.333	0.333	0.333	0.333	0.308
WHOM	3	0.571	0.136	0.333	0.333	0.200	0.297	0.333	0.333	0.333	0.333	0.333	0.333	**0.323**
	3.1	0.493	0.297	0.367	0.582	0.547	0.558	0.333	0.400	0.540	0.429	0.493	0.559	0.467
	3.2	0.196	0.086	0.265	0.109	0.190	0.122	0.333	0.200	0.163	0.143	0.196	0.089	0.174
	3.3	0.311	0.618	0.367	0.309	0.263	0.320	0.333	0.400	0.297	0.429	0.311	0.352	0.359
	3.3.1	0.163	0.122	0.163	0.400	0.443	0.168	0.276	0.200	0.500	0.327	0.285	0.413	0.288
	3.3.2	0.540	0.558	0.297	0.200	0.387	0.484	0.128	0.400	0.250	0.413	0.198	0.260	0.343
	3.3.3	0.297	0.320	0.540	0.400	0.169	0.349	0.595	0.400	0.250	0.260	0.518	0.327	0.369

Note: Last column gives average. The weights for top level criteria *what*, *how*, and *whom* are made bold, and that of next level are underlined.

4 Aggregation

The organization of criteria in a value tree puts indicator into different logical groups and levels. This structure helps in step-wise aggregation like overall score for a higher level criterion and can be obtained by aggregating scores for criteria down the level. For example, the score of *user friendliness* (under *whom* dimension) is an aggregation of scores in *transparency*, *understandability*, and *sensitivity* criteria. The overall score in *whom* dimension is obtained by aggregating the score of *policy relevance*, *media attractiveness*, and *user friendliness* criteria. Similarly, the overall scores in *what* and *how* dimensions can be constructed by finding scores down the level. The overall score for all criteria can be obtained by aggregating the scores in all the three dimensions; *what*, *how* and *whom*.

For aggregation across criteria in multiple dimensions, linear additive technique is not suitable as it imposes perfect substitutability assumption across dimensions. Under linear additive model one indicator which scores poor (corresponds to score 1) in one criterion, say data quality, and excellent (corresponds to score 5) in another criterion, say data availability, will have an overall score, which is same as another indicator which scores good (corresponds to score 3) in both the criteria.[14] This is counter-intuitive, as a poor quality data, even if readily available may not suffice the purpose; and, hence, this indicator must have an overall score less than the indicator which is good in both the criteria.

An alternative to linear addition is displaced ideal (DI) technique developed by Zeleny (1974), which is based on the notion that better system should be closer to ideal.[15] The overall score is calculated as the inverse of the Euclidian distance measuring shortfall from the ideal. Ideal denotes maximum score in all dimensions. For n number of indicators along n dimensions, ideal, I would be

$$I = (x_1^*,\ x_2^*,\ x_3^*,\ \ldots x_n^*)$$

x_j^* = condition for maximum score; i.e., max (x_j) for maximization criteria or min (x_j) for minimization criteria. By normalizing to the scale of [0,1], 0 being least favored and 1 being most, the ideal point would be defined by unity vector, I=(1,1,…1). The diametrically opposite point would be known as origin or least–favored point O=(0,0,…0). For presentation purpose, in Figure 3, a two-dimensional criteria scenario (c_1, c_2) is considered with two options (p_1, p_2) having scores (x_{11}, x_{12}) and (x_{21}, x_{22}), respectively.

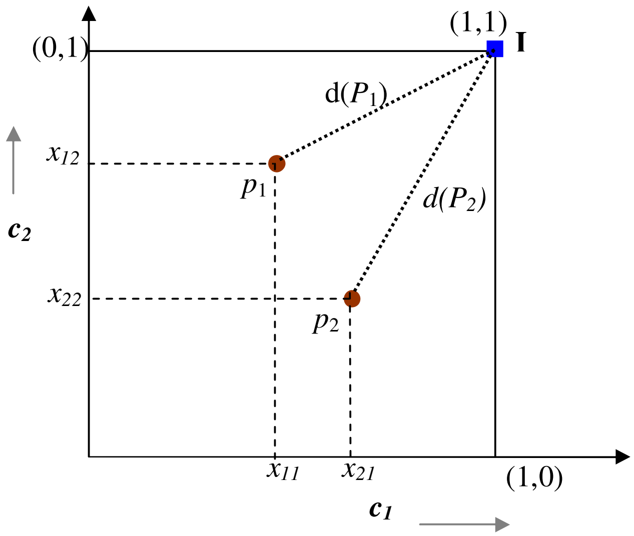
The distance of option p_i from ideal point is $d_i = \sqrt{(1-x_{i1})^2 + (1-x_{i2})^2}$. Hence, p_1 will have higher, equal or lower rank than p_2 if $d_1 > d_2$ or $d_1 = d_2$ or $d_1 < d_2$, respectively.
The overall score, based on displaced ideal technique is given by,

$$S_i^{DI} = 1 - \left(\frac{\sum_{j=1}^{m} w_j \left(1 - x_{ij}\right)^2}{\sum_{j=1}^{m} w_j} \right)^{\frac{1}{2}} \tag{3}$$

For equal weights, (3) is simplified to,

$$S_i^{DI} = 1 - \sqrt{\frac{1}{n} \sum_{j=1}^{m} \left(1 - x_{ij}\right)^2} \tag{4}$$

Figure 3 Displaced Ideal Method



Now, considering the example of two equally weighted criteria (*data quality*, *data availability*) and two indicators with scores of (poor, excellent) and (good, good), the overall scores can be calculated by both methods— linear additive and displaced ideal (see Table 5).

Table 5 Comparison between overall scores obtained through linear additive and displaced ideal methods

	Data Quality	*Data Availability*	*Overall Scores*	
			Linear Additive	*Displaced Ideal*
Indicator 1	1 (0.0)	5 (1.0)	0.500	0.293
Indicator 2	3 (0.5)	3 (0.5)	0.500	0.500

Note: The value in the parenthesis show normalized value, which is calculated as value=(actual–minimum)/(maximum–minimum). So normalized value for a score average is (2-1)/(5-1)=0.25. Similarly normalized score for poor, good, very good, and excellent are 0.0, 0.5, 0.75 and 1.0 respectively.

Under displaced ideal, *indicator* 2 fared better than *indicator* 1. This result supports what commonsense would suggest. Displaced ideal method of aggregation can be used along the braches of value tree (Figure 2) to find the overall score of each indicator. Figure 4a shows how lower level branches aggregate to give overall score, which gets further aggregated to get overall score in *what*, *how* and *whom* dimensions. Figure 4b shows the final resultant of three dimensions giving the final score for each indicator.

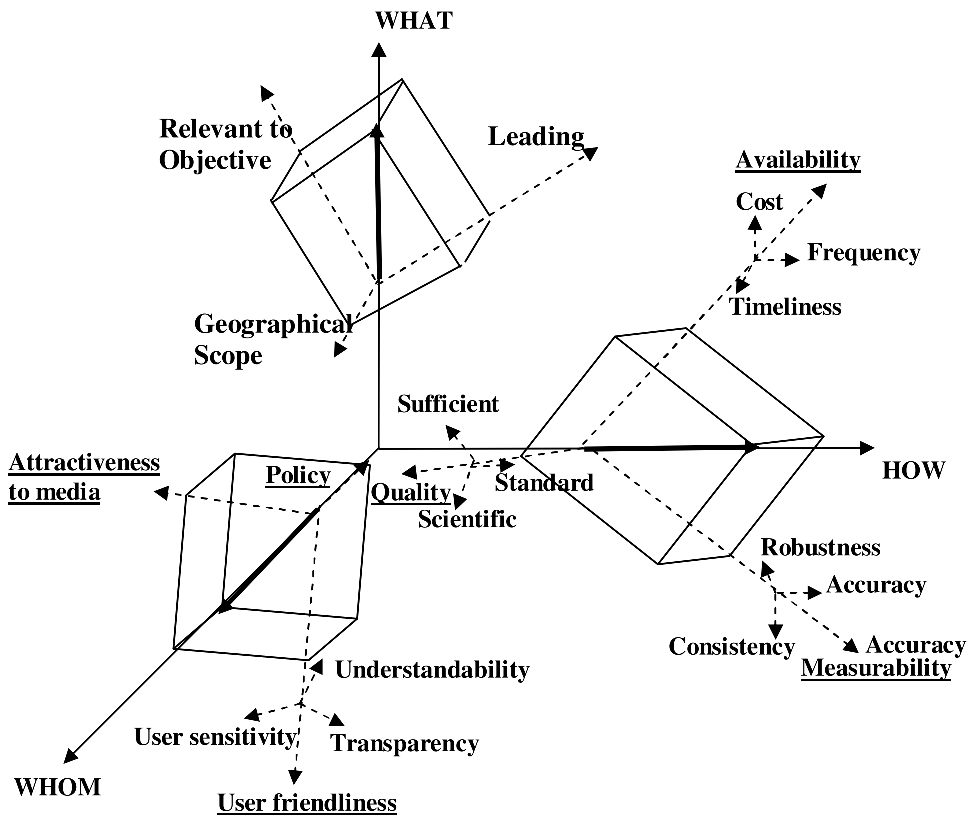


Fig 4a Components of Criteria Value Tree

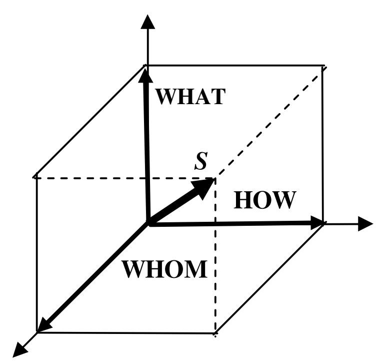


Fig 4b Resultant

5 Conclusion

The study has proposed a framework for criteria for selection of sustainable development indicators (SDIs). This framework is built from the basic functionality of the indicator, which is to transfer relevant information from physical space to users through communicable data. Accordingly, a three-dimensional What–How–Whom (WHW) framework has been conceptualized. Criteria under different dimensions are suggested and the complete value tree has been constructed. The exhaustiveness of WHW framework is tested by establishing correspondence of criteria considered in past initiatives with the proposed framework.

MCA methods are reviewed to ascertain their applicability in SDI research. AHP technique turns out to be most suitable for finding weights of criteria relative to each other. By involving experts in the field, the weights of different criteria in the proposed WHW framework are determined. For obtaining the overall score of indicators from the individual score for each criterion, 'displaced ideal'-based aggregation method is proposed over the usual linear additive model. Perfect substitutability assumption in linear model allows an indicator performing low in one criterion to neutralize its poor performance through an equivalent high performance in other criterion. However, these extremely behaved indicators have less practical use compared to indicators which performs moderately in all criteria. The newly proposed 'displaced ideal' method takes care of this concern.

Notes

[1] OECD (1993) specifies indicator as a value derived from parameters, which provides information about a phenomenon. UNCHS (2004) defines indicators not to be data, rather models which simplify a complex subject to a few numbers that are easy to grasp and understand. Indicators are considered to be small windows that provide the glimpse of big picture (Sustainable Seattle, 2004).

[2] Here, multiple levels mean the criteria may form layers among themselves such that some turn out to be top level criteria and others from the branch criteria. Multi-dimensionality means, being at the same level, one criterion may not be directly related to other and both of these may not have any direct relation to third.

[3] Australia (1998) has grouped the indicators into five categories a) Important, b) feasible, c) credible d) understandable and e) usable. OECD divides the criteria into three lists a) policy relevance and user utility, b) analytical soundness, and c) measurability. South African initiative divides the criteria into a) scientific, b) functional and c) pragmatic.

[4] Nathan and Reddy, (2008) is a companion paper where the authors have established the conceptual framework to develop SDIs to assess sustainability of resource use in Mumbai.. 'Urban sustainability' is split into three dimensions: 'economic efficiency', 'social wellbeing' and 'ecological acceptability'. The different domains of resources identified for the purpose are energy, water, land, atmosphere, population, finance, housing, infrastructure, waste and miscellanies.

[5] Since the present research is intended to study urban sustainability, the indicator initiatives chosen here relate to the same.

[6] Dominance occurs when one option performs at least as well as another on all criteria and strictly better than the other on at least one criterion.

[7] One example from outside the domain of SDI research is Human Development Index (HDI), where for instance life expectancy is used to represent the health dimension.

[8] Rank reversal is the simply by adding another option to the list of options being evaluated, the ranking of two other options, not related in any way to the new one, can be reversed.

[9] The MCA methods other than AHP, viz. methods based on dominance or conjunctive and disjunctive models are applicable at the stage of scoring the options. Linear additive model turns out to be complicated because of the simultaneous variation of weights and scores.

[10] To calculate the geometric mean across the rows the entries along the row are multiplied and the corresponding cube roots are obtained, which are 1.000, 0.550 and 1.817 respectively. Each of these values are normalized by dividing with the sum i.e. 3.367 to get the weights such that $\sum w_j=1$.

[11] Details of the institution profile are kept anonymous; these can be provided on request to interested readers.

[12] The discussion on potential list of indicators and the scoring of these indicators to arrive at the final list is out of the scope of the paper.
[13] For instance, in case of 10 indicators there would be 45 entries for each criterion, and for the proposed WHW framework, there would be 765 entries by the decision makers.
[14] Assuming equal weight for both the criteria data availability and quality
[15] In a proposed development of HDI measure, Displaced Ideal technique has been used by Nathan et al (2008), and Mishra and Nathan (2008). Also, Sharma (2008) has used the same technique in construction of an index for financial inclusion.

References

Ambiente Italia (2003) 'European Common Indicators – Towards a Local Sustainability Profile', *Final Project Report*, Ambiente Italia Research Institute, Milano, Italy, May.

Anon (1998) 'Environmental Indicators for National State of Environment reporting – human settlements', Department of Environment, Australia.

Bell, S. and Morse, S. (2003) 'Measuring sustainability: Learning from Doing', London: Earthscan, p.50.

Belton, V,. and Gear, T. (1983) 'On a short-coming in Saaty's method of analytic hierarchies', Omega, 11, pp.228-30.

Boulanger, P-M. (2007) 'Political uses of social indicators: overview and application to sustainable development indicators', *Int. J. Sustainable Development*, Vol. 10, Nos. 1/2, pp.14–32.

CIFOR (1999) 'Guidelines for applying Multi-Criteria Analysis to the Assessment of Criteria and Indicators, The Criteria and Indicator Toolbox Series, Center for International Forestry Research.

DEAT (2001) 'National Core Set of Environmental Indicators", *For State of Environment Reporting in South Africa*, Phase I, *Scoping Report*, National Department of Environmental Affairs and Tourism, May.

Dhakal S. (2002) 'Report on Indicator related research for Kitakyushu Initiative', Ministry of Environment, Japan.

French, S. (1988) Decision Theory: an Introduction to the Mathematics of Rationality, Ellis Horwood, Chichester, pp. 359-361.

Golden, B., Wasil, E. and Harker, P. (eds.) (1989) The Analytic Hierarchy Process: Applications and Studies, Springer Verlag, New York.

Goodwin, P. and Wright, G. (1998) Decision Analysis for Management Judgement, second edition, John Wiley, Chichester.

Ireland (2003) 'Measuring Ireland's Progress', Volume 2, 2003 – Background Report, Central Statistics Office, Government of Ireland.

Keeney, R.L, and Raiffa, H. (1976) Decisions with Multiple Objectives: Performances and Value Trade-Offs, Wiley, New York.

Laloë, F. (2007) 'Information at the interface between supply and demand for indicators: the use of a model with estimated parameter values as a base for indication in multicriteria analyses', *Int. J. Sustainable Development*, Vol. 10, Nos. 1/2, pp.33–45.

Mishra, S, and Nathan, Hippu S.K. (2008) 'On a class of Human Development Measure' Indira Gandhi Institute of Development Research (IGIDR), Working Paper, WP-2008-020.

Nathan, Hippu S.K., and B. Sudhakara Reddy (2008) 'A conceptual framework for development of sustainable development indicators.' Indira Gandhi Institute of Development Research (IGIDR), Working Paper, WP-2008-003.

Nathan, Hippu S.K., Srijit Mishra and B. Sudhakara Reddy (2008) 'An Alternative Approach to Measure HDI.' Indira Gandhi Institute of Development Research (IGIDR), Working Paper, WP-2008-001.

OECD (2003) 'OECD Environmental Indicators: Development, Measurement and Use', *Reference paper,* Environmental Performance and Information Division, Organization of Economic Co-operation and Development, Paris.

Parris, T.N. and Kates, R.W. (2003) 'Characterizing and measuring sustainable development', *Review of Environment and Resources*, Vol. 28, pp.559–586.

Pinter, L., Hardi, P. and Bartelmus, P. (2005) Sustainable Development Indicators. Proposals for the Way Forward, International Institute for Sustainable Development, p.3, available from: http://www.iisd.org/publications/pub.aspx?pno=769.

Rey-Valette, H., LaloeÈ , F. and Le Fur, J. (2007) 'Introduction to the key issue concerning the use of sustainable development indicators', *Int. J. Sustainable Development*, Vol. 10, Nos. 1/2, pp.4–13.

Savage, L.J. (1954) *The Foundations of Statistics*, Wiley, New York.

Sharma, M.(2008) 'Index of Financial Inclusion' Working Paper No. 215, Indian Council for Research on International Economic Relations.

Shim, J.P.(1989) 'Bibliographical research on the analytic hierarchy process (AHP)' *Socio-Economic Planning Sciences*, Vol 23, pp.161-7.

Sustainable Seattle (2004) 'Indicators of Sustainable Community -A status report on long-term cultural, economic, and environmental health for Seattle/King County', Washington.

UNCHS (2004) 'Urban Indicator Guidelines', *Monitoring the Habitat Agenda and Millenium Development Goal*, United Nations Center for Human Settlement, UN Habitat, Nairobi.

UNCSD (1996) 'Indicators of Sustainable Development Framework and Methodologies', Division for Sustainable Development, United Nations, New York.

Von Neumann, J, and Morgenstern, O. (1947) Theory of Games and Economic Behaviour, second edition, Princeton University Press, Princeton.

Winnipeg (1997) 'City of Winnipeg Quality of Life Indicator', *International Institute of Sustainable Development* (IISD), Vol? .

Zahedi, F. (1986) 'The analytic hierarchy process: a survey of the method and its applications', *Interfaces*, Vol 16, pp.96-108.

Zeleny, M. A. (1974) 'Concept of Compromise Solutions and Method of Displaced Ideal', *Computational Operation Research*, Vol 1 (?): pp 479-496

Appendix

Table 1A The explanation of criteria with examples

1. WHAT	**1.1 Relevant to Objective:** How close the indicator is to the subject matter it intends to indicate? How close it is to its objective? *Example:* Let the objective is to measure 'economic efficiency' in 'energy' domain for a city. Between two indicators: Per capita energy consumption in Mumbai and Energy consumption per unit GDP; the later is more relevant as it involves both input (energy) and output (GDP), which is a characteristic of efficiency.	
	1.2 Geographical Scope: Whether the indicator's scope matches with the geography in question? In case of a sampling, scope would signify representativeness of the sample. *Example:* Let the geographical scope of the SDI initiative is for Mumbai Municipal Corp (Greater Mumbai). Between two indicators, percentage of houses without piping water in Mumbai Sub urban area and percentage of houses without in-house toilets in Greater Mumbai, the later will get a higher score in geographical scope, as the former only represents part of the Greater Mumbai.	
	1.3. Leading: This refers to the inherent characteristics of Indicator to be leading, which means the ability of indicators to guide future actions. *Example*: Proven reserve to production ratio will score more in leading characteristics than ratio of production capacity to actual production	
2. HOW	**2.1 Quality**	**2.1.1 Scientific:** This characteristic refers to the theoretical soundness of the indicator. *Example*: The official data on houses without electricity which does not account theft of electricity may not be scientific.
		2.1.2 Sufficiency: Sufficiency refers to adequate information to provide suitable picture of the situation not too much information, nor too little; not over aggregated, or under aggregated. *Example*: Fuel mix of all income groups will be of too much of information whereas household shares using non-commercial fuels in total may be just right.
		2.1.3 Standard: The indicator must be standard enough for comparison with other geographical regions. *Example*: Fuel wood use may be typical to India, which may not be comparable with other countries
	2.2 Availability	**2.2.1 Cost:** This characteristic refers to the cost of making the indicator available. *Example:* An example of a costly indicator is to find the exact energy

		use pattern for each income group.
		2.2.2 Frequency: This refers to the frequency of availability of data. *Example:* Some indicators based on NSS data on energy consumption by households are available once in ten years, which is too less a frequency; may be a yearly data would be of right frequency.
		2.2.3 Timeliness: This says on timeliness of availability of data. *Example:* Census data in India is useful, but it is not timely because of delays.
	2.3 Measurability	**2.3.1: Accuracy:** This measures the degree of closeness of the indicator to the exact situation. *Example:* An indicator like proven reserve to production ratio may not be an accurate indicator due to inaccuracies in proven reserve.
		2.3.2: Consistency: The indicator must be consistent over time. *Example:* An indicator like energy affordability will not be consistent if the criteria for affordability will not be consistent
		2.3.3 Robustness: This is quality of being able to withstand stresses, pressures, or changes in procedure or circumstance. *Example*: An indicator like accident fatalities may not be robust indicator as it is subjective to the measurement methodology
3. WHOM	**3.1 Policy Responsiveness:** This refers to the relevance of the indicators for policy. *Example:* Between renewable share and non carbon fuel share, the former may be more relevant from policy point of view in India as we have a separate ministry for new and renewable energy and they get more policy attention.	
	3.2 Attractiveness to Media: This parameter evaluates the media interest with the particular indicator. The media actually brings the indicator in public notice. *Example:* Between the T&D loss and loss in conversion at production site, the former may be more attractive to media; because the later may be constrained by the technology whereas the former accounts for theft and maintenance issues.	
	3.3 User Friendliness	**3.3.1 Transparent**: Transparency means the indicator must be accessible to users. *Example:* Between two indicators; proportion of people availing public transport and proportion of people having pvt. vehicles; the first of the indicators may be more accessible than second.
		3.3.2 Understandable: This refers to the simplicity of the indicator. It must be easy to understand. *Example*: Between GHG emissions per unit use of energy and Renewable energy share in total production the later may be easily understood.
		3.3.3 User Sensitivity: This characteristics indicates the sensitizing capability of the indicator. *Example:* Share of population without electricity is more sensitive indicator than population using non commercial resource for cooking

Table 2A The Matrices to be filled (only blank boxes to be filled)

Criteria

	What	*How*	*Whom*
What	1		
How		1	
Whom			1

Data Availability

	Cost	*Frequency*	*Timeliness*
Cost	1		
Frequency		1	
Timeliness			1

WHAT

	Rel. to Obj.	*Leading*	*Geo. Scope*
Rel. to Obj.	1		
Leading		1	
Geo. Scope			1

Data Measurability

	Accuracy	*Consistency*	*Robustness*
Accuracy	1		
Consistency		1	
Robustness			1

HOW

	Quality	Availability	Measurability
Quality	1		
Availability		1	
Measurability			1

WHOM

	Policy rel.	*Media attraction*	*User friendliness*
Policy relevance	1		
Media attraction		1	
User friendliness			1

Data Quality

	Scientific	Sufficient	Standard
Scientific	1		
Sufficient		1	
Standard			1

User friendliness

	Transparency	*Understandability*	*User sensitivity*
Transparency	1		
Understandability		1	
User sensitivity			1