EVALUATING THE IMPACT OF WATERSHED MANAGEMENT PROJECTS: USING INDICATORS AS AN ALTERNATIVE TOOL

Michael W. Bollom • IGBP • New Delhi

INTRODUCTION

Natural resource management projects have traditionally been assessed through benefit-cost ratios. Calculating these ratios tends to be expensive, time consuming and difficult to implement. Computing these ratios requires that the total benefits of a project are divided by its costs. If the ratio is greater than one, the project is considered a success. For example, if there is an investment of Rs. one lakh in an irrigation scheme and the benefits derived from it are worth Rs. two lakhs, then the benefit-cost ratio is two.

Internal rates of return are benefit-cost ratios that have been discounted for time. Discounting is extremely important because project investments are often separated from payoffs by large periods of time. For example, Rs. one lakh invested in a forestry project today may, after a sufficiently long period of time, produce Rs two lakhs of return (benefit-cost ratio of two). The problem, however, is that it is not meaningful to compare Rs. one lakh invested today with Rs. two lakhs recovered thirty years from now. In all likelihood, Rs. two lakhs thirty years from now will be worth less than what Rs. one lakh is today (i.e., a benefit-cost-ratio of less than one). For this reason, internal rates of return are adjusted for time (or "discounted").

Even if all benefits are specified, taking exact measurements of these changes is often very resource intensive. Moreover, there is the added problem that not all the benefits are easily quantifiable

Adjusting for time is in itself a difficult issue, but estimates can be made (often using long-term inflation or depreciation rates). It is more difficult to quantify project benefits (whether they are reaped today, or in three decades). The first difficulty lies in determining all project benefits. For example, what are all the benefits of an irrigation project? Would a list of benefits only include increased crop yields, or should it also include improvements in public health, lower migration rates, higher incidence of school attendance, changing gender relations, etc.? Even if all benefits are specified, taking exact measurements of these changes is often very resource intensive. Moreover, there is the added problem that not all the benefits are easily quantifiable. How can the benefit of phenomena like changing gender relations be quantified? Yet benefit-cost ratios require quantification.

Indicators can be a fast, cheap and easy alternative to making the direct, quantified measures of project impacts needed for benefit-cost ratios and internal rates of return. Simply put, an indicator is a proxy measure – one easily-measured phenomenon, which is closely related to a target phenomenon that is more difficult to measure. If some phenomenon can be measured directly, indicators are unnecessary. It is only when a direct measurement cannot be taken that an indicator needs to be employed.

Indicators are also used to indirectly measure concepts that are abstract and/or complex. This is often the case in the realm of natural resource management. For example, how does one directly measure bio-diversity? The concept itself is quite abstract, and even when precisely defined, its direct measurement would be a Herculean task. In this case, the presence and vitality of one particular species might be used as an indicator of the entire region's bio-diversity. In the American Pacific Northwest, the spotted owl population has been used as an indicator of bio-diversity in coastal forests.

This article presents indicators as an alternative to traditional methods of evaluation such as benefit-cost ratios. The first third of the article discusses indicators in general. The remainder of the article then presents detailed information on nine indicators that have been developed by the Indo-German Bilateral Project "Watershed Management" (IGBP) to evaluate its Representative Watershed (RWS) Programme.

The Qualities of a good indicator

The difficulty with indicators lies not just in selecting ones that can be easily measured, but in selecting valid indicators – that is, indicators that are suitable proxies for the objectives ultimately being measured. Gauging the validity of an indicator can be difficult. For example, The Economist magazine has for several years used an indicator of foreign exchange rate volatility based on the local currency price of a McDonald's Big Mac® relative to its price in the United States. While The Economist itself presents arguments as to why this may or may not be a valid indicator of exchange rate valuation, there is reason to believe that it is a valid indicator of future expected exchange rate fluctuations.[1]

Indicator selection in the realm of natural resource management is difficult because there is no universally agreed upon set of indicators

Indicator selection in the realm of natural resource management is difficult because there is no universally agreed upon set of indicators. But, as Heidrun Traeger (1997, 79) argues, the search for a universal indicator set leads to a "dead end" because the choice of indicators depends upon the objectives of the programme being assessed and the natural conditions of the indicator application.

This means that the indicators chosen need to be programme-specific. That indicators should only seek to measure change concerning project objectives is heavily emphasised in the literature. The point (which is very much linked to the issue of causality) is that if a programme does not seek to affect a particular phenomenon, one can hardly attribute changes in the phenomenon to that programme. For example, if a project aims to preserve biological diversity, it makes little sense to employ indicators that measure ground water quality. This may lead observers to believe that the project has been responsible for an increase in ground water levels that has occurred recently.

In addition, indicators by themselves (e.g., the number of spotted owls in a fixed geographical area) have little meaning. Indicators only have meaning in the light of specified targets and threshold values. For example, biologists might determine that local bio-diversity is threatened when the number of spotted owl pairs in a one hundred square kilometre of coastal forest falls below twenty. Targets and thresholds must be specified for all indicators.

Useful indicators must be comparable across time and space. In order to assess change in one case over time, a researcher must be able to ascertain whether current measurements show significant variance from previous measurements. In order to assess variation across cases, the researcher must be able to show variance longitudinally. Both attempts to measure change require indicators that produce easily comparable data points. This is especially true as data sets grow larger.

In order to facilitate comparison, quantifiable indicators are often preferred. This preference is also because quantifiable indicators are often thought to be more reliable (i.e., they will produce the same results no matter who undertakes the assessment or the conditions under which it is undertaken) and replicable (i.e., deployable at different programme sites and in different settings). While reliability and replicability are of utmost importance, they do not require quantification. Indicators need only to be explicit (used here to mean "clearly defined and specified").

Given explicit specifications, qualitative/descriptive indicators are not necessarily inferior to numerical indicators. In fact, there are many instances when only qualitative indicators can be used. For example, odour, taste and colour might be the quickest, cheapest and easiest indicators to assess the quality of local drinking water. Odour is not directly quantifiable, neither are taste or colour. While some qualitative descriptions of water odour might be totally incomparable in nature (e.g., poems about the aesthetic qualities of different water samples), this does not need to be the case for all qualitative descriptions. Instead, a research protocol can clearly define a limited number of descriptive terms which would be employed to describe each water sample (e.g., sulphuric, brackish, etc.). Such a set of terms would be used to specify and compare the smell of water samples across space and time. If desired, such clearly specified qualitative indicators can, in the end, be numerically coded and used in a statistical analysis.

Indicators also need to be precise (or finely calibrated) measures in order to be useful. An indicator may be a valid measure of some phenomenon, but if it fails to register anything except very large changes, it is of little value. For example, the number of tractors owned in a village is an indicator of growing wealth among farmers, but such an indicator would fail to register small or even modest increases in income.

Even if an indicator is precise, it also needs to be responsive. That is, the indicator must register change soon after it has occurred. For example, children's height-for-age is an excellent measure of malnutrition, but the effects of malnutrition do not show up immediately in stunted growth—investigators will obtain such information only after several years.

If the assessment undertaken is only for the consumption of a select audience, only that audience needs to understand the indicators and accept them as valid

Some authors argue that only "widely accepted" indicators should be used. The logic behind this argument is that if the consumers of an evaluation do not believe that the indicator is valid, or if they do not understand it, they will not accept the results of the study. This is a good point, but it does not really support the argument that indicators need to be widely accepted. If the assessment undertaken is only for the consumption of a select audience, only that audience needs to understand the indicators and accept them as valid. Care only needs to be taken to select indicators that the target audience will accept as valid. Regardless of how widely accepted the indicators are, they should be devised with ease of comprehension in mind. Highly complex or abstract indicators will cease to have meaning, especially when results are shared with lay audiences.

Composite indicators versus indicator sets

In a quest for extreme parsimony many composite indicators have been developed by development organisations. Composite indicators amalgamate a wide range of information on different (but generally related) subjects into one index. A good example of this is the UNDP's Human Development Index (HDI), which combines information on rates of literacy, life expectancy and real per capita GDP to produce a single quantitative indicator of socio-economic development. While composite indicators are extremely easy to digest and compare, developing and computing them can be complex and time consuming. In addition, such indicators can be quite confusing and deceptive, given the manner in which the components of the index are combined.

The alternative approach, which is used in this article, is to use sets of individual indicators. Since one indicator by itself is generally not enough to assess impact across all project objectives, especially when the project has broad goals, practitioners must use indicator sets. The set must be assembled such that it contains indicators, that measure change in regard to all relevant indicators. These sets must contain a delicate balance between offering the maximum amount of information without overwhelming the consumer with too much data or exhausting project resources.

Causality

Employing a well-designed indicator set, investigators will be able to measure various aspects of change at programme sites. Often, however, it is difficult to determine whether the measured changes can be attributed to programme activities. The physical and social phenomena being dealt with in natural resource management and the conditions under which they are being measured are far too complex to isolate causality using only indicators. In fact, using indicators alone, it is likely that investigators will draw misleading conclusions. This is because non-programme factors (e.g., large-scale economic, social and environmental factors) are often responsible for the observed changes.

According to the noted expert in the field of impact assessment, Krishna Kumar, a solution to the causality problem is to measure "net impact". This entails subtracting all impacts that are caused by external forces from any measured changes. However, Kumar counters,

"Experience shows that the two methodological strategies for measuring net impact—quasi-experimental design and statistical controls . . . have not proven practical for use in agricultural and rural development projects. Both strategies pose major conceptual and methodological problems that are difficult to resolve satisfactorily. Moreover, they require massive and expensive data collection efforts which must be conducted over extensive periods of time." (Kumar 1989, 6)

The indicators themselves can provide information regarding the magnitude and direction of change. Qualitative studies then determine whether these changes can be attributable to project activities

Kumar's solution to this problem is to supplement data provided by indicators with qualitative studies. The indicators themselves can provide information regarding the magnitude and direction of change. Qualitative studies then determine whether these changes can be attributable to project activities. Kumar (Ibid., 7) recommends essentially three different methods for carrying out such studies: 1) in-depth, unstructured interviews with key informats such as government officials, project management staff, beneficiaries, and local leaders; 2) community meetings and focus group discussions with the same people; and 3) direct observation by experts. Such methods rely on the premise that local people understand their own environment. While such techniques are not foolproof, they can help establish a reasonable degree of certainty as to whether programmes are responsible for observed changes.

Whenever possible, this article supports Kumar's recommendations. When that is not possible (especially vis à vis phenomena that lay people do not readily observe, like erosion), the use of controls is necessary. Controls are often, however, expensive to use because they increase (often double) the number of measurements that must be taken. In addition, controls are often unable to filter out the effects of larger socio-economic trends.

Resource constraints

Given that there could be an almost infinite number of potential indicators to measure the impact of watershed management programmes, evaluators have the luxury of choosing ones that suit their unique needs. Most watershed management programmes, however, face resource constraints. As such, the programme must be executable under the following resource constraints.

Equipment costs

Equipment costs can vary greatly from indicator to indicator. An indicator the use of which entails only observations or oral surveys requires no tools (apart from writing utensils). A silt monitoring station, on the other hand, requires expensive equipment that is also costly to maintain.

The need for highly skilled research staff

Regardless of the tools employed, some indicators are more difficult to execute than others. While many factors can make a tool more or less difficult to use, the ease of use is defined here in terms of the amount of training beyond literacy that is necessary to employ an indicator. (It is important to note that this is different from the skill level necessary to design the indicator or to analyse the data from the indicator). For example, a literate villager can learn how to read a rain gauge and enter the amount onto a table. On the other hand, conducting a Participatory Rural Appraisal requires a great deal of formal education, in addition to specialised training.

Man hours

It is possible for the use of an indicator to require no special tools or training to use, yet its use could still be labour intensive. For example, a relatively unskilled person can measure ground water levels using local wells with very simple tools. However, he must take a sample every day of the year. Such an indicator is incredibly labour intensive.

The number of field visits required

Resource requirements are also affected by the number of site visits required to execute an indicator. This is because travel and per diem costs in the field can be expensive. Some indicators can be fully deployed in just one site visit. This is especially true when participatory methods are used. Other, more scientifically-measured indicators require multiple site visits in order to record change over time.

These four factors combine to determine the total resource demands of a particular indicator. Labour costs are the product of the number of man hours required and the necessary skill of that labour. Total costs are then a sum of this product with equipment costs and field visits:

Total Costs = (Man Hours * Skill Level) + Equipment Costs + Field Visits

All else being equal, indicators with lower total costs are preferred.

The indicator set

The following nine indicators were chosen by the IGBP to create an indicator set for evaluating watershed management programmes. This section examines these indicators in detail. Each indicator is discussed in terms of the objectives towards which it measures progress, how it should be implemented in the field, and recommendations for the continued use of the indicator.

Soil loss

While this indicator is resource intensive, both in terms of man-hours and equipment, it is recommended for continued use. This is due to the lack of a better alternative, and for the sake of continuity. (The IGBP has already invested in the infrastructure to use this indicator and trained people to operate it. In addition, the project has already accumulated large data sets.) This indicator seems impractical, however, for the replication stage of a programme when cost containment becomes more important.

Target objectives

This indicator measures topsoil conservation. The assumption is that increased soil runoff in local streams indicates a higher rate of topsoil loss. In addition, the recharge of ground water resources can be inferred from this indicator—all else being equal, less topsoil being washed downstream implies that more rainwater is being absorbed into the ground, thus recharging the ground water supply. Soil runoff also indirectly measures the extent of vegetative cover—less topsoil runoff implies thicker vegetative cover (vegetative cover prevents runoff by reducing splash erosion and holding soil in place with roots).

Soil runoff also indirectly measures the extent of vegetative cover – less topsoil runoff implies thicker vegetative cover

Measurement procedures

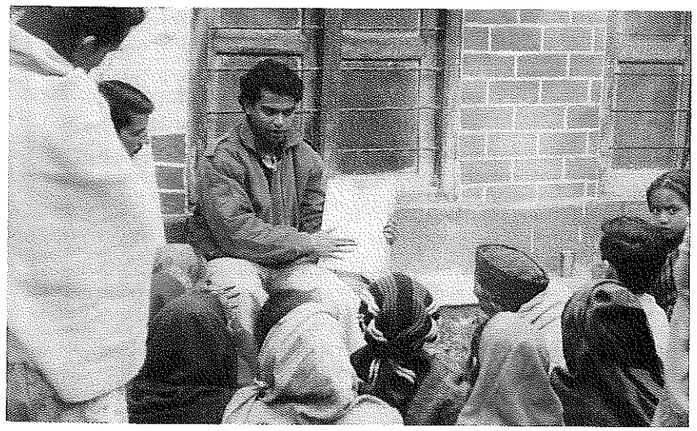
The ultimate goal of the hydrological monitoring undertaken here is to collect data demonstrating: how much water runs off in the stream that drains the watershed for a given amount of rainfall, how quickly this occurs after a rainfall, and how much sediment is carried away by the stream. To do this, a crew of silt observers must work around the clock to gather data on rainfall in the watershed, in addition to measuring depth of the stream at the drainage point, stream velocity and sediment concentrations in the stream.

In order to demonstrate changing runoff rates, this data must be collected over very long periods of time (not less than ten years). Successful anti-erosion treatments will result in decreased levels of discharge (meaning more water soaks into the ground). If this is the case, then the discharge that does occur will be spread out over a longer period of time (i.e., the water drains more slowly). In addition, silt loads will be decreased.

Outlook and recommendations

This indicator has many strengths. It produces valid, reliable data about soil (and water) conservation. It is also extremely responsive and sensitive – improved soil conservation will be evident in the fine-grained measurements taken during the very next rain. In its favour, very few (perhaps only one) monitoring stations are needed to take measurements for the area being treated. In addition, once the monitoring stations are in place, the measurements can be carried out by relatively unskilled labour. Finally, if meticulously carried out, monitoring this indicator can produce a very objective, and highly quantifiable database.

Unfortunately, this indicator is only suitable for evaluating pilot projects, not monitoring projects being implemented on a large scale. The indicator is very labour intensive. In an ideal situation, measurements must be taken every hour, twenty-four hours per day, three hundred and sixty five days per year. In order to carry the measurements out properly, this indicator also requires the use of some expensive equipment – instruments to measure rainfall, stream depth and stream flow. Given the remote, rural setting of most watershed programmes, maintenance of this equipment can be difficult. Poor maintenance and/or lax personnel will quickly corrupt databases. As with other scientifically measured indicators, this indicator also requires the use of controls.



Those desiring an indicator more suitable for monitoring large-scale projects at the replication stage might want to explore the use of remote sensing (i.e., satellite imagery). A series of satellite images taken over time can be compared to chart changing land-use patterns. Since satellite images (old and new) are often available (perhaps through the Remote Sensing Department of the local government), remote sensing would be fast and inexpensive. There may, however, be difficulties in obtaining such images if there are "national security" concerns involved. Such a method would

also still require the creation of a computer model to interpret the images, and initial field visits to help create a key to the satellite images.

Ground water

Target objectives

This indicator measures ground water conservation. If ground water levels are maintained (or even augmented), then ground water resources are being sustainably utilised. In addition, this indicator indirectly measures topsoil conservation. All else being equal, a higher level of ground water is at least partially the result of less (or slower) water runoff. This, in turn, results in lower levels of topsoil erosion. Water table levels are also an indirect measure of vegetative cover. As previously stated, vegetative cover prevents runoff, thus recharging ground water, by checking splash erosion and by holding soil in place with roots. If ground water levels are being sustained, constant or increasing vegetative cover may be responsible.

If ground water levels are maintained (or even augmented), then ground water resources are being sustainably utilised. In addition, this indicator indirectly measures topsoil conservation

Measurement procedures

A water level sensor should be used to measure the depths in village wells. Such a device is accurate up to less than a centimetre. Measurements could be taken in local wells, or special monitoring wells could be dug. While the former is less expensive, it is less reliable. The problem with such a low-tech methodology is that local people disturb the level of water in wells by drawing water from them. This problem could be circumvented if readings were taken early in the morning, before the day's first water is drawn out.

Since water levels in a well fluctuate from day to day and month to month, readings need to taken frequently, for the duration of the programme and beyond, as is done in the Silt Monitoring Stations. And as with the other extractive indicators, controls must be used. Taking the required number of readings will be expensive in terms of labour and, as mentioned, reliability is a problem.

Programme engineers need to select where in a watershed the ground water level should be monitored. Their decisions will be based on the knowledge of where programme impacts are expected.

Height-for-age

Target objectives

Anthropometric indicators are generally broken into three sub - indicators – height - for-age (otherwise known as stunting), height-for-weight (wasting), and upper-arm circumference. Height-for-age is selected here as the single best anthropometric indicator for health (primarily nutrition) because it registers long-term health status.[2] This is because growth cycles, which are missed due to periods of poor health, cannot be recovered. This growth is simply foregone forever. Children who have foregone growth cycles will register as significantly shorter than statistically established averages. Height-for-weight and arm circumference both give information about current nutritional status only. The stunting and wasting data have the added bonus of being objective, so it is less subject to criticism of bias.

Height-for-age is also an indicator of wealth for the very poor. The very poor often spend any increased wealth on food, which will register as increased height. Following the same logic, distributional analyses of height-for-age along gender and class lines are indicators of gender and economic equity. Finally, height-for-age is indirectly a measure of soil and water conservation – in a rural community, improved health is often linked to the raw materials of farming.

Measurement procedures

Children of known ages must be measured for height. Techniques for doing this are well documented and easily available, as are the international standardised tables that detail the distributions of height-for-age for children of different ages (for example, see the web page by Bender and Remancus listed in the Reference). A sample can be taken from several villages (this must include control villages), randomly or in terms of representativeness. Children of particular age groups are measured and their ages noted. Since standardised tables of height-for-age measures are only applicable internationally up until puberty, younger children must be the ones measured.

When the IGBP's evaluation team was preparing to visit a watershed, they asked partner NGOs to inform parents that the team would need to know the ages of their children. While very few children in the watersheds had proper birth certificates, most parents did have vaccination cards that listed their children's date of birth. These cards were accepted as accurate since children begin their vaccination sequences within a few months after birth, when the date of birth is still fresh in the mother's memory. The team also accepted cases where the accompanying parent simply recited a birth date, if she or he seemed quite confident. In cases where the birth date seemed uncertain, the case was excluded from the final database.[3]

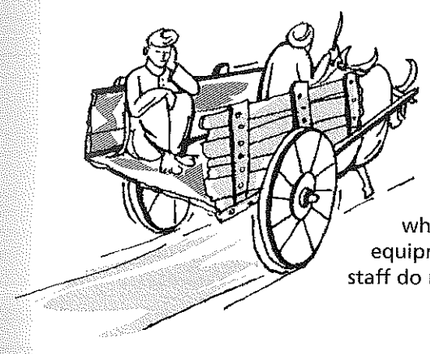
Stunting is determined by counting the number of children who fall below two standard deviations of accepted norms. Gender parity is determined by analysing how the measurements for female children fare in comparison with males

The preliminary Programme Evaluation Protocol recommended that two-year-olds be measured. Given the small size of the villages being studied, the team's sample size would have been too small had this limitation been maintained. Instead, local people were told to bring children who were old enough to walk, up to five or six years old. In retrospect, it was a disaster to include children less than two years old. In the best cases they were unable to stand still; more often their shrieks of terror set off a chain reaction of crying which cascaded through to much older children who would have otherwise remained calm.

In order to encourage people to bring their children (and in the interest of giving something back to the community) the Principal Investigator hired a local doctor to accompany the team on its site visits. The doctor's presence undoubtedly increased the level of participation in the study.

The data collected is then analysed. Stunting is determined by counting the number of children who fall below two standard deviations of accepted norms. Gender parity is determined by analysing how the measurements for female children fare in comparison with males. Economic equity is determined by looking at the overall spread of the scores – if the standard deviation of the scores is high, equity is low.

Finally, future evaluations need to set aside control groups. It may be more difficult to convince people in untreated areas to bring their children for measuring (they will not even know the NGO or state department people). The presence of the doctor (and the bag of sweets!) may compensate for this.



Outlook and recommendations

Of all the indicators tested, this one has the potential to be the most powerful, especially when considered in terms of resource constraints. For a relatively small investment of time and monetary resources, an evaluation team can gather a fine-grained database that is rich in information on nutrition, social equity and gender equity. (For almost no added costs, the evaluation team can also take measurements for wasting, which offers added information about the same issues.) Only a few pieces of equipment need to be purchased (costing less than Rs. 12,000), and the executing staff do not need to have any special skills. The study can be carried out quite

quickly (half a day per village), while freely available software (EpiInfo6 from the Center for Disease Control) makes analysis quick and easy.

According to international organisations like the World Health Organization, this is a valid measure of nutritional status.[4] Distributions of data on nutritional status can be examined to gain information on gender and social equity too. Unless an evaluation team has sampling difficulties (and if they follow the procedure discussed above, they will not), this is also a reliable indicator. The stunting and wasting data have the added bonus of being objective, so it is less subject to criticism of bias.

This indicator does have some weaknesses, however. It is not a responsive indicator—changes in the level of community health will take several years to show up in anthropometric surveys. In addition, height-for-age will not be a useful indicator of health or wealth in those communities that are already relatively healthy and wealthy (since human height eventually approaches physical limits). Finally, like all extractively executed indicators, controls are necessary for the proper implementation of this indicator. Nevertheless, the continued use of this indicator is recommended without any reservations.

Ownership of consumer durables

Target objectives

This indicator measures the level of wealth in a watershed. The assumption is that as general levels of wealth increase, the local population will purchase more consumer durables. The ownership of various, highly visible consumer durables is used as an indicator for several reasons. First, consumption levels of non-durables (alcohol comes to mind) are difficult to determine as people often do not monitor or remember their consumption rates.

The assumption is that as general levels of wealth increase, the local population will purchase more consumer durables

The second reason why this particular indicator has been selected is because people often wish to conceal their personal income. Highly visible consumer durables, especially the larger ones such as farm animals or bicycles, are difficult to conceal so they can be easily counted. The distribution of consumer durables is then used to measure economic equity.

Measurement procedures

Ownership of particular consumer items can be measured by a household survey or through a more participatory approach. The latter is recommended because it is faster, cheaper and more likely to uncover the truth. Obtaining figures of statistical significance does not require that the entire watershed be surveyed. Depending upon the number of villages, a sample of villages can be surveyed. If the villages in the watershed are very different in terms of socio-economic makeup, representative villages could be non-randomly selected. A skilled practitioner of participatory rural appraisal techniques can assess the levels of ownership of various consumer durables in the village in a PRA session lasting no more than two hours.

The consumer goods surveyed need to be selected with local culture and levels of wealth in mind. Appropriate consumer goods to survey will be those that local people aspire to buy, but are just out of their reach. For example, the number of snow shovels owned by Rajasthani villagers will probably not change even if their level of wealth increases dramatically. The goods surveyed must also be of the kind which others in the village would be readily aware of (e.g. a bicycle, more so than jewellery).

Prior to conducting evaluations, partner organisations were asked to help the evaluation team assemble a list of consumer durables that a few people in the watershed possessed, but most aspired to own. The beneficiaries were told that when the evaluation team arrived, the team would survey these items in the selected villages.

In order to develop the best possible survey list of consumer durables, a group of local people must be consulted. Evaluators must look for items that few own, but many would choose to own, given a modest increase in wealth. Such items will help evaluators measure change over time. If almost everyone in a village already owns an item, it is useless as a survey item. For example, if a village is already saturated with radios then it probably still will be when a second survey is executed several years hence. If this is the case, evaluators will not be able to register changing levels of wealth, even if they have occurred.

Evaluators would be well served to survey between twelve and fifteen items. This number is still a bit large, but the original list needs to be longer. This is because some items may need to be withdrawn from the list during successive evaluations. For example, many people in Kattery RWS presently want to own a mixie (blender). Five years from now, when a follow-up evaluation is being conducted, the mixie may have been superseded by a superior tool that does the work of a mixie and a grinder. If this were to happen, it would not make sense to continue surveying for mixies and the item would be omitted.

The lists of durables should then be turned into pictorial surveys by an artist. In such a survey, each item on each survey is represented by a separate picture. This was done both for illiterates as well as to stimulate discussion. In case the pictures were ambiguous, a caption was also included, in both the local language and in English. Since this article recommends that reconnaissance visits be dispensed with in the future, it is important that the evaluation team be prepared to construct these pictorial surveys in the field.

These surveys are used during the participatory sessions. During the discussions themselves, it works out best if the surveys are reserved until the end, when people are loosened up enough to talk about monetary issues. The team member leading the discussion asks the group to fill in the survey out together. They are asked to debate, one item at a time, how many of each are presently owned by the villagers. This number is then recorded, by one of the group members, in the left-hand box under each picture. In order to measure changing rates of ownership, time series data on changing ownership levels of consumer durables should be collected through periodic visits, several years apart.

When all the durables had been fully surveyed, the group should then be asked to discuss the reasons for change. For example, why do fewer people own cows than they previously used to? How is it that there has been an explosion in the ownership rates of kitchen implements like grinders and mixers? If they answer that people have more money now, they should be asked about the sources of the new-found wealth. Future evaluations should also use control groups to help pinpoint the mechanisms for changing levels of wealth.

Outlook and recommendations

This indicator is highly recommended for future use. It allows evaluators to gather data on a very delicate issue (wealth). It is extremely inexpensive to implement, requiring only the purchase of poster paper and coloured pens. It also takes little time to implement – no more than an hour per village. The skills necessary to

execute the indicator (the ability to lead a PRA in the local language) are somewhat sophisticated, but they are the same as the skills required for some of the other indicators in this set.

In addition, this indicator should be a valid measure of wealth for all but the poorest people. Given the nature of information obtained in group discussions, this data is also quite reliable. Reliability will, however, start to falter if the village is too large or the item surveyed for is too pervasive. For example, Mellodyarahatti in Kattery had almost seventy households. It was difficult for the groups to come up with accurate estimates of ownership for widely held items such as mixers.

To its detriment, this indicator is not very responsive – it takes time before beneficiaries transfer new found agricultural wealth into consumer goods. In addition, this indicator will not measure changes in wealth at the very lowest rung of the economic ladder. Those most in need will first spend increased resources on food. After that, they will expend their resources on shelter and debt repayment. This problem is taken care of in part by using this indicator in conjunction with an anthropometric indicator (see the previous section). A survey of consumer goods also fails to register productive investment that come with increased wealth (e.g. the purchase of fertilisers).

...it takes time before beneficiaries transfer new found agricultural wealth into consumer goods. In addition, this indicator will not measure changes in wealth at the very lowest rung of the economic ladder

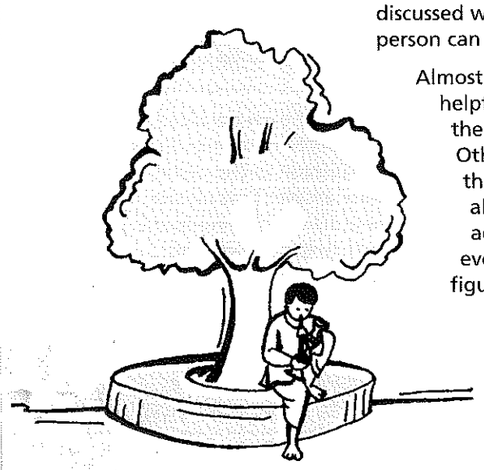
School attendance

Target objectives

This is a proxy measure for levels of education. In all but the worst cases, children become more educated the longer they attend school. Given that very poor people do not send their children to school, this is also an indicator of wealth – as the poor acquire more resources, they will send their children to school. The distribution of attendance data along gender lines also serves as an indicator of gender equity. Once again, this is an indirect indicator of soil and water conservation—in a rural community, the rising level of wealth necessary to attain higher levels of education is in most cases linked to the raw materials of farming.

Measurement procedures

Members of the evaluation team should take a single day's attendance at all the schools that serve the selected villages. This would be accomplished by simply arriving (unannounced, if possible) at the schools in question, and requesting a head count. In the spirit of participation and sharing data, attendance figures should be discussed with the principal/director of the school at the time of collection. This person can offer an interpretation of the data.



Almost all of the principals with whom team members spoke, were quite helpful. After asking about the nature of the evaluation, many simply opened their attendance registers and let the team collect the needed information. Others went to their record keeping area and supplied the information to the team. This latter method of operation highlights the weakness of our alternative method – data supplied to the evaluation team by school administrators may be subject to tampering. (The same is true, probably even more so, of enrolment figures. This was the reason why attendance figures were preferred to enrolment records in the first place.)[5]

Although the evaluation team did not actually carry out the head count method, visiting schools in Arki and Kattery gave rise to a concern about the invasive nature of the of the head count method. This problem was not evident when this method was first developed by the principle investigator after his visit to Karkara RWS, Bihar.

In Karkara the schools were fairly small and informal, as one might imagine village schools in a poor rural area to be. The secondary schools visited in Arki and Kattery were, however, much larger and more formally run. Had team members asked, the headmasters of these schools may have bristled at the idea of disrupting their classes to count the number of students, especially when the attendance had been already taken in the morning.

So both the methods of collecting attendance data have shortcomings. Head counts are invasive and official records can be falsified. If the local schools are small and informal, the head count method is recommended. If not, evaluators can attempt to be present during the normal morning attendance session (to do a parallel count). If neither of these options work, evaluators should use the daily attendance records available at the school.

If the local schools are small and informal, the head count method is recommended. If not, evaluators can attempt to be present during the normal morning attendance session (to do a parallel count)

Outlook and recommendations

Attendance is certainly a valid measure of schooling. School attendance is also a fast, cheap and easy indicator for measuring education. It requires very little time, no special equipment and can be carried out with low levels of training. This indicator is not, however, particularly reliable. Attendance on any one day is subject to many factors, including the weather, festivals and cropping patterns. As such, attendance figures gathered even on two successive days might be quite different. For this reason, it is still recommend that the data gathered be discussed with an administrator — he or she will know if the attendance on a particular day is unusual or not. (If a high degree of reliability is required, then enrolment figures should be used instead.) Another weakness of this indicator is that it requires the use of controls.

Use and maintenance

This indicator (referred to in the rest of the article as "Use") is very closely associated with the indicator "Outsiders". For this reason, many of the comments made here will also apply to "Outsiders". The text will make clear which statements apply to both. "Use" and "Outsiders" are the process indicators in this set. They do not attempt to measure the impact that a programme's activities have had. Instead, they give some idea how successfully the activities themselves are functioning. With such information, evaluators can better understand the mechanisms through which programmes have affected the treatment area, and also in some manner predict the future impacts.

Target objectives

This indicator measures project sustainability. If many of the units that have been installed under some activity are not functioning or improperly maintained while outside support is still coming in, it is likely that even fewer will function after outside support is withdrawn. Such an activity is not sustainable.

Measurement procedures

The execution of this indicator is highly time consuming. Determining the extent to which programme activities are being used and maintained requires that members of the evaluation team visit at least a sample of units from every activity in the watershed. During the test evaluation this entailed a lot of travelling, much of it over difficult terrain. If an evaluation must be completed very quickly, investigators may prefer to omit this indicator.

The evaluation team must survey all of a project's activities in order to determine how frequently they are being used and how well they are being maintained. At this

point sampling becomes an issue. During the present evaluation, if an activity had only a few units in the watershed, the evaluation team surveyed all of them. If this was not possible, only those units in the selected villages were surveyed. For some activities, so many units existed that even this was not possible. In such cases, random visits were made to the units.

Before going to the field, evaluators must define the terms "heavily used", "lightly used", "well maintained" and "poorly maintained" for each activity. For example, the evaluation team decided that smokeless stoves must be used for cooking everyday to be defined as "heavily used". Explicit definitions prevented qualitative appraisals from becoming too subjective. Developing these definitions can challenge the imagination, especially when it is difficult to determine what it means to "use" or "maintain" units of a particular activity. For example, how does one use a check dam? In such cases, only levels of maintenance can be checked for. Definitions of both use and maintenance should be formulated in consultation with the people who have designed the activity.

Outlook and recommendations

This indicator gave the evaluation team a clear understanding of how well the IGBP's work is proceeding, on an activity-by-activity basis. While it relates little about the project's impacts, it is essential for understanding the mechanisms through which the project is making an impact.

This indicator is a valid, if only partial, measure of programme sustainability—when units in an activity are being heavily used and highly maintained, it is very likely that the activity is sustainable. While it is not perfect, this indicator is also sufficiently reliable. A single field visit may miss periods of high or low use. To circumvent this the procedures for this indicator combine a field survey with participatory discussions. In addition, this simple indicator requires no special equipment and, because it is a purely participatory indicator, it does not require the use of control groups. For these reasons this indicator is strongly recommended, except in the case of extreme time constraints.

Dependence of programme activities on outsiders

Like Use, "Outsiders" is a process indicator that produces information about the sustainability of a activity. While Use presented information regarding use and maintenance of units in an activity, Outsiders tells evaluators whether the operation and maintenance of an activity is dependent upon outside expertise, funding, etc.. If it is, the activity will probably cease to function after project funds dry up.

Target objectives

If a project's local programmes are operated and/or managed by outside personnel, then levels of active local participation are lower. In addition, the presence of outsiders is an indicator of project sustainability and replicability. If local people cannot manage and operate an activity by themselves, it will eventually collapse when outside support is withdrawn. In addition, if an activity cannot be run without the help of outsiders, it is also less likely to be widely replicable in other areas.

Measurement procedures

In many cases, deciding whether someone is an "outsider" is quite easy. A foreign consultant brought in from abroad is an outsider. In some activities, however, the definition of who is actually an outsider can be difficult. Essentially, an "outsider" is someone who would not be involved with an activity were it not for programme funding. This can even include local people who carry out certain tasks under the

employ of the programme. For example, in the IGBP's Arki watershed one of the SUTRA staffers is a young woman who has always lived there. In all the activities analyzed she is an "outsider" in the sense that she will cease to carry out her current responsibilities once the IGBP withdraws funding and the SUTRA office shuts down.

Gathering the data involved informal discussions with people who use and/or operate units from all NGO and state department activities. The visits for Use and Outsiders were undertaken simultaneously. While investigating use and maintenance issues, the evaluation team asked questions regarding the people who keep each particular activity in operation. This involved obtaining answers to a series of questions: Who operates the units in this activity on a day-to-day basis (and where do they come from)? Who maintains them? Who supplies the spare parts? Where do the finances for on-going operations come from? If group action is necessary, who organises it? The answers to many of these questions required follow-up interviews with the people named, in order to determine why they are in the watershed and who pays their salaries. Information gathered in the field was verified during the participatory sessions, which took place later.

Outlook and recommendations

Participation and sustainability are very difficult concepts to concretize and measure. Like Use, this is a valid, if partial, measure of both — if the operation and maintenance of a activity is not dependent upon outsiders, then the activity has a much higher potential for survival when the programme funds cease. Given a thorough investigation of who operates and maintains an activity, this indicator should also be fairly reliable. While various evaluators may begin their investigations with different units, their questions should lead them to similar answers regarding those responsible for operation and maintenance of the overall activity.

Outsiders is, along with Use, a time-consuming indicator. On the other hand, no special equipment is required. Time constraints aside, this was a very important indicator for understanding the sustainability of a project's various activities. If, however, time constraints are severe, and if information about processes is significantly less important in comparison to studying impacts, it should be omitted from the indicator set.

Participation and sustainability are very difficult concepts to concretize and measure

Replication

Target objectives

This indicator is a measure of replicability. If local people replicate some programme output without support, it implies that there is a local demand for the units, a willingness to pay for them, and the necessary skills to construct, use and probably maintain them. In such a case, the programme is definitely replicable, at least in the surrounding areas. It is also likely to be replicable in other locations with similar geo-climatic conditions and socio-economic resources.

Measurement procedures

The methodology for this indicator is rather ad hoc. Evaluators simply need to scan for and inquire after evidence of programme outputs that have been upgraded or replicated without programme support. Inquiry is probably the best starting point, especially for information regarding replication, because copying of programme units may be taking place in remote areas, or areas outside of programme coverage. Any leads should be followed up and personally confirmed by evaluators. While gathering data on replication, evaluators should also inquire about facilities that have been upgraded or modified. All leads should be personally confirmed. Evaluators should also look for evidence of up-gradation and modification when they are conducting surveys for the indicators Use and Outsiders.

Table A
Final Evaluation of the Individual Indicators•

Indicator	Objectives Measured	Validity[a]	Reliability[b]	Precision[c]	Responsiveness[d]	Equipment Costs	Training[e]	Man hours[f]	No. of field visits[g]	Needs further refinement?	Comments
Soil Loss	Topsoil conservation	++	++	++	++	---	++	---	---	N	Highly labour intensive.
Ground Water (participatory method)	Ground water conservation	+	+	---	++	++	-	-	++	N	Not Recommended — Metric too imprecise to be of use in most projects
Ground Water (scientific method)	Ground water conservation	++	++	++	++	---	++	---	---	N	Highly labour intensive.
Height-for-age	Health, Wealth, Gender parity, Social equity	++	++	++	---	+	-	-	+	N	Slow to measure changes in objectives. Not useful in areas with pre-existing levels of socio-economic development.
Consumer Durables	Wealth, Social equity	++	+	++	-	++	-	-	++	N	Less useful in extremely poor areas.
School Enrolment	Education, Gender parity	++	-	++	-	++	-	+	+	Y	Collection of comparable time-series data sets difficult.
Use and Maintenance	Sustainability/ Replicability	+	+	-	++	++	---	-	++	N	Low validity due to incompleteness of the indicator.
Outsiders	Sustainability, Replicability	+	+	-	++	++	---	+	++	N	Low validity due to incompleteness of the indicator.
Replication	Replicability	+	-	-	-	++	---	+	++	Y	Very incomplete indicator. Not valid for relatively expensive schemes.
Social Capital	Sustainability	+	-	---	++	++	---	-	++	Y	Limited reliability. Incomplete indicator.

• In this table, the superior rating (whether that means highest validity or lowest equipment costs) is represented by "++". This is followed by "+" and then "-". "---" is the most inferior rating.

[a] How well does the indicator measure the objectives?

[b] Results not dependent upon identity of the investigator.

[c] How fine grained is the indicator's metric?

[d] How quickly does the indicator register changes in the objectives being monitored?

[e] The level of training, skill or education necessary to implement the indicator.

[f] The total time necessary to implement the indicator.

[g] Regardless of the man hours, how many field visits are necessary to fully implement the indicator.

Outlook and recommendations

To its credit, this indicator is fast, easy and requires no special tools to execute. Given the nature of what is being investigated, it also makes little sense to use a control group with this indicator (cutting expenses even further). Although it is a valid indicator (at least for smaller activities)—it will be of little use vis à vis resource intensive activities. It is simply not realistic to think that local people will have the capacity to finance and execute large investments such as check dams or community wells. These are activities which, at least in the Indian context, can only be implemented by the state, or large NGOs.

Another weakness of this indicator is that it is not particularly reliable. It relies too much upon luck—evaluators must be told by some informed party that a case of replication exists, or the evaluators must stumble upon the replicas themselves. In addition, this is only an incomplete indicator of replicability. Replicability depends upon many factors, including geography, climate, level of development, socio-economic institutions and the structure of the state. Even if an activity is being replicated in the watershed being evaluated, this does not mean it is replicable elsewhere. An evaluation can only conclude that a activity is potentially replicable.

In the end, this indicator is recommended, but only for want of a better alternative. Other monitoring and evaluation specialists would be well served to either expand upon this indicator, to make it more reliable and applicable to a wider range of activities, or develop a new one.

Social capital

...social organisation and mobilisation often occur through what are called "demand groups" — loosely organised pressure groups that spring up around some contentious issue, then quickly disappear after the conflict has subsided

Target objectives

While the indicators Use and Outsiders produce information that is crucial to the determination of sustainability, they do not address the issue of social organisation and mobilisation. When programme investments are on common land or public land (which is generally the case), then they are, in practice, owned simultaneously by everyone in the watershed, and no one.[6] Commonly owned resources are difficult to manage and maintain. In the absence of some norms or institutions, there is no way to prevent over use, ensure maintenance, solve disputes, etc. Solutions to these problems must be found if a resource is to be used, sustainably.

Measurement procedures

Post-Independence India has not provided strong, decentralised political institutions to manage local soil and water issues. Appropriate government departments generally exist at the state level, but these are largely distant bureaucracies, not local democratic institutions. Under the IGBP's overall plan, partner state departments and NGOs were supposed to co-operate towards the goal of building up the sort of social capital that could locally manage watershed issues (Honore and Chaturvedi, 1997).

As defined here, "Social Capital" can be any organisation or institution that facilitates group co-operation towards a social goal.[7] Anything from an interest group or chamber of commerce to a village council or political party can act as a vehicle of social organisation and mobilisation. Even if the indicator Social Capital only incorporates such easily identifiable organisations, measuring its strength would be difficult. Should evaluators count the number of such organisations, the attendance at the organisations' meetings, or the number of meetings that they hold? Although this would be difficult, it could be done. But are these even relevant pieces of information? Just because organisations exist and hold meetings, it does not mean that they have any capacity for social mobilisation.

This problem is made even more complex by the nature of Indian politics, where social organisation and mobilisation often occur through what are called "demand

groups"– loosely organised pressure groups that spring up around some contentious issue, then quickly disappear after the conflict has subsided.[8] Because demand groups are usually in the dormant stage, it is not possible to measure their strength through a survey of existing social organisations.

With these ideas in mind, the evaluation team began to develop an indicator of social capital while working in Arki RWS. Instead of focusing on processes (i.e., groups, meetings, attendance), the team looked for outcomes. Investigations revolved around the question: Had local citizens recently confronted any watershed related problem? (i.e., had there been any instances where groups of local citizens attempted to solve some soil and/or water conservation problem?) This would be used as evidence that Social Capital exists.

Since many people did not really understand the issue of soil conservation, the team chose to focus on water issues, which are more concrete and of immediate importance to local people. Villagers were asked about problems they may have had with water. When the team uncovered some water problem, they asked how it was dealt with. The team was looking for verifiable stories of groups who had come together and successfully addressed some water problem. Appropriate cases were followed up and noted down in detail.

Outlook and Recommendations

This indicator is a valid measure of social capital — if people have demonstrated the ability to pursue grievances regarding watershed issues, then social capital is present.

Unfortunately, this is not a reliable indicator. The questions that the investigators need to ask are necessarily vague, so respondents may not always understand what the evaluation team is looking for. In addition, not everyone in a watershed may even be aware of the Social Capital that exists. If evaluators do not interview the right people, they will not obtain the necessary information.

In addition, it may be difficult to determine change with this indicator. If Social Capital was previously non-existent in a baseline survey and then it registers in a subsequent survey, evaluators are safe in assuming that change has occurred. Evaluators can also assume that change has occurred if identical types of social capital are found in the "before" and "after" surveys, but the intensity of the social capital has changed. If, however, one or more forms of social capital were found during a baseline survey, but completely different ones are found in a subsequent survey, evaluators will not be in any position to determine if the level of social capital has increased or decreased. It is very difficult to say whether one type of social capital is stronger than another.

While the indicator is inexpensive to use in terms of equipment (absolutely none is required), it must be executed by someone with a sophisticated understanding of social capital, local society, governmental structures and the programme being evaluated. In addition, this person must be a skilled interviewer, although local language capabilities are not essential for such interviews.

Finally, this indicator cannot be executed with any great speed. It involves a great deal of open-ended interviewing with various types of people—from farmers to local government officials. In addition, many of these interviews require subsequent discussions with additional informants. These interviews can, however, be carried out concurrently with other parts of the evaluation. Since it is a participatory indicator, it does not require the use of control groups.

This indicator is recommended with reservation. It is the best available alternative for measuring the existence of social organisation and mobilisation. It can be abandoned when a more reliable and less time consuming alternative is developed.

CONCLUSIONS

Indicators work! After testing and refining the Programme Evaluation Protocol under field conditions, the author who was the Principal Investigator in the present study does not hesitate to claim that this indicator set can be used to measure physical and socio-economic realities in rural watersheds, although not always as quickly, cheaply and easily as originally hoped. Exclusive of travel time, a team consisting of the PI and two assistants was able to execute the preliminary PEP in two Representative Watersheds—Arki in Himachal Pradesh and Kattery in Tamil Nadu—in just less than twenty days . This was done with very few expensive tools and with research assistants who had received very little specialised training.

Of the nine indicators in the set, four—Height-for-Age, Consumer Durables, Use, and Outsiders—are highly recommended. Two more—Soil Loss and Ground Water—are recommended with some reservations. This is because they do not meet all of the original selection criteria—that the indicators be "fast, cheap and easy to use". (For example, both Soil Loss and Ground Water are labour intensive and a final analysis can only be undertaken after years of data collection.) Finally, while they are usable in their present forms, it is recommended that Attendance, Replication and Social Capital could be further modified due to confront reliability problems. Table A contains summary information about the individual indicators.

As a final note, it is important to repeat that the point of this article has not been to claim that indicators are a monitoring and evaluation panacea. They are simply one of the many tools that are available. The strength of this method is its economy vis à vis time and monetary resources. For this reason, small programmes may prefer it. Most indicators do not, however, offer the sort of precision and detail that can be obtained through a full benefit-cost study. Nor are all indicators as easy to execute, as originally imagined. Just as with traditional, benefit-cost analyses, most of the indicators discussed here must be executed by college graduates or professional consultants.

The hope is that programmes and monitoring and evaluation specialists will now be able to benefit from the IGBP experiences. Those interested in carrying out a similar evaluation can follow the simple guidelines set out in the Programme Evaluation Protocol which is available with the IGBP.

END NOTES

1 For example, see The Economist, April 12, 1997 (75)

2 According to international organisations like the World Health Organization, this is a valid measure of nutritional status. Height-for-age as an indicator of health may face objections from people who do not believe that the heights of indigenous people can be compared to standardised tables constructed by the United Nations or the United States Department of Health. The basis of such an objection might be that local people are genetically shorter or taller than Western people. According to research financed by the Food and Agricultural Organization of the United Nations this is not true (Bender and Remancus November 13, 1997). At least for the first ten years of life (through puberty) children throughout the world are the same height, all else being equal.

3 The original plan was to calculate the children's ages using an ageing chart. Such a chart lists notable local events (whose exact dates are known) that occurred near the time that the children were born. A parent is then asked to specify how the birth of their child relates to these events. In this way the child's age can be estimated quite accurately. When the team was informed that the children in both watersheds had vaccination cards, this plan was abandoned. If birth records are not available, evaluators will need to devise a birth chart during the first several days of their stay in the watershed. This should be done with the help of local people and the NGO.

4 Height-for-age as an indicator of health may face objections from people who do not believe that the heights of indigenous people can be compared to standardized tables constructed by the United Nations or the United States Department of Health. The basis of such an objection might be that local people are genetically shorter or taller than western people. According to research financed by the Food and Agricultural Organization of the United Nations this is not true (Bender and Remancus November 13, 1997). At least for the first ten years of life (through puberty) children throughout the world are the same height, all else being equal.

5 After collecting some enrolment figures in both Arki and Kattery, the original reluctance to use them was justified. One school in Kattery had offical absentee rates (enrolment minus attendance) of between twenty and forty percent. That this many students are missing for just the day is unlikely. It is more probable that many of these "absentees" are simply on the enrolment roster, but they do not really attend school.

6 I am quite aware of the distinction between common and public property. In India, however, the state is often very removed from the management of public lands, in which case public and common lands are treated very similarly by local people.

7 Ideas about social capital have been drawn largely from Robert D. Putnam's Making Democracy Work (Princeton: Princeton University Press, 1993).

8 The term "demand group" was coined by Rudolph and Rudolph in their book In Pursuit of Lakshmi (Chicago: University of Chicago Press) 1987.

REFERENCES

Arya, Swarna Lata and J.S. Sharma. 1995. Socio-economic Implications and Participatory Appraisal of Integrated Watershed Management Project at Bunga. Chandigarh: Central Soil and Water Conservation Research and Training Institute Research Center.

BAIF Journal. 1997. "A Study of the Inter Relationship between Livelihood, Development and Health". BAIF Journal [Pune, India] 17:2 & 3 (13-15) 1997.

Barrow, C.J. 1997. Environmental and Social Impact Assessment: An Introduction. London: Arnold.

Batiwala, Srilatha. 1994. Women's Empowerment in South Asia: Concepts and Practices. New Delhi: Food and Agriculture Organization.

Beaton, B., A.Kelly, J.Devany, R. Martorell, and J. Mason (eds.). 1990. "Appropriate Users of Anthropometric Indices in Children: A Report based on an ACC/SCN Workshop." Geneva: UNIES.

Bender, Bill and Sandy Remancus. "Anthropometric Tutorial." Anthropometry Resource Center (Funded by FAO/SADC project GCP/RAF/284/NET).

http://www.od.com/anthro/tutorial/tutoc.html (downloaded on November 16, 1997).

Bollom, Michael W. 1998. "Impact Indicators: In Search of Fast, Cheap and Easy Indicators to Assess Physical, Social and Economic Change caused by Watershed Management Projects". New Delhi: Indo-German Bilateral Project "Watershed Management" (Technical Paper No: IGBP-WSM 86/98).

Casley, Dennis J. and Krishna Kumar. 1987. Project Monitoring and Evaluation in Agriculture. Baltimore: Johns Hopkins University Press.

1988. The Collection, Analysis and Use of Monitoring and Evaluation Data. Baltimore: Johns

Hopkins University Press.

Chambers, Robert. 1994. "Participatory Rural Appraisal (PRA): Analysis and Experience." World Development 22, 9: 1253-1268.

1983. Rural Development: Putting the Last First. New York: Longman Scientific and Technical.

FAO. 1992. Sociological Analysis in Agricultural Investment Project Design. Rome: Food and Agricultural Organization.

GTZ. 1996. "Process Monitoring (PrM): Work Documentation for Project Staff". Eschborn, Germany: GTZ.

GTZ and SDC. 1997. "International Workshop on Sustainable Land Management: Guidelines for the selection of Indicators and Monitoring Methods (Workshop Proceedings)." No place of publication.

Gautam, Madhav. 1996. "Development of a Tool for Impact Assessment on Beneficiaries in Food for Work Projects." Working Paper No. 2 in "Rural Community Infrastructure Works Programme. Kathmandu: GTZ – Food for work.

Honore, Guy. 1996. "Strategy and Approach" New Delhi: Indo-German Bilateral Project "Watershed Management".

Honore, Guy and K. Chaturvedi. 1997. "Joining Hands: Enhancing collaboration between Government Organizations, Non-Governmental Organizations and People." New Delhi: Indo-German Bilateral Project "Watershed Management".

International Institute of Rural Reconstruction and FAO Regional Office for Asia and Pacific. no date. (circa 1996). Resource Management for Upland Areas in Southeast Asia. No place of publication listed.

The IUCN International Assessment Team. 1997. An Approach to Assessing Progress Toward Sustainablity. Gland, Switzerland: ICUN.

Kerr, John M., Dinesh K. Marothia, Kata Singh, C. Ramasamy, and William R. Bentley (eds.). 1997. Natural Resource Economics: Theory and Application in India. New Delhi: Oxford University Press.

Kumar, Krishna (ed.). 1993. Rapid Appraisal Methods. Washington D.C.: World Bank.

1989. Impact Indicators: 'General Issues and Concerns' A.I.D. Evaluation Occational Paper No. 30. Washington D.C.: United States Agency for International Development.

Milllesen, Britha. 1995. Methods for Development Work and Research: A Guide for Practitioners. New Delhi: Sage Publications.

Narayan, Deepa. 1993. Participatory Evaluation: Tools for Managing Change in Water and Sanitation. World Bank Technical Paper No. 203. Washington D.C.: The World Bank.

Picciotto, Robert and Ray C. Rist (eds.). 1994. Evaluation and Development: Proceedings of the 1994 World Bank Conference. Washington D.C.: World Bank.

Putnam, Robert D. 1993. Making Democracy Work: Civic Traditions in Modern Italy. Princeton: Princeton University Press

Rudolph, Lloyd and Susanne. 1987. In Pursuit of Lakshmi: The Political Economy of the Indian State. Chicago: University of Chicago Press.

Schönhuth, Michael and Uwe Kievelitz. 1994. Participatory Learning Approaches: Rapid Rural Appraisal Participatory Appraisal, An Introductory Guide. Roßdorf (Germany): TZ-Verlagsgesellschaft.

Schmidt, U. 1993. Participatory Impact Monitoring. Eschborn, Germany: GTZ.

Smith, Thomas B. 1985. "Evaluating Development Policies and Programmes in the Third World" Public Administration and Development 5 (129-144).

Srinivasan K. And Abusaleh Shariff. 1997. India: Towards Population and Development Goals. Delhi: Oxford University Press for United Nations Population Fund.

Stephens, Alexandra. 1990. "Participatory Monitoring and Evaluation: Handbook for Training Field Workers." Bangkok: Craftsmen Press.

Thompson, John and Jules N. Pretty. 1996. "Sustainability indicators and soil conservation: A participatory impact study and self-evaluation of the Catchment Approach of the Ministry of Agriculture, Kenya." Journal of Soil and Water Conservation (July – August).

Traeger, Heidrun., 1997. Indicators of Sustainable Land Management: A Literature Review. Eschborn, Germany: Gesellschaft für Technische Zusammenarbeit (GTZ).

United Nations Department for Policy Coordination and Sustainable Development, "Indicators of Sustainable Development: Framework and Methodologies." gopher//gopher.un.org/00/esc/cn17/1996/backgrnd/INDICATO (Downloaded on November 15,1997).

United Nations Development Programme. 1997. Human Development Report, 1997. New York: Oxford University Press.

Valadez, Joseph and Michael Bamberger. 1994. Monitoring and Evaluating Social Programs in Developing Countries: A Handbook for Policymakers, Managers and Researchers. Washington D.C.: The World Bank.

World Bank. "Designing Project Monitoring and Evaluation" http://www.worldbank.org/html/oed/lp008.htm (Downloaded November 13, 1997).

1997. Expanding the Measure of Wealth: Indicators of Environmentally Sustainable Development. Environmentally Sustainable Development Studies and Monographs Series no. 17. Washington D.C.: The World Bank.

1996. Social Indicators of Development, 1996. Baltimore: The Johns Hopkins University Press.

1994. World Development Report: Infrastructure for Development. New York: Oxford University Press.