Indicators

Discussion Paper

With indicators being given high emphasis in many offices in the development community, this discussion paper includes a challenge to the use of indicators by UNV Programme Officer in Bhutan Chris Whitehouse, followed by a response in defence of the indicator, by UNDP Head of the Poverty Unit in Bhutan, Thomas Winderl.

The Ants and the Cockroach

A Challenge to the Use of Indicators

Chris Whitehouse

Emphasis on accountability and transparency and the introduction of logical frame approaches are being adopted by all the leading international aid and development institutions. The focus of this paper is on what some have claimed to be a key element of this approach – the indicator. It is argued that use of indicators can be time-consuming and expensive, can result in programme design being skewed away from the most effective and towards the most measurable, and that indicators are, most worryingly, an essentially flawed concept.

What are Indicators?

As more and more development organisations strive in an ever more competitive environment to procure funding support, they have to be able to show that they can use the funding efficiently and effectively, and that the funds used will produce results. Whereas in the past, donors may have been satisfied to know that their funds have been put 'in' to certain programmes areas in certain developing countries, today the situation is quite different. It is not enough to know that funds went in; donors want to know what came 'out' from the funding; the focus has shifted from inputs to outputs, and from outputs to outcomes.

Measuring inputs was easy enough - just count the dollars spent for this training programme, that equipment purchase, and so on. Measuring outputs is only marginally more difficult - it entails measuring the number of people trained, and the number of computers now sitting in district government offices. It is yet more challenging, and, it has been suggested, more useful, to measure outcomes, i.e. the extent to which, through their training programmes, or through the presence of computers, government offices are working more effectively, and the citizens are thereby benefiting from a more transparent and effective government support.

One key tool in measuring outputs and outcomes is the indicator. In effect, this is a measuring stick of development programmes. Indicators could be, for example, the percentage of district level civil servants who are computer literate to a certain level, the total number of computers that are able to access the internet, the level of satisfaction of rural people in the service they receive from government offices, or, at its broadest, the Human Development Index for the country.

Criteria for "Good" Indicators

To be effective, indicators have to satisfy certain key criteria. Are they measurable? Can we identify who will do the measuring? Will the measures be accurate? Do they truly measure the output/outcome which they claim to measure? Will the cost of the measuring be reasonable (relative to the cost of the inputs)? One must remember, of course, that it is not enough to take the measurements at the end of the programme; we need a baseline (pre-input) measure, against which any improvements can be compared. Take, for example, a computer training for civil servants. A number of indicators might come to mind.

- *Number of trainees trained* - easy to measure, an accurate figure is attainable - but fails to measure the quality or use of the training.
- *Average satisfaction score from an evaluation survey at the end of the programme* - this will help measure the quality and use of the training, but now the number of trainees will be lost, and it doesn't measure the long-term effect of the training. And is it accurate? Maybe the trainees want to be nice to the facilitators? Or they are filling in the form at the end of an alcoholic closing ceremony?

- *Average time taken for civil servants to carry out a certain process* (e.g. the issuing of a driving licence, or locating the tax details of a citizen). This is the first indicator which enables us to have a pre-training baseline figure, against which a post-training figure can be compared. It also relates closely, perhaps, to the rationale for having the training programme in the first place. However, we are now losing the direct link with the training programme, as any changes in the scores when compared to the baseline (whether the scores are better or worse) maybe due to external factors - e.g. the reliability of the national telephone service, or government staff motivation levels, or as a result of changes in salary structure.

The issue of indicators is a complex one, and over the next few paragraphs, attention will be drawn to the various challenges that the indicator approach now faces. Mention has been made of the difficulty of identifying truly effective and accurate indicators, but it is likely that, in the course of time, the selection and wording of indicators will improve. It is not, therefore, the aim of this paper to address this particular issue. Rather, it seeks to assess the extent to which indicatorism is a welcome and useful approach at all. At the end of this paper, conclusions will be offered on the value of this indicator-approach, and on how the challenges facing this approach can best be addressed.

The queen ant announces that, henceforth, it is not enough for each ant to just 'chip in' to the tasks at hand. Next time there's a cockroach to be moved, she is going to measure each ant's contribution

Illustration of Indicators at Work - The Ants and the Dead Cockroach

Imagine this scene: Thousands of ants working to bring a dead cockroach towards their nest. An amazing feat of teamwork, where these very small animals are pulling an object many times their own weight.

Now, choose any individual ant, and you will find that it may be doing any number of things; it may be a key player, pulling the dead cockroach in the right direction, towards the nest. It may unwittingly be one of a minority pulling it in the wrong direction. It may be running around doing nothing, but just feeling very excited to be near the action. It may, through its very presence, be encouraging the other ants in their work. Conversely, its comments might be irritating the other ants, and lowering their motivation. It may be helping to clear the route, flattening the path and thereby easing the movement of the cockroach… or it may, in its excitement, be making the path more difficult. So how do measure this ant's contribution to the task at hand? If we ask it, it will say it contributed actively, as it indeed shares the colony's dream of having the cockroach at the door of their nest. It can provide us with supporting information, for example, it can tell us for how many minutes it (itself) was pulling at the cockroach (but in which direction, it might not be able to tell us!). It can advise us of how many pieces of debris it moved (but off the path, or onto the path?). It can tell us how many suggestions it made (but did they help or hinder its colleagues?).

Or we can be more 'objective', and ask the other ants. But maybe there will be factors that will influence their reply - they might not like this particular ant; perhaps it had previously been alleged to have contributed to the collapse of the west wing of the ants nest. They may not have noticed this ant. They may be too busy with their activities, and give a quick but inaccurate reply.

What else can we measure? Well, we can go to a higher level of assessment. For example, we can measure the average speed of progress of the cockroach (but this will be affected by such factors as the weather, the gradient of the slope, the time of day, the nutritional levels of the ants, etc.) Or we can just assess whether the objective was met - did the cockroach arrive at the ants nest door?

It is apparent that the broad measures are more objective but more vulnerable to external factors. The more finely tuned measures (specific to the one ant) are less vulnerable to external factors, but are more vulnerable to subjectivity.

Now if, rather than looking at the role the ant played in the cockroach move, we now look at the contribution, for example, of UNDP to ICT development in Bhutan, we face the same problems.

- UNDP can quote with pride the increasing computer use in the country, a noble outcome - but one cannot deduce from this that UNDP actually contributed to it. The cockroach did move, yes, but did this ant actually help?
- Alternatively, UNDP can look at its outputs, and can give the numbers of people trained, and the number of computers now installed in the districts. But did this contribute to the improved computer use? Or was it irrelevant? Indeed, was the improvement in computer use despite (rather than because of) UNDP's interventions? The cockroach moved, yes, but might it have moved even quicker and easier if our friend the ant hadn't made 27 (unhelpful) suggestions, or hadn't moved 94 twigs along the path?

The Ants Story Continues - and Indicators are Introduced

Now, let's return to the ants. The queen ant announces that, henceforth, it is not enough for each

ant to just 'chip in' to the tasks at hand. Next time there's a cockroach to be moved, she is going to measure each ant's contribution, and their rights to gnaw at the flesh of the dead cockroach will be awarded in proportion to their contribution to the effort. Indeed, since the task will be so great to assess each ant's individual performances, she has identified 10% of the ants to be carrying out the monitoring on her behalf. Each ant will have to report to their monitor ant, as well as to be observed by the monitor ant during their work, and the monitor ants will collate the findings and report to her every quarter.

Of course, this initiative made a lot of sense. The queen had noticed that some of the ants were not pulling their weight, were free-riding the system. Indeed, she had begun to feel that those who ate the most were the very ones who did the least work; perhaps they were getting too fat and so got exhausted more easily. And wasn't this bad for morale? This new initiative was very timely, also, as the number of ants in her nest was rising, while at the same time the supply of dead animals for food was getting scarce. No, it was really high time to make change.

Never before had so many ants been so eager to be seen to be doing something useful.

That evening, there was considerable unease in the nest. They were all tired, of course, after the day's work, and the prospect of being watched all the time in the future was, to say the least, unsettling. Plus, as one of the more mathematically minded ants pointed out, they would now only have 90% of the workforce actually pulling the cockroach; the other 10% would just be patrolling up and down, taking notes. Then again, they would have less time to do the pulling, as they would, each one, have to report to their monitor ant on what they each had done. Still, the queen was probably right: the lazier ants would either have to start working or leave the colony, and through the supervisory scheme, they would all put in a bit more - so at the end of it all, there would be more food for fewer ants! Yes, she was right. By the time they went to sleep, all but the laziest ants were happier than they had ever remembered.

The Ants Story Continues - the Ants and the Dead Beetle

Next day was the beginning of a new era for the ants, and they woke early, looking forward to a new life. And when the report came through the nest's intercom system that a dead beetle lay 20 metres east of the nest, they rushed to work. Of course, knowing they had to report their activities to their monitors, they each made sure to greet their monitors, just to be sure they were recorded as having gone to work. They each carried with them a notepad, pencil and stopwatch, so that, whenever they were pulling, or clearing debris, or cheering their colleagues on, they could note down how many times they had done each, when and for how long. After all, it would be disastrous if they concentrated so much on their work, that they forgot to record it - they could find themselves ineligible to feast on the beetle at dinner-time!

The first delay came when the monitors decided to measure the beetle's weight - by finding out what the minimum number of ants needed to hold the beetle full off the ground. They then had to record the distance from the cockroach from the nest. The best way, they found, was to have all the ants stand head to tail in a long line, and the monitors counted how many ant-lengths lay from the beetle to the ants nest door - it came to a staggering 198,865,314! It was now mid-morning, and the ants were tired from both these activities. The whistle blew, and it was time for action! With a sense of exhausted relief, the ants rushed to move the beetle… but the scene was one of unprecedented chaos. Never before had so many ants been so eager to be seen to be doing something useful. The monitors held their front legs to their ears, such was the din of thousands and thousands of well-meaning but contradictory suggestions. And the beetle all but fell apart - so many ants pulling with all their might, in opposite directions. Such was the force that at one stage, over 600 ants were nearly crushed to death when the right wing of the beetle suddenly came off.

Eventually, the beetle, now in pieces, lay nearer to the nest door, again the distance was measured; and again the various parts of the beetle were weighed. Then the exhausted ants formed long queues to submit their reports to their monitors, who in turn had to collate the many thousands of figures, which were then submitted to the queen. She was of course delighted - it was really working. And she was pleased, too, to find that none of the ants had been found to be lazy.

The only slight disappointment was that the flesh of this beetle was found to taste not quite right, and as they went to sleep, many of the ants complained of feeling a little dizzy. Was it all the excitement and noise? Or was it just exhaustion? Or was there something wrong with the flesh of the beetle? In their enthusiasm that morning, none of the ants had noticed the strange coloured liquid which one of those big humans had pasted onto a leaf near where the beetle had died.

The moral of this story is that a balance has to be maintained between the time and effort spent doing what needs to be done, and that spent reporting what you have done. The queen was right to note that there were imperfections in the previous way of things; but she perhaps allowed (encouraged?) her

colony to shift their focus too significantly towards measuring indicators.

The Dangers of Indicatorism

The key problems with indicators fall into three main categories: the time and resources spent on identifying and measuring indicators; the skewing effects that inclusion of indicators may have on programme design and implementation; and, perhaps most seriously of all, the logical fallacy which many development workers fall victim to, of confusing concomitance with causality.

1. Time and resources spent on indicators

The story of the ants above clearly demonstrates what many of us already recognise in the real world, that additional requirements in terms of reporting results will entail costs in terms of time and money. Yet, at the same time, all projects and programmes aim to limit the amount of money 'lost' on administration, monitoring and evaluation to a bare minimum. A balance therefore has to be struck between a perfect monitoring and reporting system, and the desire to spend as much time and money as possible directly on project activities.

Therefore, even if indicators can be helpful to the development world (which is questioned further in parts two and three of this section), we need to ensure that (1) they measure what they are supposed to measure; and (2) that the cost of measuring is not out of proportion to the project budget.

As noted earlier in the introductory section, and also illustrated in the ants' story, there is a play off between indicators which are easy to measure, directly related to outputs, and which are de facto trivial (e.g. number of people trained), versus the more difficult to measure indicators of outcome, which are far more significant, yet at the same time are more vulnerable to external factors (e.g. improvements in government service to the public; or improved standards of living amongst the rural poor).

A review of project document log-frames shows a spectrum from the banal to the stratospheric:

a. one project might have the activity given as 'ICT training for 20 IT managers'; the output is '20 IT managers trained'; and the indicator is 'number of IT managers trained'. Means of verification, an important element in any logframe, would be 'attendance record'. Yes, they did the training, the trainees were trained, and we can know it was a success by counting the number of people trained - but was the training course well designed? Was it useful? These questions remain unanswered;

b. at the other extreme, the activity could be the same, and the output (or really, perhaps, the outcome) defined as '20 IT managers able and motivated to provide full IT service to all the government offices in their districts'. The indicator could then be something like 'number of breakdowns of computer systems which remain unrepaired 24 hours after reported breakdown' - now, this indicator would very clearly reflect the quality of the training - but maybe other factors will come into play as well? What will happen to these measures if one of the trainees has to attend a funeral on the other side of the country? What if the computers were using pirated or poorly installed software? What if a lightning storm caused an electricity surge across half the country? Yes, we have a fine indicator, but does it measure only what we want it to? The answer is 'no'.

I can anticipate the response - oh, you have chosen some particularly weak indicators; but I would say that such weaknesses tend to be the rule, rather than the exception. I look forward to the time when I can see a project document whose indicators are measurable, will truly measure what they purport to measure, and which are neither trivially meaningless nor largely prone to external factors beyond the control of the project.

Let's return to the ants again - what went wrong? Firstly, they invested heavily, to the detriment of efficiency, in measuring and monitoring. Indeed, 10% of the labour force were taken out, in order that they could carry out this monitoring. Normally, time and resources invested in indicators should be kept to a minimum, and must remain proportional to the size of the programme being measured (did the ants really need to weigh the beetle, and to measure its distance from the nest, and to each one keep record of their contributions to the team effort?). But another key error of the ants was that they focussed on the measurables, and ignored the most important (but least easy to quantify) aspect of their task. They measured weight, distance and total ant-hours spent making suggestions and pulling, but ignored the most important but least quantifiable element - did the beetle taste OK? In fact, the most appropriate indicator for the ant colony may have been something as simple as nutritional levels of ants; or, at the broader level, average life expectancy in the colony. But then, as discussed above, these measures, although more appropriate, would be vulnerable to external factors - in this instance, to beetle demographics, climate change, use of pesticides by humans, etc., over which the ants cannot possibly have any control.

2. The Skewing of Project Design and Implementation Behaviour

More worrying than the above, is the fear that the designers of future project and programme designs may be tempted to engineer their activities and outputs towards 'measurable' achievements. That is to say, focus on indicators can skew development programmes during their design stage. Current programmes may survive, as we implant backdated indicator baselines and target figures,

which become a useful annex to well-designed programmes. However, many of these good but otherwise 'wishy-washy' programmes may not have been approved had indicatorism been the fashion a decade earlier. We have seen a shift away from the small independent projects building a bridge here, giving literacy training to a women's group there - and rightly so. It is recognised that we need a more integrated approach, where a bridge, and literacy training and many other components should fit into a whole - i.e. into an integrated programme. The danger with indicatorism is that it will result in a de facto re-division of the integrated whole into only the measurable parts. The 'wishy-washiness' that binds an integrated programme together is the very element which will fall away when we focus only on the measurables.

A related worry is that project designers, especially if they know that they, the same people, may be responsible for monitoring the project during implementation, may feel tempted to under-estimate the targets to be reached. Should we aim at 15 out of 20 of the NGOs achieving growth? No, better aim lower - perhaps we should aim for 5 out of 20! And then, great success, the project is completed and 10 NGOs have seen growth! Wow!

Yet another worry is that the priorities of project management personnel during implementation are also likely to be swayed by this emphasis on indicators. Rather than invest time and effort in something which common sense would suggest was necessary or desirable, they may feel tempted to focus on those activities which will most speedily and easily achieve targets set - why risk allowing the lower class (and lower educated) people into the training programme? If we are aiming for 75% minimum to pass the final exam, then better to go for those trainees with a higher chance of passing. And certainly, don't let's get any women trainees, as they will be bound to miss some classes due to family commitments. Although gender issues were covered in the document text, there are no disaggregated targets for this indicator, so going for healthy educated men makes the most sense, doesn't it?

One is reminded of the absurd situations which arose in the Soviet Union. For example, one of the targets set for a shoe factory was 'number of shoes manufactured in a year'. Since it took time and effort to adjust the machinery after making left shoes, to making right shoes, and the target only specified total number of shoes (not pairs of shoes), then it just made sense, didn't it, to churn out only left shoes for the whole year? We can laugh at the Soviet Union now - but would we like to have people laughing at the development community in the same way, in years to come?

Focus on indicators, therefore, can have detrimental effects on project design, and on implementation - so, if you are going to have indicators, and are going to take them seriously, enormous care must be taken to ensure that the wording of the indicators is sufficiently tight that the project focus with respect to indicators is exactly matching with the project focus as a whole; and that, that project focus would remain valid even if indicators were not on the cards.

Can anyone advise on which indicators the queen should use? They would need to be indicators of factors over which the queen and her colony can have control, they should be measurable, AND (most awkwardly) should be sufficiently directly relevant to the needs of the colony that introduction of these indicators will not skew the colony's programme activities away from their prime needs.

3. Scientific Validity - Concomitance, Causality and Control

The final challenge to indicatorism is perhaps the most serious. The introduction of indicators at first sight appears to demonstrate a logical, scientifically valid system of auditing, monitoring, proving transparency and accountability. Yet this, as shall be explained below, is definitely <u>not</u> the case.

Firstly, in order to see the fatal flaws in the logic of indicators, we have to draw a very important distinction between <u>concomitance</u> and <u>causality</u>. When an event A is followed by event B, it is tempting to say that A caused B. When the sun sets, and darkness follows, we deduce, especially when the same thing happens time after time, that the setting of the sun causes the darkness that follows. The identification of the causal link is strengthened when logic is used to support the justification - the sun setting is known to be due to the revolving of the earth, such that the sun is now 'below us', then of course it will become dark.

However, one can also think of instances where A is followed by B, but we know (through our scientific understanding) that A doesn't cause B. For example, the chiming of Big Ben for 6 o'clock in London may be followed every evening by the departure of the 18:01 train to Bristol. Does the chiming cause the train to depart? No - and this can be proved if you sabotage Big Ben to stop it chiming one day, yet still you will see the train depart; or, you could sabotage the engine of the train, and find that even after the clock chimes, the train doesn't go.

Scientists are very aware of this distinction between concomitance and causality - before any medicine is approved for sale, causality has to be proved - i.e. it has to be shown, not only that those suffering a headache are cured when they take medicine YYY, but also that their headache is not cured if

they don't take medicine YYY. Indeed, modern experimental methods require that double blind tests are carried out - out of a group of 50 volunteers, 25 would be given medicine YYY, and the other 25 would be given an identical-looking but harmless and inactive placebo, where neither the patient nor even the person administering the treatment can know who is getting the real medicine and who is getting the placebo. It would only be through this kind of test that YYY could be demonstrated to work

In the development world, through indicators, we also hope to test the validity of treatment YYY (e.g. a training programme for civil servants) as a means to arrive at situation ZZZ (i.e. improved service for the public). But what do we do? We provide YYY, and then claim, with zero scientific basis, that situation ZZZ was as a result of the provision of YYY. We fail completely to have a control group - to be able to compare what actually happened to the target group with what would have happened if they hadn't received this programme.

Does this make sense? Let's use a stupid example to show what I mean: I go to a 5-year-old child, who wants to be taller. I say to his parents that I can help. First let's measure his height. Then, let me give him three carrots to eat, once a week, for three years. Then, at the end of three years, let's measure him again. If he's taller, then we know that carrots make you tall.

Or, I can go to the government, and offer a course in hydrology for all government hydrologists. Without having any control group (i.e. of hydrologists who are not receiving this training), then how can I show that, simply because they are offering improved hydrology service five years later (outcome indicator), that our hydrology course had any positive influence? Only if you offer a 'placebo' course (e.g. in Asian Cookery) to another group, and if the hydrology-trained people fare better than the cookery-trained ones five years later, can you prove that your course was successful.

It is not enough to show improved scores as against a baseline, because progress (whether of the child, getting taller; or of the hydrologist getting wiser) will often happen even without carrots, or without training programmes. We need to have a control group, outside of the support programme, against which to compare any improvements.

It has been a cause for concern for a long time, even before indicators became fashionable, that project reports of income generation programmes, for example, would highlight growth in income for their target group as a success (without seeing what negative impacts there may have been in neighbouring villages). But now we are taking a further, and more dangerous step, of claiming scientifically measurable progress - of giving numbers to 'prove' success. It is very tempting for us to think that if a certain training programme is followed by an increased efficiency in an office, that the training was a success; even more tempting if the project management shower us with pre-training and post-training measures of office efficiency. But, without a control group against which these figures can be compared, these figures are meaningless. It would be a cause for grave concern if those funding such programmes were so impressed by these 'figures' that they would prioritise further support for this programme, to the detriment of other programmes which might be just as valuable, but where the project managements are not flooding the donors with figures.

Conclusion

Three weaknesses in the indicator approach have been discussed, illustrated by a story about ants in a colony seeking to improve performance through introducing targets and indicators.

To summarise the three key challenges:

1. Properly defined indicators are time-consuming to create, and expensive to measure. This time and money would be well spent if yielding useful and valid results, but, given the next two problems, it appears not to be the case.
2. Emphasis on indicators can result in the designs of future programmes being engineered towards achieving measurable results, and thereby disqualifying otherwise valid programmes and projects which are targeting improvements in areas which are by their nature un-measurable. Furthermore, selection of activities for achievement of outputs and outcomes may be made to a greater extent on the basis of the way the indicators defined, and only to a lesser extent on the effectiveness and efficiency of achieving those outputs and outcomes;
3. Indicators are invalid unless, through comparison with a control group, it can be shown that any improvements identified are due solely and completely to the programme being assessed. The claim that this system of indicators has any logical validity is flawed, and the development community risks damage to its reputation if ever this pseudo-science were to be assessed in the public domain.

So, Do We Drop the Logframe Model?

However, let us not throw the baby out with the bathwater. The logical framework is a highly valuable tool for development workers: it is through this, and most notably through the flow chart which often precedes logical framework design, that we can truly understand the logic of a project or programme. By seeing how well the inputs relate to the

activities they are supporting, and how the activities really relate to the outputs they are designed to address, and how, in turn, those outputs lead logically towards the achievements of the outcomes and on to the objectives… all this is really useful. And there's no argument with the idea of having checks on each of those arrows of logic within this flowchart, i.e. checking what assumptions there may be in the project design (are we assuming the government will be supportive of X; are we assuming that the participants of X training programme will be released by their supervisors to attend the training, and so on).

But what is being challenged here is the validity and efficacy of bringing in a measuring stick approach, a superficially scientifically valid measure, where no such validity exists, of investing significant time and resources in measuring something which proves nothing; of introducing a component to the logical framework which adds nothing to the logic - but, rather, can lead to project designers feeling obliged to skew their programming away from what is needed or wanted, towards something which (only apparently) is measurable.

Let's not get carried away by indicators - let's limit ourselves, in the logical framework approach, only to the central pillar of that approach: to logic and common sense.

Response to Chris Whitehouse

Thomas Winderl

A pot of chicken soup

Imagine you feel suddenly hungry. While the hydrologists in Chris Whitehouse's analogy were well-versed in Asian cuisine, you would like to stick to something less extravagant. You want to re-heat the succulent chicken soup (no, not for the soul) from last night. While the soup is on the gas stove, you want to monitor progress of the re-heating process (agreed, nobody calls it that way). After all, you hate cold soup, and you don't want it to boil over and burn, nor to gulp down a lukewarm soup. So what do you do? Right, you choose an appropriate indicator. There are many to choose from. You might want – for example - to visually monitor the soup surface in the pot, and remove the pot once the soup starts to boil. You could use your extended experience in soup re-heating, and use a time frame (e.g. 3 minutes 30 seconds for a ½ liter pot) as an indication that the soup is ready to eat. If you are a connoisseur - and a rather adventurous one as well -, you might stick your finger into the pot every minute or so to feel progress. We could easily think of more possible indicators.

Let's take another day-to-day example, this time not gastronomic but rather meteorological. You intend to take a stroll to the local supermarket, and want to know what clothes to wear, and if you need to take an umbrella (doesn't apply to Brits). What do you? You would probably look outside the window, see the sun, see the trees moving in the wind, see people wearing T-shirts in the streets, and conclude naturally that it is a sunny, but windy summer day. There is (normally) no need to measure the exact temperature, or the speed of the wind. These casual indications tell us what we want to know: Not to wear the fur winter coat, not to take an umbrella (except for the Brits), but to dress lightly, and maybe to take a windbreaker or a light jacket against the wind. Even if you normally would not say "Darling, I'll go and monitor the indicators for the weather", this is exactly what we do every day.

Indeed, talk about indicators and result-orientation is a part of contemporary development speech. No self-respecting expert in human development would go without it. They seem to be – as so many trends before - a temporary development fashion, like "participatory processes", "structural adjustment", "grass-root based work", or "microfinance".

While result-orientation and indicators emerged in the 1980s, I argue below that they are here to stay, and with good reason. Much more, indicators – and monitoring progress in general – **come natural to human beings**. We have been using them, are using them, and will be using them all the time. If you want, they have been around since the cave men, and might even have been a significant part of what makes the *homo sapiens* a highly successful species. True, in the caves – nor nowadays for that matter – we didn't call them indicators, but rather common sense. A lot of work needs to be done still to de-mystify indicators, and look at them as a standard tool for planning and monitoring progress in any development too difficult to appraise immediately.

Argument 1:

***More* time and resources**

While reheating yesterday's soup from the fridge, few people would design an elaborate monitoring process with complicated indicators to monitor the heating process. Nobody with a sound mind would call for high-tech instruments, or extensive surveys among soup particles. You would not call in a team of soup-cooking specialist to help you re-heat your soup. You can do it all by yourself. It is relatively easy and not very time consuming.

Chris Whitehouse is arguing that time and resources invested in indicators must remain proportional to the size of the program. I couldn't agree more. However, the reality is that there is *not* enough time and funds invested in designing and monitoring clear and relevant indicators. As we know, decades after the "invention' of logical frameworks and indicators, the inadequate and skewed logical frameworks and bad or plainly wrong indicators still dominate the development business. And here I utterly agree with the author that bad indicators are worse then no indicators at all.

Let us look at the resources spent on thinking up and monitoring indicators: As a rule of the thumb, guidelines for monitoring recommend to earmark about 2-3% of the project budget to be used for this purpose. In development reality, however, it is highly doubtful whether many projects and programs spend that even that much time. In a 3 million USD project, that would mean 90.000 USD dedicated to creating meaningful indicators and monitoring them, which is hardly the case in the contemporary development business e.g. in UNDP.

The same holds true for wasted time. It is true that the selection of a good indicator requires a collective – and time consuming - thinking effort. This is not easy. While creatively thinking up good indicators is the time-consuming part, most methods to monitor change using an established indicator are easy, quick and cheap (and they usually don't involve dangerous things like sticking your finger in

boiling soup). But if a whole month is spent by the project designers to come up with quality indicators and monitor them in a 5-year project, this would amount to only 1.7% of the total time involved in the project implementation. And it is next to impossible to find programming exercises where even one week is invested in thinking up indicators. It just does not happen.

However, Chris Whitehouse is – indirectly - pointing out an important reality. It is not the case that either too much time or too many resources are invested, but that the result is generally of very poor quality. A common mistake is the definition of too many indicators. What do you do if half of them point at a strong improvement, and the other half at a deterioration of the situation? Nothing is won, and we would be better off without them, using our common sense and informed assumptions. Other indicators indicate the wrong thing. They were not thought through, and lack creativity. In these cases – which are without doubt the majority in most development organizations – the monitoring of these indicators brings no additional benefit. Time and resources are lost, which could have been spent more effectively on the project itself.

Chris Whitehouse is absolutely right when he argues that "a balance has to be maintained between the time and effort spent doing what needs to be done, and that spent reporting". However, we should interpret his argument as a case for *more* resources and time spent on careful planning and monitoring, rather then less. Given the choice between putting your money into an activity where impact is proven through a (mostly lengthy) narrative using a host of (hopefully not wild, but informed) guesswork, and well thought-through outputs, outcomes and impacts with indicators, a 3-4% share of time and resources is well spent indeed.

Argument 2:

Skewed implementation behaviour

Chris Whitehouse's second set of arguments – that the focus on indicators can skew development programs during the design state - is only partially valid. Let me point out some of the author's misconceptions:

Firstly, the paper's worry about a **re-division of integrated projects** into its measurable parts is lacking any base. If indicators are taken seriously, they do not only include input, output and outcome indicators, but also long-term goals. And only extremely poor logical frameworks allow for multiple goals. The norm is to have a clearly defined goal at the top level of a logical framework. As we know, if outcomes are not contributing to the overall goal of a project, they should be deleted from the framework. Rather then dividing integrated projects, the logical frame work (not so much the indicators) force project designers – if applied properly and rigorously – to focus on a coherent set of outputs and outcomes to achieve one goal.

Second, the author seems to lament that the "**wishy-washiness**" of old-style project design will fall away when focusing only on indicators. While it is true that old projects without indicators might sometimes have resulted in valuable development improvements, they did so *despite* – and not because of – the lack of indicators. The approach of "let's do something, and something good will probably come out of it" is not an option any more. This romanticizing image of positive spill-off of random project activities is clearly – and rightly so – a matter of the past, mostly due to the overall poor results of development aid and the increased accountability of donor countries to their clients, the tax-payers.

An extremely valuable aspect of including indicators in project design is the **process** itself. It forces the designers to define better what outcome is intended. While it is easy to set as a goal the increased capacity of the Royal Civil Service Commission, it is much harder when thinking up indicators what is really meant by it. What is the RCSC actually doing? What aspect of what the RCSC is doing do we want to enhance? Are we happy with 5 people working more effectively and efficiently, or are we targeting 50 people in the RCSC? Often, the post-design introduction of indicators poses.

Third, the paper argues that there is a strong incentive – if the designers and the implementation agents are the same – to **under-estimate the target**. This is indeed true, if one does not link the framework with the inputs, that is the funds used to achieve the targets. Using Whitehouse's example, it might look acceptable to spend 75.000 USD to help 15 NGOs achieving growth (5000 per NGO). However, if the designers set their target at 5 NGOs, that is 15.000 USD per NGO, the donor organization should decide that this is too much, and refuse the funding of this activity. Agreed, the linkages of inputs and outcomes in logical frameworks are in development practice still weak, but this does not indicate a flaw in the concept. On the contrary, the tendency is more and more to ask: How much money do we need to spend to achieve outcome X? Could we do it in a more effective way? What other options do we have at our disposal?

And fourth, the author describes a valid problem, the **focus on indicators** in project implementation. Citing the convincing example of Soviet shoe factory churning out right shoes in order to achieve the target set, he provides us with a hilarious description of the importance of getting your indicators rights. Assuming the intended outcome of the Soviet policy was the provision of sufficient, cheap

and high-quality pairs of shoes to its citizens, the chosen indicator was evidently flawed on the output level. It requires careful and creative thinking – and anticipating the fallacies of "indicatorism" for the implementation phase – to create indicators which capture the intended output, outcome or goal. Rather then being an argument against indicators, it is a good example of the peril of brainless indicators.

Argument 3:

Picasso, not Einstein

Chris Whitehouse's third – and potentially most devastating – argument is that indicators aspire to create a logical, scientifically valid system of monitoring and auditing, providing transparency and accountability. Once again, it is not his line of argument but his underlying assumptions which are flawed. While indicators – together with their twin, the logical framework - aspire to introduce a more structural and logical thinking into the complex realities of projects, it is a misconception that indicators pretend to set up a scientifically valid system. His argument is based on the popular myth that indicators – maybe because they operate with numbers – are science. They are not.

Let me explain more in detail two of the fallacies in dealing with indicators: First, they are not – and do not claim – to be scientific. Second, they do not normally measure progress or regress.

We cannot claim for most indicators that they are anyway close to **scientific** measurements. The creation of indicators is by no means a scientific action. Although most indicator-people will agree whether a particular indicator is better or worse, there is no systematic way how to decide among indicators which are equally good or bad. In short: far from being Science, the development of indicators is art combine with a large portion of systematic, logical think, and an even larger portion of common sense. If you look at the process of how indicators are being thought of, you will see elements: On the one hand, a certain number of tools are necessary to help in the process: problem trees, log frames, etc. But on the other hand, the process demands a high degree of creativity, out-of-the-box thinking, or de Bono's "lateral" thinking. The choice of the "right" indicators is an art rather than a science

The second misconception is related to the first one: indicators do not **measure** progress. One doesn't have to look too closely at the word itself to find out that indicators – well, as the word tells us – "indicate" a direction. They tell us which direction a change possibly takes, or whether there is hardly any change. If the average time the ant tribe takes to bring in a juicy beetle over the year is half the time it took them last year, it indicates to us – beyond reasonable doubt – that they are doing much better then before.

Indeed, the better an indicator, the better it matches the actual direction of a change. A good indicator will represent the actual development as closely as possible. However, this is not always the case. Let's look at another example: Ms. Poor wants to be rich in 10 years. Having defined 5 million Dollars as her personal threshold of wealth, she can measure her assets easily by looking monthly at her bank account statement. By monitoring this simple indicator (although no one would call it that way, it's just common sense), she knows with certainty where she stands in regard to her ultimate goal. This is one of the few cases where the indicator "USD in bank account" comes close to measuring her wealth.

There are very few indicators – however – which manage to come so close to indicate a change. Ms. Poor could have taken another indicator: the value of the cars she is driving. We can assume (yes, that's what indicators do!) that the richer she grows, the more expensive (or bigger? that would be another indicator) her car will be. But it's not as good as the bank account. It could happen that – because she saves all her money – she initially sells her expensive Porsche, and keeps on buying second-hand vehicles. She could look at the dividends her money creates. But stocks go up and down, and – although the general direction might tell us a bit about her financial status quo – we might get a distorted impression as well.

Here is a last example how indicators indicate and not measure. UNDP supports the Royal Government in its negotiations to join the World Trade Organizations. Finding an indicator is a bit tricky here. A possible solution could be the number of years the WTO grants Bhutan before it has to phase out subsidies for fertilizers. What a wacky indicator, you might think. But think again: Subsidies for fertilizers are generally a major point in accession discussions. One could readily assume that the better the Government negotiates with the WTO, the longer the transition period for fertilizer subsidies might be. Together with a few more indicators, this might well indicate to the donor agency how well spent their money was spent. The point here is: The length of the transitional period for fertilizer subsidies clearly does not measure at all the skills involved in the negotiations. It measures, well, the length of the transitional period. But still the indicator would be – among others – a clue how well the Government negotiated.

* * * * * * * * * *