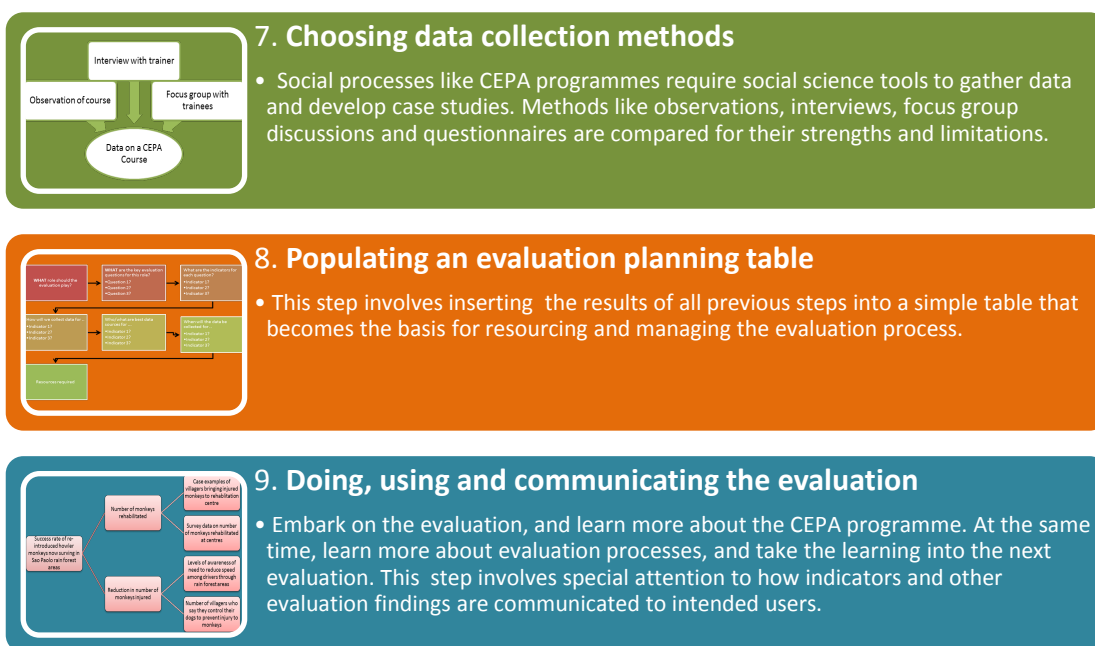




This section of the toolkit is a guide through the steps of designing an evaluation. It contains practical design tools and examples of evaluation questions and indicators.

At times it refers to the context in the Case Studies, to points in the Key Ideas Pages, and to various reference documents on the CD.



STEP 1: CHOOSING AN APPROACH TO THE EVALUATION

We find it important to decide early on what broad role evaluation will play in the CEPA programme. Will it be formative, summative, or developmental? ¹

Typically, evaluation plays a *summative* role in a process that is linear and limited to the three discrete steps in Figure 1.

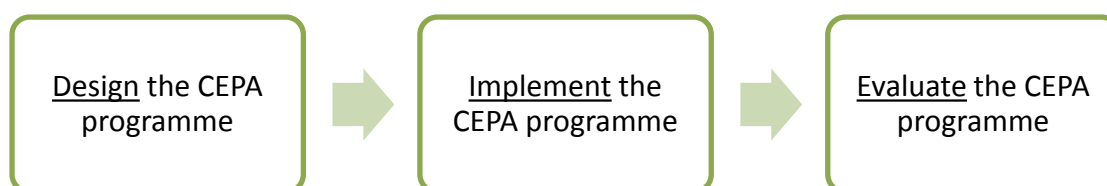


Figure 1: Linear Model of Programme Evaluation

A second phase of the programme could continue after the evaluation, in which case the evaluation could be said to be *formative*, if it shapes the second phase. If summative and formative evaluations are approached in this linear fashion, CEPA practitioners often find the provisioning for evaluation inadequate, for example, we might not have collected the necessary data for the evaluation from the start of the programme, or we might not have adequate opportunities to stop and reflect as the programme rolls out. This is a common situation, and the Cape Town Green Schools Audit Programme (in the *Case Studies Folder* on the CD) provides one example of where this may have been the case.

¹ Refer to the *Key Ideas Page: Approaches to Evaluation*. There we explain these broad roles that evaluation can play in a CEPA programme.

Another approach to evaluation is *developmental*. In this case, evaluation is built into all phases of a programme and is planned for at the same time as the CEPA programme is being designed. Evaluation data is collected throughout and there are regular programme pauses to review and reflect. This approach could be reflected as follows:

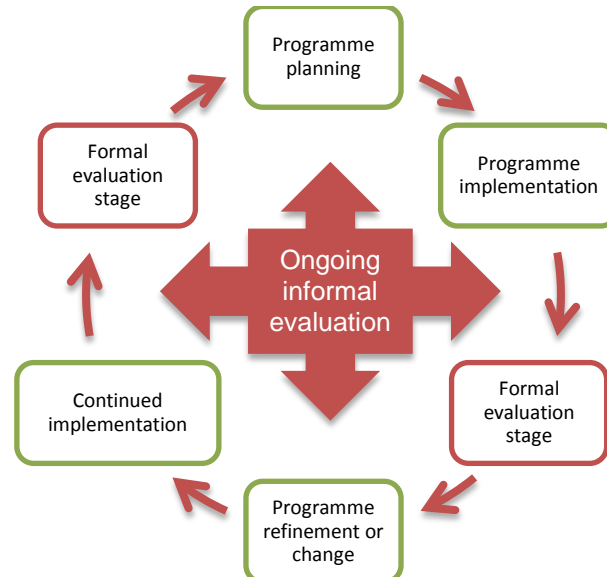


Figure 2: Developmental Evaluation

We find that complex systems² such as those in which we conduct CEPA programmes are not easy to map out and influence in a predetermined way. Developmental evaluations are therefore useful because they allow for practice-based learning, as we evaluate both our models of change and our programmes regularly in a process of continual action and reflection. They provide us with short feedback loops of information, which allow us to adapt and evolve our programmes and respond intelligently to the complexity of the situations in which our CEPA programmes play out.

However, developmental evaluations do require a different way of thinking about how we work, and adequate planning. All CEPA programme participants as well as managers must understand the role of evaluation in order to regularly contribute evaluation data; time and resources must be set aside for internal staff and, from time to time, for external evaluators; and CEPA staff must build evaluation into their daily routines. It is therefore not always possible to take a developmental approach. The chosen approach to evaluation will depend on:

- the phase of the CEPA programme
- the available resources
- the interests of the various stakeholders in the evaluation,
- the models of CEPA processes and change and what they should achieve, and
- the research paradigm informing the process.

² Refer to *Key Ideas Pages – Understanding Complex Systems*.

In addition to choosing a formative, summative or developmental approach to the evaluation, evaluation teams should decide on their research approach. Will the evaluation follow an experimental design, an interpretivist case study approach, a participatory action research design, or a combination? These (research methodology) choices have a great influence on how we go about the evaluation, and the kinds of questions that we ask. For example, a pre-test post-test experimental design requires one to set up baseline pre-tests and control groups beforehand. For a participatory action research-based evaluation, one needs to involve a broader than usual range of evaluation participants right at the start, when the research evaluation questions are formulated.

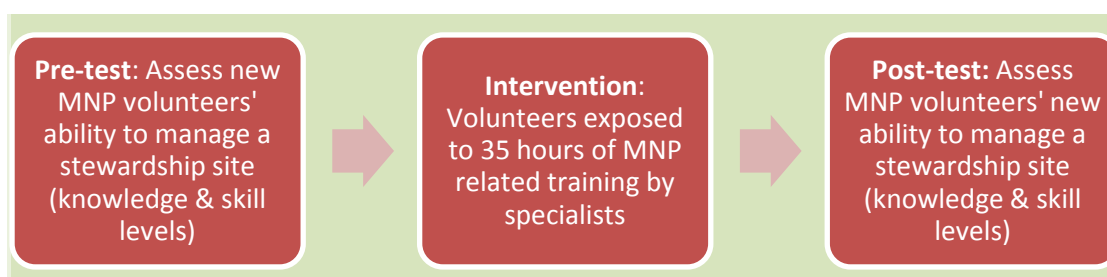


Figure 3: Illustration of a *potential* pre- /post-intervention evaluation component in the Edmonton Master Naturalists (MNP) Programme³:

Table 1: Choosing the Evaluation Approach

What role should evaluation play in your CEPA programme?		
	The benefits of this would be ...	The practical design implications are ...
A summative role		
A formative role		
A developmental role		
What research design is most appropriate for this evaluation?		
Experimental Design		
Case Study Based		
Participatory		
Other		
Combination		

³ Refer to the *Case Study Folder* on the CD, for a description of the Master Naturalists Programme.

Complete this table, then map the evaluation process as you see it unfolding, perhaps using one of the diagrams in this section (e.g., a linear evaluation process, a pre-test post-test design, or a developmental evaluation process). Also return to this table later, however, as you may want to change your approach once you have worked through the next steps.

STEP 2: PLOT THE LOGICAL FRAMEWORK FOR THE CEPA PROGRAMME

In this step we outline how one can plot the theory or logic of how a CEPA programme is meant to work in the form of a logical framework (usually abbreviated to ‘log-frame’). It is very useful to do this at the start of planning a new CEPA programme with built-in evaluation. We have also used it as starting point for designing an evaluation at any stage of a CEPA programme that is already underway. Once the logical framework has been plotted, we can identify single-loop evaluation questions related to the various elements of the log-frame. The process of drawing up the log-frame is very valuable in itself if done with the CEPA programme staff.

How do we know and demonstrate that a CEPA programme contributed to the change we intended? A well-crafted programme logic offers a basis for evaluating progress against intended outcomes and impacts. One of the reasons for this is that CEPA programmes operate in complex environments where the scientific certainty of proof is seldom attainable. Unlike in a laboratory, influences and forces in real-world contexts and communities are mostly beyond CEPA practitioners’ control. Therefore, evaluation is generally more about documenting a programme’s *contribution* to change, than about *proving* causal links or attribution.

This is where the programme’s logical framework is helpful. Using the basic ‘inventory’ template for a log-frame, and working backwards (from impact to resources moving left across the columns), identify and list the key components of the CEPA programme’s logic, as a basis not only for planning and implementing the programme, but also for designing evaluation questions.

The best time to build evaluation into a CEPA programme plan is in the initial programme planning stages. One of the many advantages is that one then knows what sort of data to collect and one can plan for it accordingly.

However, if a programme is already up and running when the need for an evaluation plan is identified, the logical framework for the programme can be plotted at any stage. Programme stakeholders often want to modify their log-frame after the results of an evaluation phase become evident.

There are various versions of a log-frame template. One example is used here (Table 2), and an alternative is included on the CD (Appendix 6). Most tabular templates use rows to order and show the relationships among components. Some number the lists within a column to aid discussion. Others have a box and arrow format to illustrate ‘causal linkages’, i.e. demonstrating how resources, activities, outputs, outcomes, and impact connect to form chains.

The first important task is to get the component parts categorised and described in a simple inventory (such as Table 2). Then, once the basic inventory table has been filled in,

experiment with identifying the relationships among the items across columns. For example:



We find it very useful to complete these tasks in a group of stakeholders involved in the CEPA programme. The process often results in enlightening discussions if stakeholders or team members have differing understandings of the programme elements and what they are meant to achieve.

Fill in a Basic Programme Logic Template

Fill in Table 2, or another logical framework format of your choice. We work backwards, starting by identifying the intended results (outcomes, incomes and outputs) before listing activities. For ideas, see the notes following the table.

Table 2: A Basic Logical Framework Development Template⁴

Resources	Activities	Outputs	Short- & Long-term outcomes	Impact
<i>In order to accomplish our CEPA activities we have and/or will need the following:</i>	<i>In order to achieve this we will conduct the following CEPA activities:</i>	<i>We expect that once completed or under way the CEPA activities will produce the following evidence of learning and action for local biodiversity:</i>	<i>We expect that if completed or on-going, this programme will lead to the following changes in 1-3 then 4-6 years:</i>	<i>We expect that if completed this programme of CEPA activities will lead to the following changes in 7-10 years:</i>
Materials and other resources required for the CEPA activities.	What is being done or will be done to create the desired change.	The most immediate intended results of the CEPA programme. <u>Each relates directly to an activity.</u>	Actual benefits or changes.	The longer-term change that stakeholders hope the CEPA programme will help to bring about.
Evaluation Questions ... Table 3				
Indicators ... Table 5				

⁴ Adopted from W.K. Kellogg Foundation, *Logic Model Development Guide*, January 2004, www.wkkf.org.

Below are some commonly used guidelines for completing the logical framework:

- **Impact** refers to the results expected 7-10 years after a CEPA programme is under way – the future environmental change we hope our CEPA programme will bring about. Impacts are the kinds of organisational, community, or system level changes expected to result from programme activities; they might include improved biodiversity conditions, increased human well-being, ecological resilience or social capacity.
- **Long-term outcomes** are results one would expect to achieve in 4-6 years. Like short-term outcomes (see below) long-term outcomes are also specific changes in attitudes, behaviours, knowledge, skills, biodiversity status or level of functioning, expected to result from programme activities. The difference is that they usually build on the progress expected by the short-term outcomes.
- **Short-term outcomes** are results one would expect to achieve 1-3 years after a CEPA programme is under way. Short-term outcomes are specific changes in attitudes, behaviours, knowledge, skills, biodiversity status, or level of functioning expected to result from programme activities.
- **Outputs** are the direct results of programme activities. They are usually described in terms of size and scope of the products and services delivered or produced by the CEPA programme. They indicate whether or not a programme was delivered to the intended audiences at the intended 'dose', scope or intensity. A programme output, for example, might include the number of classes taught, meetings held, materials distributed, or programme participation rates.
- **Activities and Resources** - The planning meetings, brochures, booklets, training workshops, and so on, that the CEPA programme needs, in order to achieve the intended results. To connect actions to results, this exercise links one's knowledge of what works, with specific descriptions of what the programme will do. In the planning stages, CEPA staff can consult CEPA specialists or refer to published guidelines for CEPA⁵, for expert-derived suggestions for CEPA activities. When listing the resources that are needed to support what the CEPA programme proposed, it may also be helpful to describe the influential factors in the context that CEPA staff would be counting on to support their efforts.

Create Evaluation Questions

We find that once we have created a logic model of the CEPA programme, it is not that difficult to develop evaluation questions. A logic model illustrates the purpose and content of the programme and therefore suggests meaningful evaluation questions. Table 3 gives some examples. As you work through it you may realise that a myriad of questions can be generated. *Deciding which questions to ask is a very important component of evaluation design, and is ideally an iterative process of consultation with stakeholders.*

In the evaluation design framework outlined in this toolkit, there are two broad sets of questions that can be derived from a programme log-frame:

⁵ *Guidelines for Environmental Education* and *Guidelines for Biodiversity Communication* are included on the CD (Appendices 2 and 3 respectively).

- Evaluation questions about the *single elements* in the log-frame (e.g., has an activity been completed, what was the quality of the output?). These types of questions are illustrated in Table 3.
- Questions about *relationships between the elements* of a programme's logic model (e.g., to what extent does a particular output result in a desired outcome?). Such 'double loop' questions serve to question the assumptions within the logic model itself, which means that one's evaluation (and programme) does not become entirely constrained by the logic model with which one started. This is discussed in step 3.

Add evaluation questions related to the elements in the programme logical framework, that you have drawn up earlier. See Table 3 for an illustration.

Table 3: Creating Evaluation Questions Using Logical Framework Components⁶

INPUTS	OUTPUTS		OUTCOMES		IMPACT
	Activities	Process	Short-term	Long-term	
Staff; Money; Training materials	Development of CEPA course	Targeted participants attended	Participants increased knowledge of biodiversity stewardship	Participants join or form communities of practice	Biodiversity loss reduced; ecosystem services increased
	Provide x interactive training sessions	Targeted content covered to a standard	Participants undertake stewardship activities	Biodiversity is effectively co-managed by City and citizens	
Key Evaluation Questions					
Was the provisioning of funding and staff sufficient, timely? Were the training materials of suitable quality, content?	Was the required CEPA course developed? Were all x sessions delivered?	Did all intended participants attend? All sessions? Why? Why not? Do the CEPA programmes communicate the issues comprehensively and effectively? Were participants satisfied with the course delivery?	To what extent did knowledge increase? What are participants able to understand and do as a result of an input/activity? How many participants signed up for volunteer stewardship/conservation action?	After 12 months, how many participants are still doing stewardship? How many groups have been formed? What is the scope and quality of their stewardship activities? How many hectares covered?	What is the status of biodiversity and eco-system services in the city compared to before the programme started? Have goals been reached? What unintended impacts have there been?
Indicators (Table 5)					

⁶ W.K. Kellogg Foundation, *Logic Model Development Guide*, January 2004, www.wkkf.org

The Benefits of a Logical Framework in Designing an Evaluation

- Can provide the framework for an evaluation plan, as it helps us to select and communicate evaluation questions and associated indicators.
- Provides a basis for discussion about the programme and the evaluation, and what we want to achieve, among stakeholders, CEPA practitioners, managers and experts.
- Helps determine and explain the relationship between an indicator and its purpose, in assessing the suitability of potential indicators to answer the key question(s) and their validity, and how effectively they represent the intended change.
- Increases the evaluation's effectiveness by focusing on questions that have real value for stakeholders.
- Helps to clarify the subject being addressed for all involved and aids in the selection and communication of appropriate indicators.
- Can guide on how to structure the explanation of an issue and the meaning of the indicators; it can be included in a report, where it may help to develop the narrative.

Finally, to the extent that the logical framework communicates the CEPA programme's logic of change, or the programme theory, it opens up the programme logic to questioning and revision, in those instances where the logic may be faulty and therefore hampering progress in achieving the intended outcomes and impacts.

STEP 3: PROBE THE ASSUMPTIONS UNDERPINNING THE PROGRAMME

It should be clear from Step 2 that plotting a logical framework for a CEPA programme is a very useful process. Among other things it reflects our assumptions about how change is likely to happen in a programme⁷. Log-frames are commonly used in development planning and may also shape our understanding of how change happens.

But what if our assumptions about what we need to do and what outcomes we will achieve, are wrong? Should we not also probe these very assumptions?

Like a pane of glass framing and subtly distorting our vision, cognitive or mental models influence what we see. These maps consist of personal and collective beliefs that are based on conclusions that we have drawn based on what we observe, our past experience, and education. We need these mental 'maps' to help us navigate through the complex environments of our world. However, all of our mental maps are flawed in some way, to a greater or lesser extent. This is only a problem if our self-generating beliefs remain untested.

Using the Ladder of Inference⁸

The ladder of inference (Figure 4) can help us to gain greater clarity on a CEPA programme we aim to evaluate, by:

- Becoming aware of our own thinking about CEPA process and change through reflection
- Making our thinking and reasoning more visible to others
- Learning more about others' thinking, through reasoning.

⁷ See *Approaches to CEPA and Change*, in the *Key Ideas Pages Folder* on the CD.

⁸ Senge, Peter, 1990. *The Fifth Discipline: The Art and Practice of the Learning Organisation*. Doubleday.

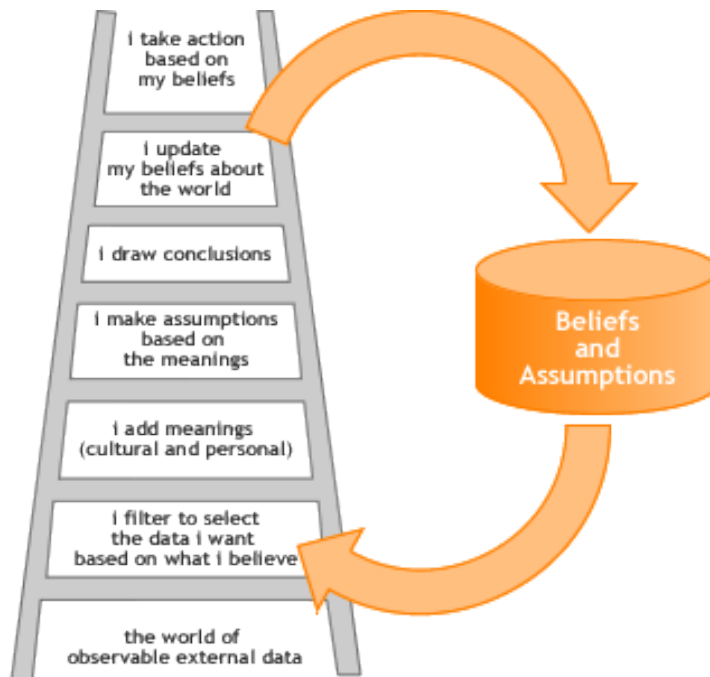


Figure 4: The Ladder of Inference⁹

To explore the possibilities of this, start at the bottom of the ladder, in the empirical world of reality and facts. From there (moving up the ladder), consider that we:

- Experience reality and facts selectively, based on our beliefs and prior experience.
- Interpret what this reality and these facts mean.
- Apply our existing assumptions, often without questioning or even noticing them.
- Draw conclusions based on the interpreted facts and our assumptions.
- Develop beliefs based on these conclusions.
- Take actions that seem 'right' because they are based on what we believe.

Without examination, this process can create a vicious circle. Our beliefs have a big effect on how we select from reality, and can lead us to ignore evidence, facts and possibilities. We could be 'jumping' to conclusions – by missing facts and skipping steps in reasoning.

Use the Ladder of Inference to encourage all evaluation participants to start with the facts and use their beliefs and experiences to positive effect, rather than allowing them to narrow or cloud their field of judgment.

We find it useful to consider the Ladder of Inference once we have developed the model of change for the CEPA programme we are evaluating, but also right throughout the evaluation. It encourages us to ask probing questions such as:

- Is this the 'right' conclusion? Why did we draw that conclusion? Is it sound?
- Are there alternative conclusions that are better supported by the facts?
- Why do we think this is the right thing to do?

⁹ ibid.

- What data have we chosen to use and why? Have we selected data rigorously?
- Are there any facts/ best practice research that we have left out? How would including them, change the conclusions?
- What are we assuming, and why? Are our assumptions valid?

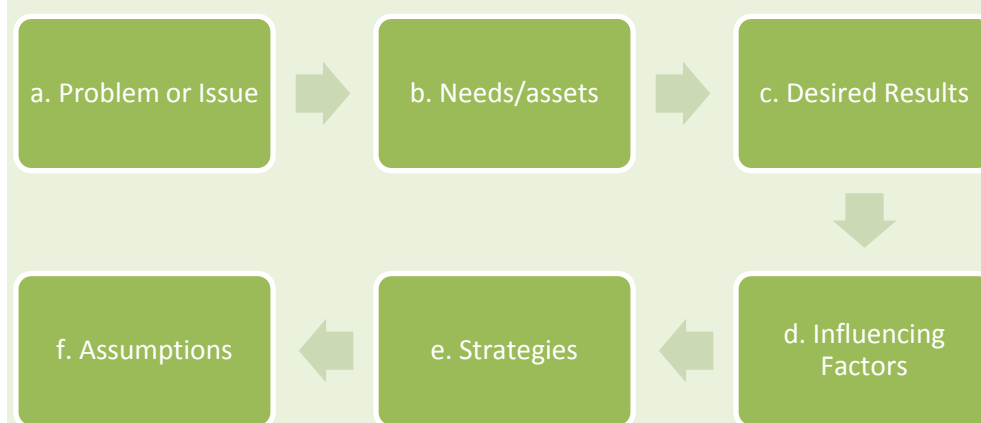
Drawing a Picture of the CEPA Programme’s Theory of Change

Now that we have considered the nature of our beliefs and assumptions, and where they come from, we are in a better position to draw another model or picture of our understanding of why a particular CEPA programme should lead to the desired change.

Is it clear *why* the selected activities, outputs and outcomes will create the desired impact among these participants? The answer to this question constitutes the CEPA programme’s *model of change*, which supports and builds upon the logical framework developed in Step 2. Successful programmes create a desired change and are built on a solid understanding of what works – Pawson and Tilley¹⁰ call this understanding, the *programme theory*.

Systematically work through the following programme and evaluation planning processes, in order to describe the basic theory that underpins the CEPA programme you wish to evaluate, and its change strategy (Figure 5 provides a possible template):

Figure 5: A Theory of Change Template



- Define the problem** the CEPA programme is attempting to address (e.g. which biodiversity issue in this City, which educational issue, the target group(s) and why they are important). Explain concisely the issue you will address. The model of change will be built upon this statement, which should illustrate how the CEPA programme will function or functions, and what it expects to achieve in the city. We try to refer wherever possible to research about the problem or issue, e.g. a State of the Environment report; consultative workshops with CEPA specialists can provide other successful programme or “best practice” information.
- Quantify the scope of the needs or assets** that led to the selection of this particular

¹⁰ Pawson, Ray and Tilley, Nick, 1997. *Realistic Evaluation*. Sage.

problem. Documenting the needs and assets helps the evaluation plan later on. It can become a baseline providing indicators that measure progress made by the CEPA programme over time.

- c. **Describe the desired results.** These are the outputs, outcomes and impacts you have listed in your logical framework.
- d. **Identify contextual factors** that could influence the outcomes, either by helping or by hindering (barriers). Are there perhaps policies that could affect your CEPA programme? Look at previous evaluations of similar programmes, as they might identify some of these barriers and enabling factors.
- e. **Why do you believe this programme will work?** Look for a rationale in research into effective CEPA programme strategies and evaluations of what worked, or didn't work, in other cities or situations like this. Connect what you plan to do, with why your approach will succeed. Funders would like to see evidence that supports the proposed solutions. Apply best practice guidelines that support plausible solution strategies for the identified problem area (for example that active 'hands-on' involvement with the issue will bring about the desired learning and behaviour change among residents and staff.)
- f. **Why will your approach be effective?** After you make the case for selecting a specific strategy from among the alternatives you researched, state why your CEPA programme strategy is needed and why it will work in your city. It should for example be apparent how the programme intends to function as an intervention in terms of biodiversity benefits. List these *assumptions* last because in this format, you have the benefit of all the information that supports your assumptions. They are then easier to spot and articulate with all the facts in front of you.

Here is a fictional example of the first processes, based on the city of Edmonton's Master Naturalist Programme (see the *Case Study Folder* on the CD):

Table 4: Towards a Theory of Change Underpinning Edmonton's Master Naturalist Programme (editor's own examples):

Describing the CEPA programme's theory of change	Possible Responses (editor's own examples)
Define the problem the CEPA programme is attempting to address	Edmonton has many special natural areas that contribute to quality of life in the city, but skilled manpower to effectively manage and protect all these sites is limited; as a result natural areas are invaded by alien vegetation and wetlands are threatened by inappropriate development which may cause reduction in ecosystem services and quality of life.
Quantify the scope of the needs or assets that made the case for the selection of this particular problem	X (number) natural areas comprising Y hectares are currently unmanaged, and the City of Edmonton has only Z site managers and no volunteer stewards at this time.

Desired results	X natural areas are effectively protected and co-managed by City staff and knowledgeable volunteers.
Identify factors in the context that are likely to influence the outcomes, either by helping or by hindering (barriers).	Willingness of many Edmonton residents to participate in programme, but the distribution of the volunteers may not match the distribution of sites that need co-management.
Apply best practice research that supports plausible solution strategies for the identified problem area.	Learning through doing, working collectively in communities of practice strengthens commitment and skills.

Complete a table like the above for the CEPA programme you wish to evaluate, then map out a theory of change template such as the one in Figure 5.

That takes us to the next part of the evaluation design, which involves preparing evaluation questions that test the assumptions underpinning the model of change.

Testing Assumptions

Assumptions are explored by adding probing ‘double loop’ questions to your logic model. By being explicit about our assumptions that underpin our models of change, we allow ourselves to also reflect back on or review these assumptions during evaluation. This adds a basis for evaluation that can be particularly helpful in explaining why a particular intervention or programme works, or fails to work. An important tool to help us identify the assumptions behind our models is the *ladder of inference*.

Also see *Appendix 5: Most Significant Stories of Change* on the CD. This valuable evaluation methodology surfaces and works with participants’ assumptions about CEPA success.

When evaluating a CEPA programme, it is important to evaluate not only whether it is producing the intended outputs and leading to the desired outcomes and impacts, but also if not – why not? Double Loop Learning Questions to add to the evaluation could include:

- Are all the underlying assumptions correct?
- In drawing up the model of change, did CEPA practitioners allow for discussion and debate of a range of theories?
- Does the model of change take into account that change is not necessarily a simple linear process?
- What unintended outcomes and impacts are evident, and what might their effects be?

Asking key questions such as these for the evaluation can be related back to the model of change and the underlying assumptions, and can help CEPA practitioners to refine and if necessary, re-define their programme (adaptive management).

Figure 6 further illustrates questions for what is called *double loop learning*. Where single loop questions are about inputs, activities, outputs, outcomes and impacts, double loop

questions are about the underlying assumptions – in this case, about the *relationship* between these single elements. The developmental evaluation process involves asking these kinds of questions on a regular basis with a number of feedback loops to facilitate continuous learning and double-loop learning.

Using the logical framework and theory of change maps you created, and after revisiting the ladder of inference, create ‘double loop learning’ questions to test the assumptions about the relationships between the elements of the logical framework, and the assumptions underpinning the theory of change of the CEPA programme you want to evaluate.

Figure 6: Examples of Double Loop Learning Questions about Assumptions

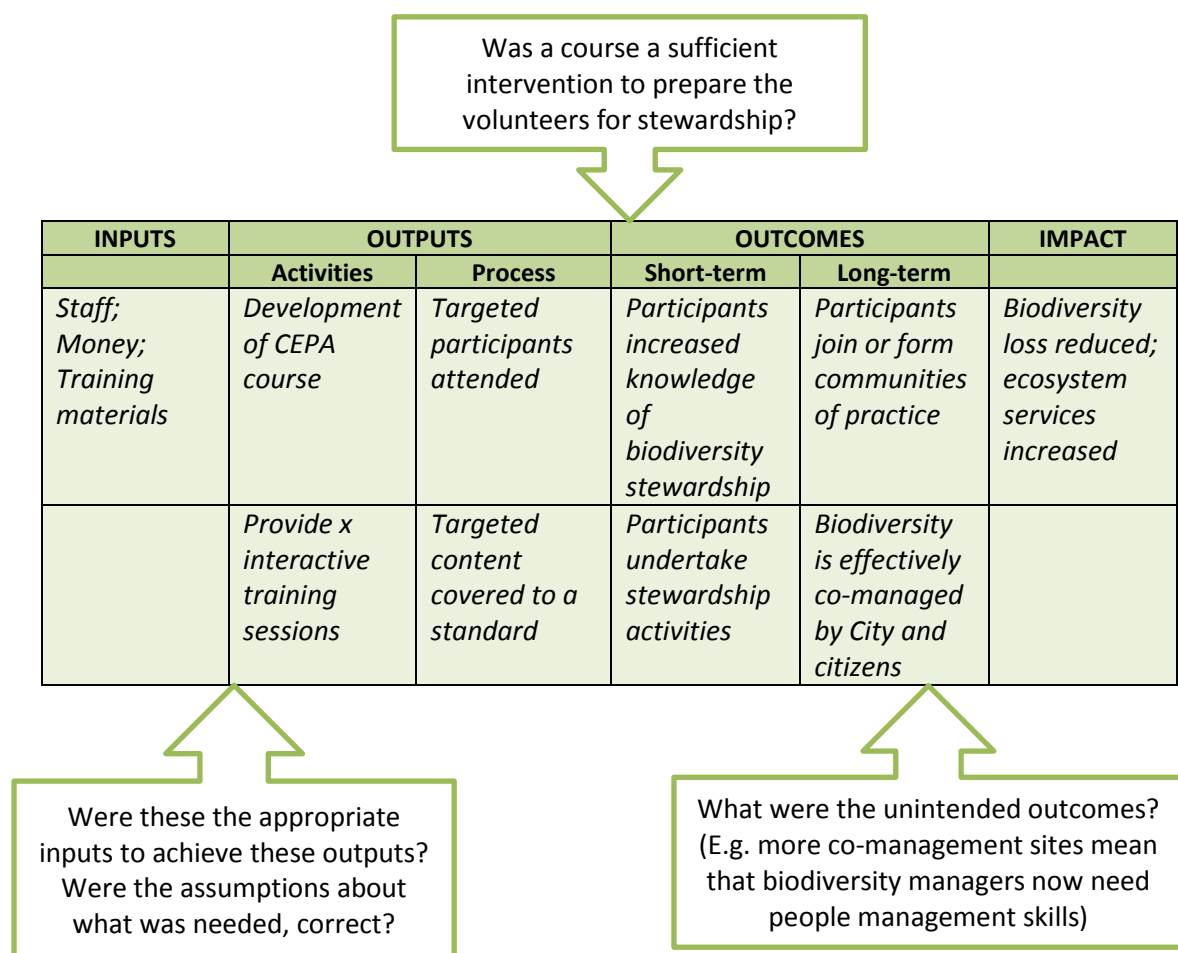
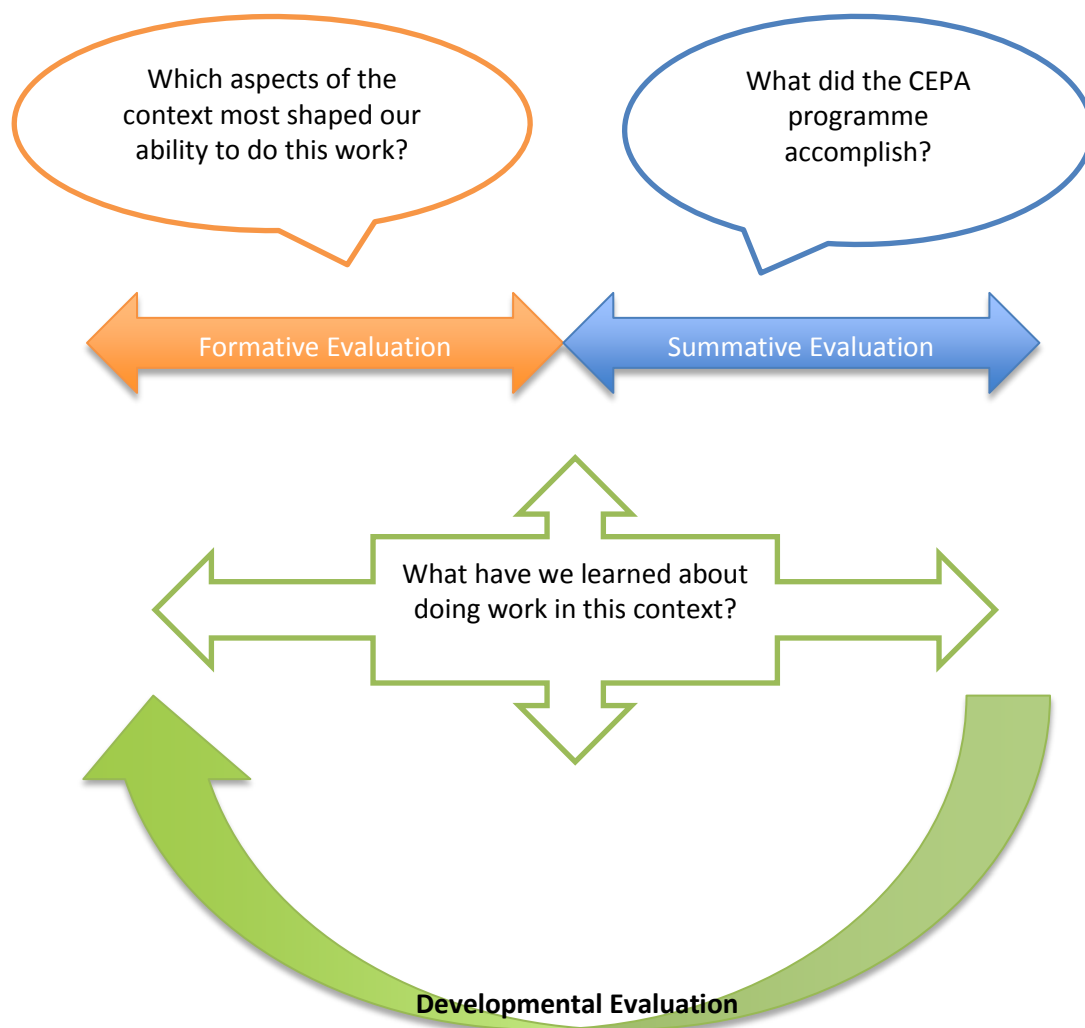


Figure 7: Overarching ‘double loop learning’ questions in different kinds of evaluation



STEP 4: UNPACK THE CONTEXT

Exploring context is about understanding how the CEPA programme functions within the economic, social, institutional and political environments in which it is set. We need to consider whether a particular model of change is appropriate within the context of the particular CEPA programme. What factors in the context might influence our ability to implement the planned programme? Did the CEPA practitioners perhaps assume a very different kind of context to the one that actually exists?

Such evaluation questions can help us explain some of the strengths and weaknesses of a programme as well as the effect of unanticipated and external influences on it. This in turn can help us explain why, or why not, a particular programme works.

Demonstration of assumptions about context

Cape Town's *Smart Living* Campaign designed an environmental resource use audit that was suitable for the home. It made assumptions about the ease of measuring energy and water consumption, and waste production, in the context of the typical family home. Here it is relatively easy, as residents typically receive monthly utility bills from the local council, which indicates their water and electricity usage from the municipal supply. They can also measure their electricity supply from a meter in the home; and they can measure the volume of waste produced by direct observation of the waste bins they leave outside the home for collection on a particular day of the week.

When the Green Audits Programme¹¹ applied the same assumptions to schools, however, these assumptions did not seem to apply that well to the new context. Students could not readily measure the amount of energy and water used or waste produced at the school. Schools consist of multiple buildings; utility bills are usually combined for different buildings; are sometimes issued quarterly rather than monthly; and could only be accessed after prior arrangement with management staff. Water and electricity meters are often in inaccessible places or out of bounds for the students. Waste is produced in multiple sites (offices, residences, kitchens, tuck shops) and disposed of in a variety of ways, on different days of the week.

Failing to take these differences in context into account, and planning adequately for them, could spell trouble for a CEPA programme requiring consumption measurement in schools.

Figure 8 below illustrates that a developmental evaluation process (and the indicators for it) would ask questions about the CEPA programme, but also about its context, and about the mental model of or assumptions about the programme and its context.

Using the second half of Figure 8 as a possible template, list all the critical features of the context of the CEPA programme you want to evaluate.

We find it useful to identify economic, political, cultural, organisational and bio-physical factors, at multiple levels. For example, economic factors at national, regional, organisational and international levels may all be significant features of a CEPA programme's context. For more guidelines on this step, see below.

¹¹ See the *Case Studies Folder* on the CD.

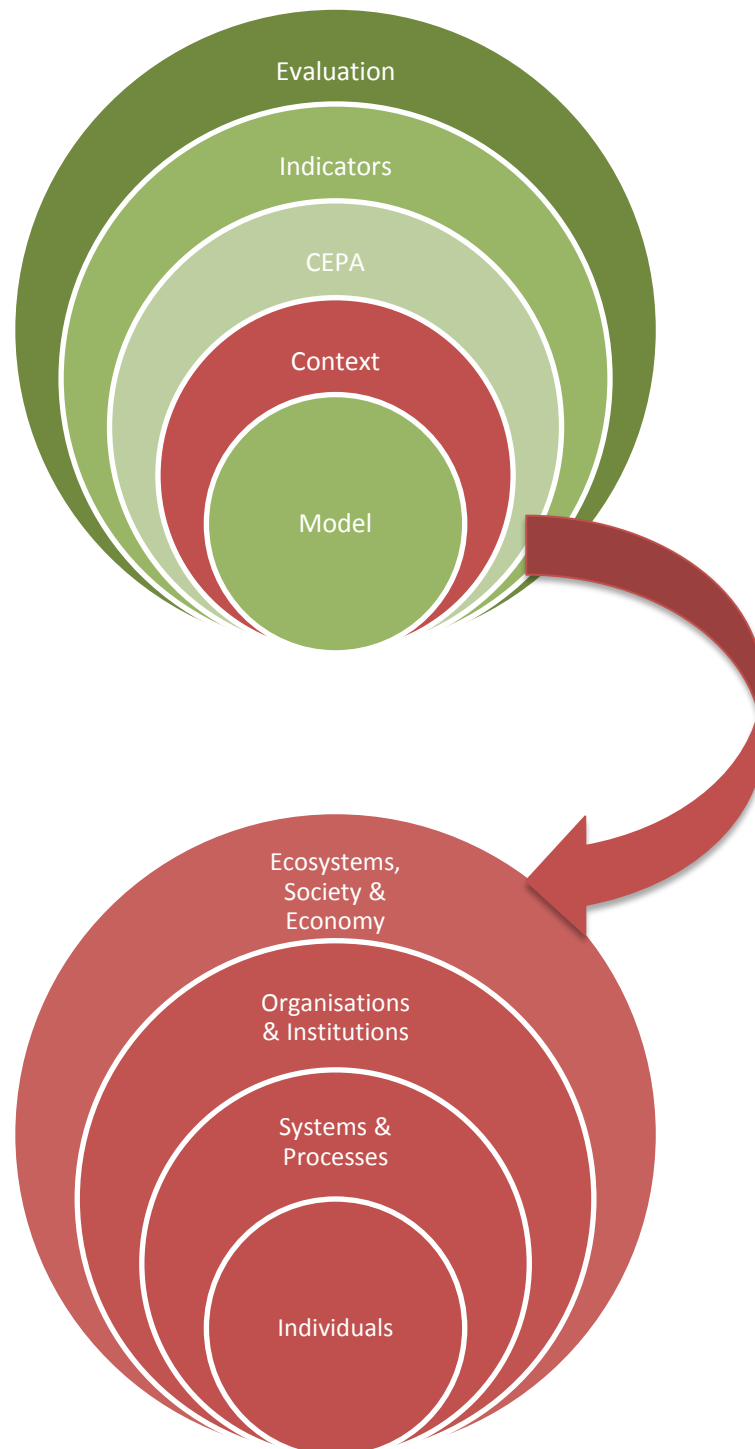


Figure 8: Aspects of Context and the Role of Context in a CEPA Programme Evaluation

We find that how we define the CEPA programme's context and what we choose to include in an evaluation of context, depends to some extent on the scope, size and duration of the programme, and to a large extent on its actual focus.

In a large scale, long term programme like Nagoya Open University of the Environment¹², for example, 'context' would certainly include the broader context of the society of Nagoya, the role of Japan's economic, business and other social systems, which may influence citizens' values and lifestyle decisions, as well as a variety of institutional role players, including various tiers of government and the education system, from schools to universities. Contextual factors such as the large scale natural disasters that have been affecting Japan would be particularly significant, for example in determining what content is on offer, and how citizens relate to this content. In other societies, the contextual factors would differ.

In a smaller scale initiative, such as the City of Cape Town Green Audits for Schools, the national economy may not be that significant, but the local economy might be, if one were to consider refurbishing schools to reduce resource consumption. The international context of donor funding for 'green energy' technology could be considered an important factor in this context, too. Local institutional contexts are also important, for example the different kinds of management evident at different schools¹³. The national school curriculum, which determines what teachers and learners should emphasise at school, can also influence the extent to which they prioritise biodiversity related CEPA activities.

Decide which contextual aspects are relevant to the evaluation you are planning, and add key questions in relation to these contextual factors.

Below are some examples of questions to ask about the context. At the start of the programme:

- Which features of the context are critical for programme success?
- Which features of the context may prevent programme success?
- Do the assumptions of the change model apply in this context?

And during the course of the programme:

- In what ways is the programme being influenced by its context?
- Which contextual factors seem to be particularly influential in shaping the outcomes of the programme?
- In what ways is the programme influencing its context? Or: Which aspects of the context are being influenced by the programme?
- Which features of the context seem to be at odds with the CEPA programme's logic and change model?
- To which features of the context does the CEPA programme seem to respond particularly well?

¹² See the *Case Study Folder* on the CD.

¹³ Ibid.

As with the linear logical framework, the *process* of producing the map is important. Whether the map accurately reflects the system is less important; mapping one's thinking about the system is important, as this creates opportunities to evaluate and refine that thinking where necessary. *Hence the mapping process is again most useful if done with the CEPA programme staff, as this helps to surface all assumptions and understandings of the programme, its change theory and its context.*

Figure 9 is an example based on the case study of a CEPA programme accompanying the re-introduction of howler monkeys (*Alouatta clamitans*) in Atlantic rain forest remnants in the city of São Paulo¹⁵. The content has been generated by the editor, drawing on the background to the case study as well as some assumed factors, which may or may not apply in the actual context. The systems map is provided simply for the purpose of demonstrating how one could represent a particular system, and the causal loops within it.

Draw one or more causal loop system maps for the CEPA Programme you are about to evaluate. Then add key evaluation questions that will allow you to test the CEPA programme as well as the underlying assumptions on which programme activities are based. These questions will then require you to look for evaluation data, best practice guidelines or expert opinion to support or refute the postulated trends, and the links between them.

For example, in the above example, an evaluation team could ask questions about whether there has been an increase in CEPA programmes with forest neighbours as well as passing drivers, whether these programmes have resulted in greater awareness among the neighbours and the drivers; and whether this awareness has in turn resulted in behaviour changes, for example, whether drivers are reducing speed and injuring fewer monkeys, or returning more injured monkeys to the rehabilitation centre.

'Double loop learning' questions could also be asked to test the assumptions that inform the programme activities. For example, based on the fictional systems map in Figure 9 evaluation questions could be asked to determine whether drivers reduce speed in rain forest areas where monkeys occur because of speed control measures (such as speed humps, signage, or prosecution by traffic police) or because of a greater awareness of the importance of the rain forest and its inhabitants?

Finally, add evaluation questions about any *unintended consequences* that might be occurring in the system. For example, a programme raising awareness about the reintroduction of endangered species in the rain forest might stimulate or increase the efforts of collectors or hunters to track down the reintroduced animals. This will be a consequence to avoid.

¹⁵ See the *Case Study Folder* on the CD.

STEP 6: DEVELOPING INDICATORS

“That which is good and helpful ought to be growing and that which is bad and hindering ought to be diminishing We therefore need, above all else ... concepts that enable us to choose the right direction of our movement and not merely to measure its speed.”¹⁶

“The search for indicators is evolutionary. The necessary process is one of learning.”¹⁷

One of the biggest challenges in developing an evaluation plan is deciding what kind of information would best answer the evaluation questions. Indicators are the measures you select to answer the questions you have posed. They act as markers of progress and success. They are central to the design of evaluation processes and for data collection and reporting.

Indicators are often likened to the icons on a car’s dashboard, that indicate (for example), at what speed we are driving, whether our headlights are on, how full the fuel tank is, and so on. A red light often signals that the car is about to cross a dangerous threshold, while an absence of red lights could mean that all is well! In a CEPA evaluation, typical indicators might be the number of participants from different groups attending a CEPA course, the degree of satisfaction expressed by participants on the course, and the level of relevant knowledge gained by them.

Indicators are not ends in themselves. The red fuel tank icon on the dashboard is not the fuel tank itself. A CEPA course, the participation in the course, and the satisfaction of the course participant are probably not end goals in themselves, either. There is something else we want to achieve through people’s participation in our courses – for example, growing their capacity to act for biodiversity.

At the same time, the nature of the indicators we choose and work towards can have a very real impact on CEPA programmes. For example, if we set a target of reaching 10,000 citizens to attend our courses, this is likely to push CEPA practitioners’ efforts towards attracting more and more citizens to courses, at least until the target is met. If this is the only or main indicator in the evaluation, it can have the effect of detracting the CEPA practitioners’ attention away from other considerations such as the quality and relevance of the courses.

In this toolkit we promote an approach to indicators that promotes reflection on practice rather than simply hitting targets.

Indicators are a central part of effective CEPA programme decision-making and adaptive management. They can provide measures of the progress and success of policies and programmes, and they can form part of an ‘early warning system’ to detect and fix problems as they arise. Indicators can be used to raise awareness about an issue. An example would be a drop in the number of observed howler monkeys in São Paulo’s rain forest. The same indicator can then be used to put responses to this issue (a reintroduction programme and related CEPA activities) into context, as is done in the case included in this Toolkit¹⁸.

¹⁶ Schumacher, E.F., 1989. *Small Is Beautiful: Economics as if People Mattered*, Harper Perennial.

¹⁷ Meadows, D., 1989. *Indicators and Information Systems for Sustainable Development*. The Sustainability Institute, Vermont.

¹⁸ See the *Case Study Folder* on the CD.

Indicators by themselves, however, provide little understanding of an issue. They always need some analysis and interpretation of what they are indicating. Just knowing that there has been a drop in the number of howler monkeys in São Paulo would not mean much, unless we knew that there was a concomitant increase in illegal capturing or hunting of the monkeys, or a disease that struck the local population, or a decrease in the area of natural habitat (Atlantic rain forest) due to urban expansion.

Indicators don't guarantee results. But well-chosen indicators, in themselves, can produce desired results. Donella Meadows gave the example of industries in the United States that started to reduce emissions in the absence of stricter laws, in response to the indicator (air pollution level per company) being made known to the public.

On the other hand, if the indicators of success are wrong, then no amount of measuring, reporting, funding, action, political will, or evaluation will lead toward the desired outcome. Compare the following two indicators – which one is likely to lead to a more effective reintroduction programme?

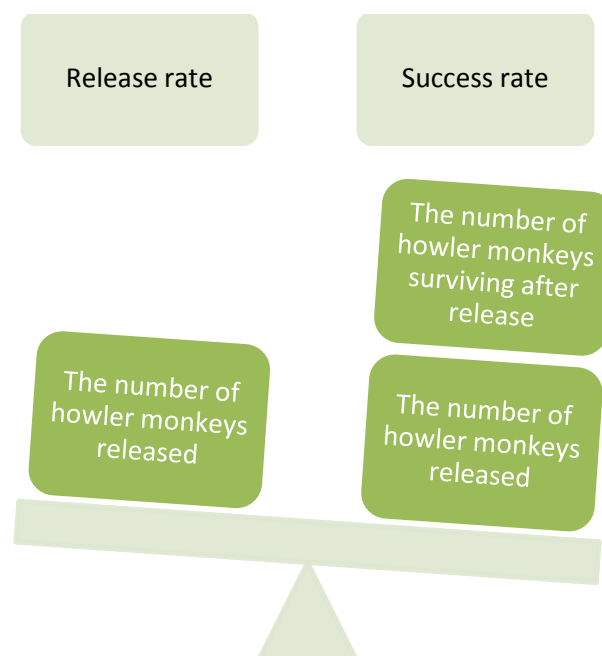


Figure 10: Comparison of Two Different Indicators for the same Programme

It was precisely because they needed to have a better indicator for success than the number released, that the São Paulo biodiversity managers in our case study introduced a system of monitoring groups of introduced howler monkeys.

The Challenge of Indicator Development

We contend that there can be no universal set of indicators for CEPA programmes that can be used in all contexts. Indicators are purpose-dependent, and the indicators we choose, will vary with our purpose. To the extent that we share purposes in CEPA activities, there will be some commonalities in our indicators, and the examples in Table 5 and Table 6 will no doubt be useful to many CEPA practitioners.

Also consider Donella Meadow's advice: *"What is needed to inform sustainable development is not just indicators, but a coherent information system from which indicators can be derived"*.¹⁹

In addition to providing some examples of common indicator examples relevant to CEPA programmes, this toolkit promotes a process for developing CEPA indicators based on:

- mapping the logical framework of the CEPA programme
- identifying the underlying assumptions and models (programme theory)
- developing indicators for different stages of the programme
- testing the results against these indicators, and then
- re-thinking or re-designing the programme *and the indicators*, if necessary.

Drawing on our case studies, we provide examples of types of indicators that could be useful in each case. *Note that these are not necessarily indicators that the case study practitioners had actually used*, but they are indicators that *could* be used in similar situations.

Illustration of an Inappropriate Indicator – Fictional Case

Imagine for one moment what *could* happen if a government were to decide that each child in the city should receive a book about the forest. Let us say that behind this is the goal of educating the city's children from a young age to understand and appreciate the forest. But say the evaluators inadvertently choose an inappropriate indicator, namely: *Every child in Year 1 should receive a book on the forest*.

To try to 'achieve' this indicator, the CEPA staff may put a large budget and all their effort into effectively obtaining and distributing the books. They are likely to have much less budget and time left to ensure that the books have good quality content, are suitable for this age and language ability (including diverse languages across the city), and that teachers are willing and able to introduce the books to the children with enthusiasm. In other words, in our imaginary example there are no indicators for quality, relevance, or use of the books.

Around the world there are examples where such a choice of inappropriate indicator has resulted in children receiving books that did not contain correct information or messages, were not attractive, were not in their home language, or failed to be promoted by teachers – and yet, the indicator – *Each child should receive a book* – would have been achieved and the programme could have been regarded as a success!

*"As you know, what usually happens is that we can only measure simple things, and then because that is what we can measure, we say that those simple things are the only real things. So we count numbers, do simple pre-post treatment surveys, look for short-term changes, measure things, and then write our report. The real things, the ways in which environmental education can change someone's life, are much more subtle and difficult to measure. You can ask questions about meaning, about influence, about impacts, and look at things that aren't visible necessarily over a short time, but become apparent over the long term. This is what we have to consider as we look at effectiveness of environmental education."*²⁰

¹⁹ Ibid

²⁰ Meadows, Donella, 1998. *Indicators and Information Systems for Sustainable Development*. The Sustainability Institute, Vermont.

What are Good Indicators?

Indicators are most useful²¹ when they are:

- Representative of what one wants to find out about the programme
- Relevant and useful to decision-making (stakeholders care about this measure)
- Easy to interpret
- Sensitive to change
- Feasible and cost-effective to obtain
- Easily communicated to a target audience.

However, just because an indicator is easy to measure, easy to interpret and cost-effective to obtain, it doesn't mean that it is a good indicator. These considerations should not limit the choice of indicators.

It is quite easy to list the characteristics of ideal indicators, and much harder to find indicators that actually meet these ideal characteristics. It is fair to say that the development of indicators is one of the most difficult parts of the evaluation planning process.

Bear in mind that indicators can take many forms. They don't have to be quantitative (numbers). They can be qualities, signs, symbols, pictures, colours.

Involve stakeholders in developing indicators

The *process* of developing indicators requires careful attention. It is strongly recommended that all evaluation stakeholders (however you define them) are consulted as early in the process as possible in order to determine the purpose of the indicators. Who would these stakeholders be? The indicator selection process works best with a careful combination of expert and grassroots or non-expert participation.

In the case of the Nagoya Open University of the Environment, for example, the stakeholders who could help determine indicators may be experts and direct users of the indicator (the CEPA programme managers and the programme steering committee), those with a broader interest in the issues surrounding the programme (e.g. environmental managers, funders and other institutional partners), and those holding relevant data (e.g. the course designers and trainers).

Consulting with these groups and identifying their needs will help to clarify how simple or complicated the indicator needs to be, and the most appropriate ways of communicating and interpreting it.

Most of us already have indicators in the back of our minds, based on issues of particular concern to us. It is important to get them out on the table at the start of the indicator development process. As indicators are selected and defined, stakeholders will express their values, purposes will be agreed upon, change models will be at play, and programme

²¹ Adapted from *Evaluation Sourcebook: Measures of Progress for Ecosystem- and Community-based Projects*, 2006, Schueller, S.K., S.L. Yaffee, S. J. Higgs, K. Mogelgaard, and E. A. DeMattia. Ecosystem Management Initiative, University of Michigan, Ann Arbor.

theories will be developed and shared (implicitly and explicitly). *The indicator selection process is the place where the legitimacy and comprehension of an evaluation are built, as people see their values incorporated into the indicators.*

The most significant change story methodology²² mentioned earlier is a useful strategy for surfacing values and developing agreed-upon indicators for a further evaluation phase.

Questions to ask during this step:

- Who are the relevant stakeholders in this programme, and do they all need to be consulted in the development or choice of indicators?
- How much ownership and decision-making power are different stakeholders going to have over the choice of indicators?
- Have the inputs, expectations and outputs of the indicator development process been clearly defined for the stakeholders?
- Do the stakeholders want to use the indicator(s) for decision-making, for reporting purposes, and/or for continuous learning? Any other purposes?

Relating Indicators to Evaluation Questions

In the preceding steps we have worked towards posing a range of evaluation questions. Once one has chosen which of these questions are most important to ask at this particular juncture in the CEPA programme's implementation, indicators should be developed to these key questions. The evaluation question defines the purpose of the indicator, and what its user wants to know about it.

One of the benefits of defining a key question is that it encourages the selection and communication of the indicators in a form that aids their interpretation. The logic of addressing a key question also encourages further analysis to explain complex issues. The more precise and specific to a situation a key question is, the more guidance it gives for the selection and development of suitable indicators.

It may be necessary to use several indicators and data sets to answer a single key question. Relying on just one indicator can distort one's interpretation of how well a programme is working. On the other hand, the total number of indicators needs to be a manageable number. Identifying a core set of indicators is a good way to proceed.

Table 5 below shows types of indicators that can be used to answer evaluation questions related to the different components of the CEPA programme's logic model.

²² See Appendix 5 on the CD.

Table 5: Establishing Indicators to Answer Evaluation Questions

INPUTS	OUTPUTS		OUTCOMES		IMPACT
	Activities	Process	Short-term	Long-term	
Staff; Money; Training materials	Development of CEPA course	Targeted participants attended	Participants increased knowledge of biodiversity stewardship	Participants join or form communities of practice	Biodiversity loss reduced; ecosystem services increased
	Provide x interactive training sessions	Targeted content covered to a standard	Participants undertake stewardship activities	Biodiversity is effectively co-managed by City and citizens	
Key Evaluation Questions					
Was the provisioning of funding and staff sufficient, timely? Were the training materials of suitable quality, content?	Was the required CEPA course developed? Were all x sessions delivered?	Did all intended participants attend? All sessions? Why? Why not? Do the CEPA programmes communicate the issues comprehensively and effectively? Were participants satisfied with the course delivery?	To what extent did knowledge increase? What are participants able to understand and do as a result of an input/activity? How many participants signed up for volunteer stewardship/conservation action?	In 12 months, how many participants are still doing stewardship? How many groups have been formed? What is the scope and quality of their stewardship activities? How many hectares covered?	What is the status of bio-diversity and eco-system services in the city compared to before the programme started? Have goals been reached? What unintended impacts have there been?
Indicators					
Quantitative Indicators					
Number of staff. Amount spent. Number of booklets produced. Quantitative content analysis of booklets.	Course developed. Number of training sessions delivered.	Numbers per group attended per session. Satisfaction expressed as a number.	Pre- and post-course knowledge test scores. Number of volunteers, hours worked.	Number of volunteers after 12 months. Number of groups. Range of activities. Hectares covered.	Conservation status of land e.g. change in species counts, change in numbers of individuals in rare, threatened and vulnerable categories. Change in volume of water from

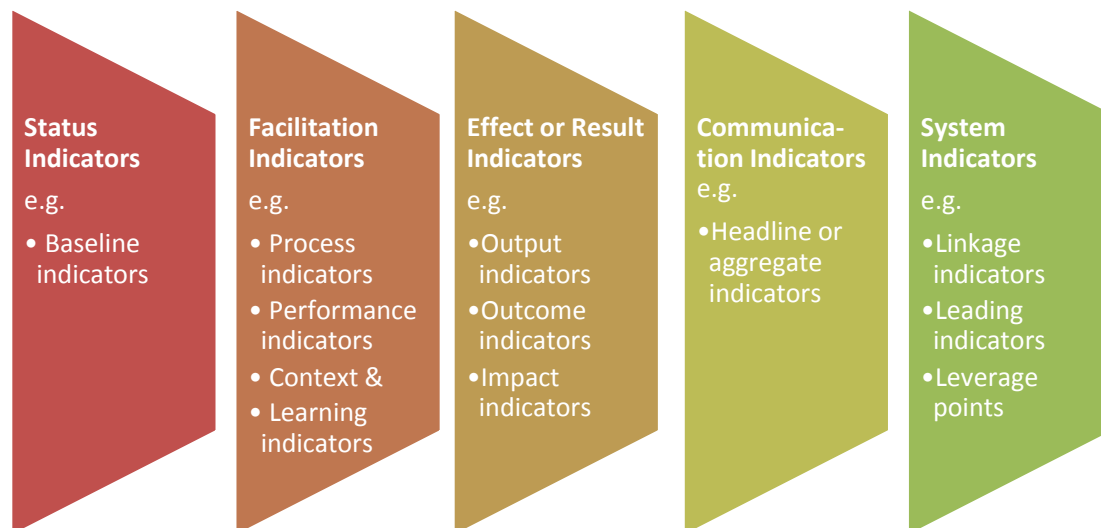
wetlands,
change in
pollution
levels.

Qualitative Indicators & Associated Methods				
<i>Comments on staff skills, capacity. Educational experts' analysis of quality and relevance of materials.</i>	<i>Participant reflection on reasons for attendance, non-attendance; scope and relevance analysis based on expert and participant input during focus group discussion.</i>	<i>Individuals' capacity based on self- and peer assessment, and expert observation of conduct in the field.</i>	<i>Groups' capacity based on self-and peer assessment, expert observation of conduct in the field. Map of areas managed vs. areas not managed, colour coding reflecting levels of management.</i>	<i>Most significant change stories. Identification of unintended outcomes, impacts through stakeholder review process(es).</i>

Types of Indicators Required According to Evaluation Phase

Five broad indicator types are used at different stages of implementation of a CEPA programme. They seek different types of data and are distinguishable by their focus on different variables relating to progress.

The five broad indicator types are:



Status Indicators

These assess variables that determine the position or standing of the CEPA programme. Baseline indicators belong to this category. Baseline indicators help to identify the starting points for change and provide reference points in identifying realistic impact indicators.

In the case of Edmonton's Master Naturalists Programme, a status (baseline) indicator could be the number of knowledgeable volunteers who are involved in the stewardship of the city's natural areas, at the start of the programme. In the case of Nagoya Open University of the Environment, a baseline indicator could be citizens' knowledge and commitment to biodiversity before they attend the Open University. In the case of the reintroduction of Howler Monkeys in São Paulo City, a baseline for the CEPA component could be the number of forest neighbours with a positive attitude to preserving the monkeys and their habitat.

CEPA practitioners interested in influencing the content of the school or university curriculum may start with a review the status of the current curriculum, by looking for indicators of biodiversity related content currently covered in the various subjects. Appendix 2 on the CD provides an example from an Australian government review of environmental content in educational resource materials.

Facilitative Indicators

These assess variables that assist, support or encourage engagement with CEPA programmes. Process questions are concerned with the quality of programme delivery and how well programmes have been implemented. Facilitative indicators show whether planned activities are actually carried out and carried out effectively and/or according to available guidelines. Facilitative indicators may measure the number of outputs generated, participant and partner satisfaction with these outputs, and other aspects of programme implementation. Context, process, and learning indicators belong to this category.

In the case of the Master Naturalists Programme, facilitative indicators would include the number of courses offered to volunteers and the quality and relevance of this training. In São Paulo the facilitative indicators would indicate whether residents have been reached by planned CEPA programmes, how many activities were offered, how satisfied various partners were with the quality of these activities, but also, what participants actually learned, if CEPA programmes were regarded as facilitating the reintroduction programme.

Effect or Result Indicators

These indicators assess variables related to initial, medium and long term achievements during the CEPA programme. Output, outcome and impact indicators belong to this category. The outcome questions and indicators often look for evidence of change in participants' awareness and behaviours over time. Impacts are the broader, long-term changes that a programme has on society and environment. The questions and indicators may look for evidence that the state of biodiversity has improved over time, or that more citizens enjoy the well-being associated with functioning ecosystems and intact biodiversity. Impact indicators assess progress towards these objectives:

- Short term impacts on individuals and organisations (e.g. changes in Cape Town school students' understanding of resource use at school, and reduced resource use in Cape Town schools)
- Longer-term impacts on practice at different levels, such as:
 - changes in practices (such as curriculum changes and the institutionalization of resource use reduction measures at schools e.g. regular recycling, installation of energy saving appliances and water wise landscaping);
 - *organisational change* in terms of policy (e.g. curriculum policy on environmental education; local government policy on urban planning) and
 - *growing* partnerships (e.g. stewardship sharing between government and residents).

The long-term influence of CEPA programmes is difficult to assess, not only because it requires long term commitments to collecting data, but also because many other factors, beyond the programme, can influence such changes. For example, in Cape Town the cost of water and electricity use may increase significantly, and if residents reduce their consumption of these resources, it would be difficult to distinguish the impact of this variable, from the (perhaps additional) impact of the Green Audits CEPA programme. Or, in

the case of the reintroduction of Howler Monkeys in São Paulo, even if CEPA programmes result in high levels of awareness of and care about the rain forest remnants among local residents, if policies and population pressure lead to rapid expansion of urban areas, the rain forest may reduce to such an extent that Howler Monkey populations cannot be sustained.

Communication Indicators

These indicators are for disseminating information relating to a range of evaluation questions in an accessible way that facilitates communication to stakeholders. Examples of communication indicators are headline or aggregate indicators, which are the sort of statements that could make it into a regional or community newspaper. Examples of headline indicators could be the number of Cape Town schools actively recycling their waste; or the number of hectares that are now under volunteer stewardship in Edmonton. The mixed nature of some stewardship groups (consisting of old and new Edmontonians) could also provide a headline indicator demonstrating widespread support for biodiversity management.

System Indicators

These indicators provide an overall picture of the state of a CEPA programme. They can provide an indication of the status of the programme, programme processes or impacts, or all of these combined. They can be very useful for communication purposes and for further programme visioning exercises. Systems change over time and we find it helpful to look for indicators to tell us about this dynamic behaviour. Systems dynamics is a field of expertise that specialises in the unfolding behaviour over time of whole systems. *System dynamics can be useful in finding linkage indicators, leading indicators, and leverage points where systems are especially likely to signal change or respond to action*²³. Metaphors are also valuable for providing a 'picture' of the overall status of the system (e.g. comparing a CEPA department to a healthy diverse ecosystem).

In the City of Cape Town's overarching Environmental Education and Training evaluation²⁴, the CEPA activities of its Environmental Resources Management Department were compared to a tree that has grown very large, with many branches and leaves (activities), but a weakness in the connections in its trunk and to its roots (alignment between departments and alignment with vision and policy intentions).

²³ See *Appendix: Leverage Points in a System*, on CD.

²⁴ See *Case Study Folder* on the CD.

Table 6: Indicator Types Using LAB CEPA Programme Examples²⁵**Examples of Status Indicators for LAB CEPA Programmes**

Indicator Type	Function	Quantitative Indicator Examples	Qualitative Indicator Examples
Status - Baseline	To describe the status of the overall CEPA picture To describe the status of the overall Local Action for Biodiversity picture	% of local government departments currently providing CEPA programmes with a biodiversity component; % of citizens who actively participate in biodiversity protection measures; % of conservation worthy land in the city that is protected and/or well managed.	Metaphors or 'one liners' describing the attitude of various citizen and city staff groups towards biodiversity, before a CEPA programme for these groups start; Photographic record of conserved and degraded sites around the city.
Status - Context	To identify the existence of CEPA support systems	Policy exists that requires CEPA programme for biodiversity in local government; Coordinator and staff appointed to assist local government with integrating LAB CEPA programmes into service delivery.	Concepts and principles in national curriculum policy on biodiversity content in schools; Strength and quality of volunteer CEPA activities support to local government.

Examples of Facilitative Indicators for LAB CEPA Programmes

Indicator Type	Function	Quantitative Indicator Examples	Qualitative Indicator Examples
Process	To identify the existence of CEPA processes and activities, and to what extent they have been implemented.	Number of citizens and range of citizen groups reached in CEPA activities; Number of press releases with an environment or biodiversity focus; Attendance at biodiversity related	Feedback from stakeholders about how programme is being implemented; Quality of responses to a biodiversity debate during a radio phone in programme; Evidence of good CEPA

²⁵ Adapted from Education for Sustainability indicator guidelines produced by Daniella Tilbury and Sonia Janousek in 2006, published by the Australian Research Institute in Education for Sustainability, with additional examples by present authors.

		events; Number of teachers who use biodiversity related materials in their teaching % of activities completed within timeframe; Number of hours spent on activities relative to priorities.	practices according to theories of change and best practice guidelines; Expert analysis on classroom teaching on biodiversity related topics; Staff opinions on whether time is used well.
Learning	To promote learning and reflection in and on CEPA programmes.		Identify markers for change - Reflective analysis of case studies of changed practice e.g. schools that reduce water consumption during CEPA projects; Identify conditions for change – Review of the process of adopting a new urban planning policy; Programmatic review of a number of smaller scale evaluations, to look for similarities, differences, patterns and trends across projects; Lessons learned in the evaluation of LAB CEPA activities are captured and shared.

Examples of Effect Indicators for LAB CEPA Programmes

Indicator Type	Function	Quantitative Indicators	Qualitative indicators
Output	To assess outputs such as training resources/course materials, and the immediate results of an activity.	Number of resources developed for LAB CEPA courses and media campaigns; Number of topics covered e.g. biodiversity, threats, ecosystem services, climate change, risk, adaptation, mitigation, resilience, etc. (content analysis).	Adherence to quality criteria in resources developed for LAB CEPA courses and media campaigns; their relevance & policy alignment.
Outcome	To assess outcomes related to	% of new teachers using CEPA-related content in the classroom; Change in attendance at relevant	Level and scope of biodiversity management

	changes or improvements that result from CEPA efforts; To what extent has the community become more aware of biodiversity issues?	events; Number of volunteer hours worked on Local Action for Biodiversity; Number of people who can name threats to biodiversity in a survey; Increases in nursery sales of indigenous and water wise plants and decreases in sales of invasive plants.	activities undertaken by volunteers; Evidence among citizens of pride in local forests and mountain; Case examples of new networks/ communities of practice.
Impact	To assess impacts that result from CEPA efforts: Is there an improvement in the status of biodiversity and Local Action for Biodiversity in the city?	E.g. improvement in water quality and numbers of endangered species in urban wetlands. For biodiversity indicators refer to <i>Biodiversity Indicator Partnership</i> ²⁶ Over a quarter of participants agree that their behaviour has changed in a specific way, e.g., that they keep their dogs out of protected forests; Number of individuals, action groups and volunteer days worked in actions to restore, remediate, or improve a natural area; number of hectares of invasive species cleared; number of wetlands or rivers restored; number of new species discovered by citizen groups; number of hectares newly placed under conservation management.	Most significant change stories which can include: Development decisions in favour of biodiversity; Participants making written reference in journals to their new sustainability practices; action projects/ changes they have made/ special events: in the form of postcards, photographs, videos, journals, and web page entries; Examples of citizens motivating others to join them in taking action.
Performance	To assess the change in the status of the overall CEPA picture in the city and region.	Increase in the number of local governments providing CEPA programmes with a biodiversity component; Numerical comparison in LAB CEPA activities across cities in one region, and across regions; The presence or absence of a number of criteria (i.e. greater	Qualitative comparison in LAB CEPA activities across cities in one region, and across regions.

²⁶ Biodiversity Indicator Partnership, www.bipindicators.net/indicators.

budgets for biodiversity conservation, more active participation among a wider range of citizens in volunteer programmes);
Can be same as status indicators for comparison to baseline.

Examples of Communication Indicators for LAB CEPA Programmes

Indicator Type	Function	Quantitative Indicators	Qualitative Indicators
Headline	To provide a 'dashboard' summary for communicating at a high level	Headlines of priority indicators for respective stakeholders; for example <i>"More than 20,000 citizens participate in Nagoya Open University of the Environment"</i> ; Use <i>Wordle</i> for an overall 'picture' of the key qualitative words describing outcomes/impacts from case studies or surveys (see cover of the toolkit).	Headlines of priority indicators for respective stakeholders; for example <i>"City of Cape Town plays a significant role in city-wide environmental awareness"</i> ;
Aggregate	To provide an aggregated summary of key indicators for communicating at managerial level	For example, <i>"Overall growth in CEPA activities follows climate change summit"</i> .	For example, <i>"Biodiversity CEPA activities significant in service delivery"</i> ; this can be 'clickable' to unpack further; can also include qualitative case study references.

Examples of Systems Indicators for LAB CEPA Programmes

Indicator Type	Function	Quantitative Indicators	Qualitative Indicators
Combinations, linkages and leverage points	Provide an overall picture of the state of CEPA programmes as systems. Indicate status, processes, impacts or all combined. Good for	Systems dynamics – change in biodiversity; change in citizen involvement over time. Systems linkages – change in biodiversity related to specific citizen values, actions or contributions.	Typically images, stories or metaphors are valuable here. For example, the use of a tree as a metaphor to describe the overall status of a programme system – is the programme weak but

communication. Indicate linkages and leverage points where the system is likely to signal change or respond to action.	Leverage points in the system (change in city development policy, re-zoning, new legislation – see the relevant Appendix on the CD and add your examples!	with many branches, is it strong within a thriving ecosystem due to all its partnerships, or is it at risk from a drought as resources dry up?
--	---	--

Questions to ask during indicator development:

- How well does the potential indicator help to answer our key question(s)?
- How does the indicator act a proxy measure, as opposed to a direct measure, in relation to what is not measurable?
- How does this set of indicators address change or the dynamics in the system?
- Are these indicators likely to foster compliance with laws, or foster learning and innovation? Which is required now?
- What are the alternatives to the indicator set?
- What are the best indicators that will influence appropriate change in this system?
- What are the resources available now and in the future for producing the possible indicators?
- Are there existing indicators that can help to answer the key question(s)?
- When will we review and if necessary revise this indicator set to keep it relevant and helpful?
- What guidance can we offer partners and participants (e.g. CEPA course planners and trainers) on how the indicators are to be interpreted?
- Are the pitfalls in the selection and use of the indicator set transparent and explicit?
- What are the incentives for developing, using and acting upon the indicator set and its findings?

Establishing Targets for Indicators

Once an indicator has been selected, it is sometime possible to agree upon specific targets to be reached as a measure of success. For example, if we want to assess whether a CEPA programme increased student knowledge of biodiversity, our indicator could specify “*at least 80% of students will correctly identify three common sources of biodiversity impacts.*” This type of targeted indicator provides a more unequivocal standard of success than one without such a target.

Indicators with their interpretative text can then be part of the definition of targets or objectives. Caution is required, though, if targets are set on the basis of a desired value of an existing indicator, especially if the indicator has been chosen principally because it is something for which there is existing data. It is important to determine the desired state of which the indicator is just an indicator.²⁷

²⁷ Meadows, Donella, 1998. *Indicators and Information Systems for Sustainable Development*. The Sustainability Institute, Vermont.

If a programme already has well-specified objectives, we may be able to extract targeted indicators from these objectives. Consider the following two alternative objectives for a biodiversity education programme:

"As a result of the CEPA outreach campaign ...

Option 1: ... *"The public will be more committed to protecting biodiversity."* This objective is not ideal from a measurement perspective: i.e., the indicator is not explicit. Which public? What does it mean to be *"committed to protecting biodiversity?"* How much *"more"* commitment will there be?

Option 2: ... *"Adult participation in voluntary biodiversity management activities will increase by 50%."* Note how this option offers an indicator for measuring *"commitment to protecting biodiversity"* that is, participation in volunteer programme activities. In addition, it includes a *'target'*, i.e., the expected increase. You can show that you have met this objective if there is at least a 50% increase in participation compared to past, baseline levels of participation.

Relating CEPA outcome/impact indicators to local and national biodiversity goals and targets can also be important, depending on one's approach to CEPA programmes.

All cities and/or countries have management objectives and policies with direct or indirect impacts on biodiversity, and reporting on progress towards these is a major role for related impact indicators. The *Biodiversity Indicators Partnership* has developed a set of indicators for assessing biodiversity status²⁸. These can be used as related targets for assessing progress towards desired impacts of LAB CEPA programmes.

However, a common problem is that local policies often lack clearly stated objectives, explicit targets or specified mechanisms for measuring progress. As a result, quantifying the indicator is not always straightforward. Different indicators may well be needed for decision-making on objectives and actions. For example, changes in the *Living Planet Index* (LPI) are an indicator of overall biodiversity loss or gain and this information is important for raising public and policy makers' awareness of the issue, but the index value alone does not explain why there is biodiversity loss or gain, or what responses are required.

Probing Questions to Ask during this Step:

- What are the existing biodiversity-relevant management objectives and targets in our city and country? Are these realistic? Adequate?
- What is the size / scope of the problem we are trying to address?
- What is the size / scope of the benefits we are trying to preserve or optimise (e.g. ecosystem services like quantity and quality of water, beauty of natural areas, tourism and recreational value, sustainability of marine and sea fisheries resources)
- Who wants to know about progress in reaching these objectives and targets?
- What outcomes and impacts of our CEPA programmes are we hoping to achieve related to these objectives and targets?

²⁸ <http://www.bipindicators.net/indicators>

- Over what time period?
- What resources do we have to achieve this?
- What contextual factors will help or hinder us?
- What are therefore realistic targets for our CEPA programmes?

STEP 7: DATA COLLECTION

At its heart, evaluation is about obtaining information and making sense of it against our chosen framework. Once we have chosen our indicators, or more likely, while we are choosing our indicators, we identify what information we will need in order to assess each of the chosen indicators. Data collection methods could include, but are not limited to: questionnaire-based surveys, focus group discussions or one-on-one interviews, and observations of CEPA programmes in action.

When weighing up potential data collection methods, consider the following:

- *practicality*
- *potential sources*
- *when* to collect data and
- the *tools/instruments* which you will need to develop, or find.

Consider how the necessary information can be efficiently and realistically gathered. When it seems impossible to gather the necessary evidence, we may need to go back to the indicator development step and find another indicator that will be easier to evaluate.

Illustration of the Need to *Plan* for Data Collection

In the City of Cape Town's Green Audits for Schools, the evaluation team ran into trouble when the evaluation was due, which happened to be at the end of the school year. They had great difficulty to reach their intended data sources, namely students and teachers, at this time of the year. Teachers were too busy marking exam papers to grant interviews. Some students were studying too hard to complete questionnaires. Other students had already finished exams and were on holiday! The evaluation team also had trouble completing resource use audits, because many of the schools' metres for water and electricity usage were in inaccessible places. To make matters worse, the team wanted to measure attitude change among the programme participants (their indicator) but found they did not have a good measure for this.

Data collection needs to begin as soon as possible, to identify and iron out difficulties early on, to establish a habit of monitoring, and to make sure one does not run out of time or data sources later. Evaluations should utilise existing opportunities for collecting data, but efforts should also be made to collect new data in innovative ways. The indicators with the greatest impact are often produced by using and presenting data in novel ways, including combining different kinds of data in ways that may not seem immediately obvious.

Building CEPA practitioners' capacity in data collection for a variety of CEPA indicators should be encouraged.

We often make use of *triangulation*, that is, the use of two or more data sources and / or

data collection methods, to measure the same outcomes. Two independent measures that ‘triangulate,’ or point to the same result, are mutually complementary and strengthen the case that change occurred.

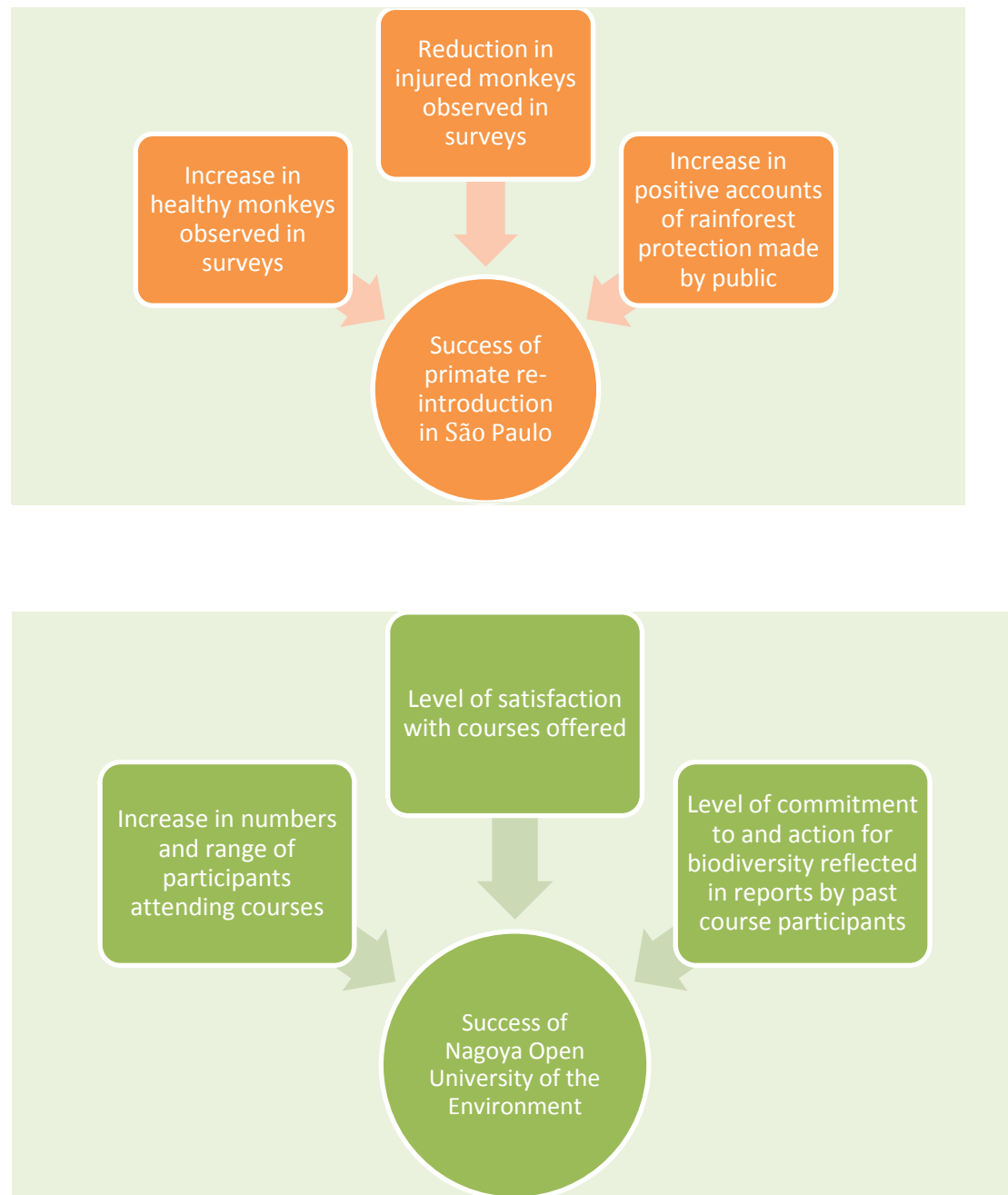


Figure 11: Two Examples of Triangulation

If one uses an experimental design for an evaluation, the standard way to account for change is to measure levels of the indicator(s) in which one is interested, both before and after a CEPA intervention. This is referred to as pre/post intervention testing (see Figure 3).

Any techniques used to make claims about change that do not rely on pre/post testing must instead rely on reconstruction, in which subjects make claims about ‘the way things used to be’. Often, these claims tend to remain unsubstantiated.

Questions to Ask during this Step:

- Are there suitable data sources for each of the possible indicators?
- Can existing data be transformed into appropriate indicators?
- How well does the available data relate to the key questions and possible indicators? (If it doesn't relate particularly well, consider triangulation with additional data sources.)
- Are the necessary agreements in place to allow data to be collected and used?
- Is there clear institutional responsibility for the continued production and reporting of the data?
- Who would be responsible for obtaining this data?
- Who will be responsible for collating and analysing this data?
- Is the data accessible and likely to continue to be produced in the future?
- Is there sufficient institutional technical capacity and resources to produce the data now and in the future?
- Is the data collected in a consistent and comparable manner over time?
- If an indicator is required to detect change, is the data collected with sufficient frequency?
- Is the data collection method appropriate to give the desired sensitivity to change?
- Do data collection and monitoring systems or agreements need to be strengthened?

Decide on the most Feasible Methods for Collecting Data

Table 7 lists the more common methods used for obtaining data to answer evaluation questions with both qualitative and quantitative indicators. One's choice of method will be determined by:

- What you need to find out
- The evaluation team's research paradigm or methodological framework – in an empiricist framework, qualitative data sources are often not highly valued or wisely used
- The kinds of data sources that are available (for example, documents or people)
- Available budget, staffing and time and associated constraints
- Possible barriers such as language, distances to travel, etc.

Also consider a suite of methods which complement each other. Each method has strengths and limitations, and often a variety of methods strengthens an evaluation.

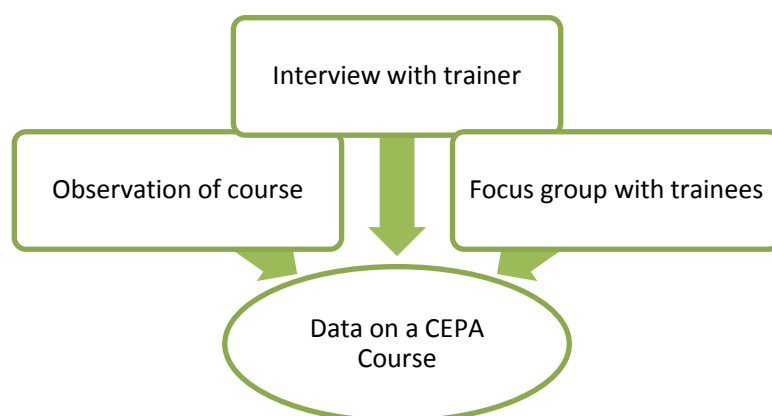
Table 7: Methods for Generating Evaluation Data

METHODS	EXAMPLES	LIMITATIONS	STRENGTHS
Observations	Observing a trainer conducting a CEPA course, or observing volunteers rehabilitating a degraded wetland.	It can be difficult to interpret what you see. For example, are the learners learning through fun, or are they distracted? Are the volunteers taking a break or unsure of how to proceed?	One can see what actually happens, rather than rely on reports of what happens.
Workshops & focus groups	Workshops with teachers to find out how a teaching resource for schools can be improved; focus group discussions with volunteers, on their wetland rehabilitation strategy.	It can be difficult to focus these meetings as they generate a lot of information, which must be accurately and adequately recorded before analysing or interpreting it.	Participants know what you're after and can assist you in finding answers to the evaluation questions; a joint exploration. Particularly useful in participatory evaluations where members seek answers together.
Questionnaires	Questionnaires to trainers and participants in the Nagoya Open Environmental University Programme, to find out their views on the courses offered.	People are often reluctant to complete questionnaires. They may fear to offend other parties. Different respondents may interpret questions in different ways, and the information obtained can be limited and hard to interpret.	Questionnaires can reach a large number of people quickly and if questions are well designed, they can produce a fair amount of information. Closed questions are easier to collate and can be analysed quantitatively.
Interviews	Interviews with individual stewardship volunteers, to find out their views and theories about their stewardship practice.	More time-consuming than questionnaires and harder to collate and analyse across interviewees. The one on one situation can encourage interviewees to simply say what they think you want to hear.	The interviewer has a chance to build a relationship, explain questions, and check their interpretation of the answers.
Tests	To check what trainees have learnt during training; a multiple choice test could be combined with a demonstration, for trainees to show what they have learnt e.g. about wetland rehabilitation.	Tests are often intimidating. It takes time to design them well. They usually test only factual recall.	One can check for specific existing knowledge on specific topics, so tests are useful for planning new activities which address areas of limited knowledge or misunderstandings.
Activities	An activity on tending a biodiversity	Activities take careful planning and can be time-	Activities are usually not as intimidating as tests

	garden with learners in the Green Schools Audit Programme, to teach them something while finding out what they have already learnt.	consuming. They should be designed so as to ascertain more than mere recall.	and can be part of the learning, while evaluating the learning.
Document Analysis	Analysis of visitor numbers recorded in staff reports; review of Strategy documents to find evaluation criteria.	The information is only as good as those who compiled the document; the original purpose and contexts of the document may limit its value if your purposes are different.	Often a quick way to access a lot of information, including historical facts which people may have forgotten. Useful for establishing trends and contextual profiles/overviews.
Participatory Appraisals	Transect walks with villagers, stopping every 100 metres to appraise the surroundings, factors affecting forest species and possible solutions.	Participatory appraisals may set up ‘artificial’ situations, or create unrealistic expectations of changes in local conditions. Strong individuals speaking on behalf of others in the ‘community’ may misrepresent others’ views.	A wide range of people is given a chance to have their say, in a non-threatening setting. More formal consultations are often experienced as intimidating.

This table has been adapted for this toolkit from its original source, *Into Evaluation: A Start-Up Toolkit*, www.capetown.gov.za, where it is listed as Tool 6. Tool 6 also has introductory information on sampling procedures, case studies and surveys, and different types of data.

Figure 12: Multiple Data Sources and Methods can strengthen an Evaluation



Analysing Quantitative and Qualitative data

Data is collected in either a quantitative (i.e. numerical) or qualitative form. For analysing quantitative data there are standard statistical procedures. We do not discuss them here, but take note of important considerations about sample size, and use a good primer on quantitative analysis and the use of statistics in the social sciences.

If you are not experienced in analysing qualitative data (data that is not numerical in nature, such as comments and general observations), obtain a good text book or guide on the topic. One general process is to: read through all the data, organise it into similar categories, e.g. concerns, suggestions, strengths, etc.; label the categories or themes; then identify patterns, or associations and causal relationships in the themes. Consider developing in-depth case studies and narratives (such as most significant change stories) with qualitative data.

Most comprehensive evaluations combine the two types of data well. Qualitative data can help you interpret the patterns and trends you observe in your quantitative analysis; quantitative analyses in turn bring perspective to the details of qualitative studies.

The level and scope of information in the evaluation report depends on its purpose and intended users and readers. A vital part of the use of the data, beyond reports to funders and senior management, is thinking through how you will apply what you learn from this evaluation phase into the next round of programme development - the learning that comes through looking at what worked and what didn't.

It is very important to question evaluation data for *double-loop learning*, for example:

- Is the data reflecting changes at source or only in symptoms?
- Does the data indicate some deeper change that needs to be made?

For each of the key evaluation questions and indicators you have chosen, indicate the methods and sources of data to provide the answers to these questions. Add a column with the names of responsible parties, a time frame for when the data should be collected and analysed, and any resources or special arrangements that would be required.

STEP 8: COMPLETE AN EVALUATION PLAN

At this stage of the process of evaluation design, you will have generated a number of evaluation questions and associated indicators, and you would have identified data sources and data collection methods with which to answer these questions. All that remains now is to put everything together in a format that shows the relationships between the various elements, and allows you to make the necessary arrangements about time and resources, in order to execute the evaluation. Such a tool also serves to communicate the evaluation plan to various stakeholders and role players, e.g. funders, managers, CEPA practitioners and evaluation team members. It is particularly useful for keeping track of the evaluation process, not only to ensure that everything happens when it should, but also to remind everyone what the purpose of the various evaluation activities and data sets are. This is easily forgotten in the hurly-burly of on-going, developmental evaluations!

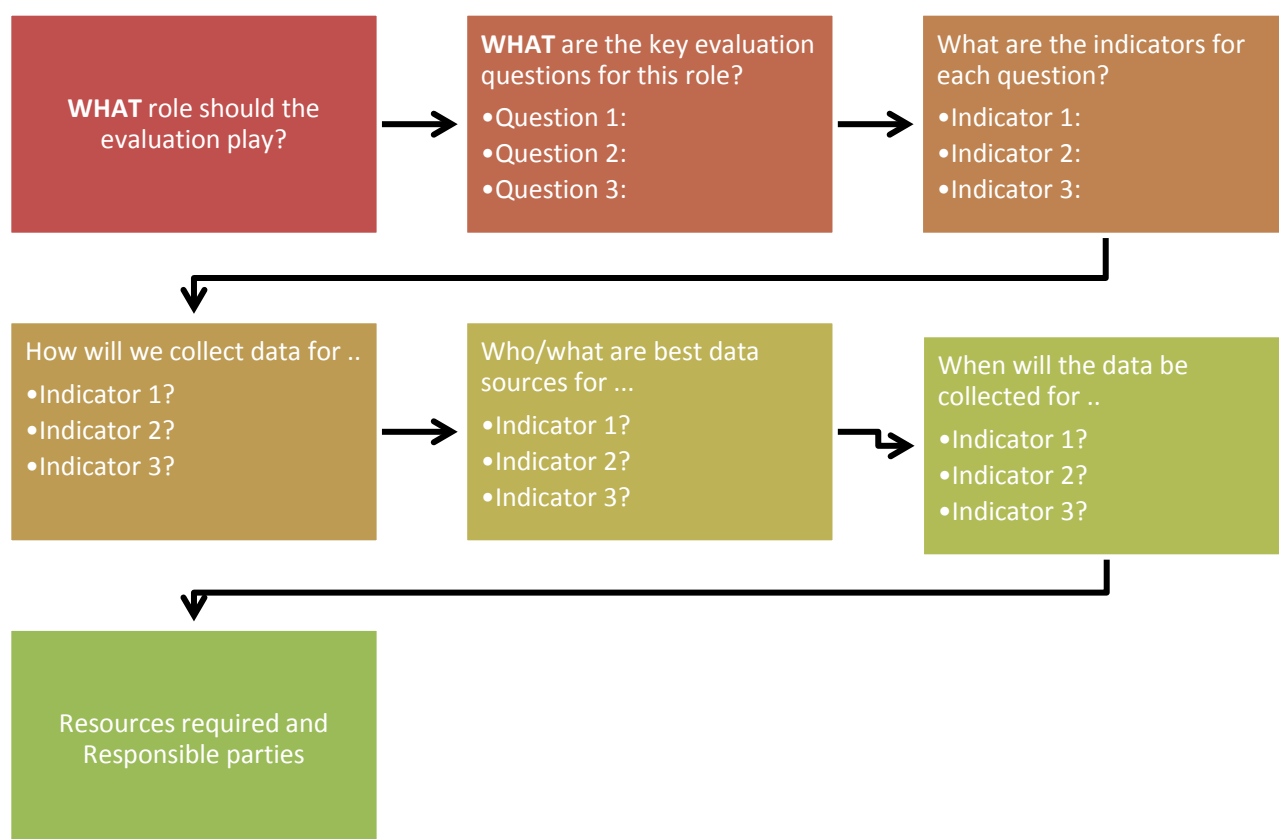
Evaluation teams often use Excel spread sheets or other software to capture their evaluation plans, in which case data could be added straight into the spread sheet. For illustrative purposes we provide a simple evaluation planning table below, that may work just as well.

Table 8 is adapted from *Into Evaluation*, the first evaluation toolkit for environmental education produced by the City of Cape Town (www.capetown.gov.za). That resource focused on working out the answers to and relationships between the first 4 of these questions in particular. This follow-up toolkit focusses in particular on Questions 5-7.

Note that at this stage, if the number of evaluation questions and indicators seem unrealistic for the available time and resources, a discussion can now be held to trim down and focus on the *most important* aspects.

Complete an evaluation plan, using the template in Table 8, or another of your choice. Make the necessary adjustments to ensure that (a) your plan is realistic, (b) you are asking the most important questions and (c), you have the indicators, the data sources and the means to answer each of these questions. Then, go ahead and do the evaluation!

Table 8: An Evaluation Planning Tool for Capturing all Basic Information



STEP 9: REPORT AND COMMUNICATE EVALUATION RESULTS

Once the evaluation is underway, collect and communicate findings to help inform and shape the CEPA programme being evaluated. In the process, also gather insights to help you improve and adjust the evaluation. Give special attention to communicating evaluation findings.

Indicators are evaluation tools and steering tools but also important communication tools. To optimise their communication value, one needs to invest time and effort in presenting and explaining indicators appropriately for their intended audience(s). Hence the skills needed for indicator development lie not solely in technical areas, but also in communication and writing. Being clear about the key questions for the evaluations is one way of ensuring that indicators are selected and communicated in a form that aids their interpretation.

Present Indicators in a Hierarchy

Indicators can be aggregated and presented in a hierarchical information system of increasing scale and decreasing specificity. One interesting way we have been exploring for presenting indicator data is in a form similar to a hypertext page. The main ‘cockpit’ shows the most critical and aggregated indicators relating to the questions of highest priority (as defined by stakeholders). A ‘click’ on that indicator opens a more detailed set of information that has contributed to the aggregate indicator. Another ‘click’ could open boxes of further information including illustrative case studies. Further ‘clicks’ could give even more specific details, such as the data sources or explanations about how the indicators have been derived. Evaluations are useful to multiple stakeholders if the entire information system is accessible to users.²⁹

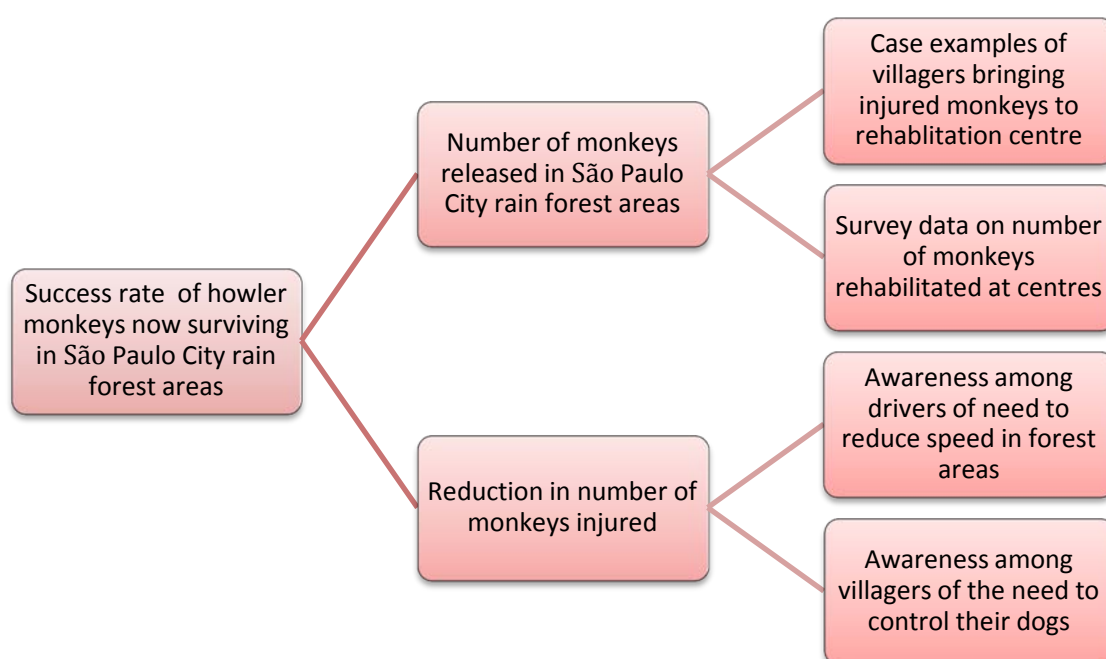


Figure 13: A Hierarchy of Possible Indicators for the São Paulo Case Study

Communicate indicators in terms of a story

It is often necessary to simplify information in order to convey useful messages to a wide audience. The art in communicating indicators is to simplify without losing credibility.

To achieve this, the overall communication of indicators can be in the form of a ‘story’ or narrative about the subject, in response to the key question(s). The narrative surrounding an indicator (set) is essential, as indicators by themselves provide only a partial understanding (hence ‘indication’) of an issue. *They always need some analysis and interpretation of why they are changing and how those changes relate to the system or issue as a whole.* Additional information allows the reader to put the indicator in context and see how it

²⁹ Meadows, Donella, 1998. *Indicators and Information Systems for Sustainable Development*. The Sustainability Institute, Vermont.

relates to other issues and areas. Information to support and explain the indicator should therefore be collected as the indicator is developed. The selection and creation of indicators should consider how they can detail and communicate the 'story'. It is also important to remember that a single indicator cannot tell us all we want to know.

Questions to ask during this step:

- How will the indicator be used?
- Who are the target audience(s) that will be using the indicator?
- Why are they being targeted? What do we want to achieve through communicating with them?
- What are the key questions that these users (may) have about the issue?
- What medium will be used to communicate? Will there be a printed report, a document on a website, a static or interactive web-page, video footage on national TV, a workshop or site visit with stakeholders, a Power Point presentation, a newspaper article, a radio interview, or a combination of some of these?