Evaluating the impacts of conservation interventions on human wellbeing

Guidance for practitioners

Emily Woodhouse, Emiel de Lange, E.J. Milner-Gulland

With: Emilie Beauchamp, Heather Gurd, Katherine Homewood, Charlotte Mathiassen, Ben Palmer Fry, Dilys Roe, Helen Schneider, Rebecca Singleton

About the authors

Emily Woodhouse* is a lecturer in the Anthropology Department of University College London.

Emiel de Lange is an MRes student in Tropical Forest Ecology at Imperial College London.

E.J. Milner-Gulland is Tasso Leventis Professor of Biodiversity at Oxford University, and leads the Interdisciplinary Centre for Conservation Science.

*Corresponding author: Emily Woodhouse **e.woodhouse@ucl.ac.uk**

Acknowledgements

Many thanks to participants of the workshop, 'Methods for evaluating the human wellbeing impacts of nature conservation interventions: a workshop for practitioners', held in London in June 2015: Rafael Chiaravalloti, Tom Clements, Victoria Dauncey, Georgina Domberger, Phil Franks, Charlotte Gough, Moise Kono Bidzo, J. Terrence McCabe, Simon Mercer, Danni Parks, Maria Belinda E. de la Paz, David Thomas, David Wilkie, Alexandra Zimmerman. We gratefully acknowledge funding from the ESRC-DFID joint fund for poverty alleviation (ES/J018155/1) 'Measuring complex outcomes of environment and development interventions' and an ESRC Impact Acceleration award to Imperial College London.

Published by IIED, September 2016

IIED promotes sustainable development, linking local priorities to global challenges. We support some of the world's most vulnerable people to strengthen their voice in decision making.

http://pubs.iied.org/14667IIED
ISBN: 978-1-78431-375-3
Citation: Woodhouse, E *et al.* (2016) Evaluating the impacts of conservation interventions on human wellbeing. Guidance for practitioners. IIED, London

iied

International Institute for Environment and Development
80-86 Gray's Inn Road, London WC1X 8NH, UK
Tel: +44 (0)20 3463 7399

www.iied.org
@iied
www.facebook.com/theIIED

Download more publications at **http://pubs.iied.org**

Summary

Conservationists are increasingly seeing the importance of carrying out social impact evaluation to ensure accountability and to learn what works for both biodiversity and human wellbeing. A single toolkit or blueprint method cannot fit the diversity of intervention types and evaluation questions, and conservationists are faced with an array of decisions about the most appropriate methods and research designs to use. This guidance aims to demystify the process of social impact evaluation and support practitioners in navigating through these methodological decisions, taking into account: the questions the evaluation aims to answer; the characteristics of the intervention; and the organisational capacity and resources available. It takes practitioners through the key steps in an evaluation: 1) thinking through the aims of the evaluation; 2) defining relevant wellbeing outcomes and indicators; 3) designing the evaluation to link outcomes to the intervention; and 4) collecting data, including applying methods to account for bias, social dynamics and ethical considerations. The guidance provides a range of real life case studies and ideas for appropriate methods and tools.

Contents

List of boxes, figures and tables

1 Introduction

1.1 Background and purpose

Conservationists recognise that for many reasons the social impacts of their work are important to understand, and that interventions need to be designed for accountability, learning, and to build an evidence base. But social impact evaluation is not widely used in the conservation world, and practitioners still often feel they lack the capacity to rigorously monitor and evaluate their projects. Instead monitoring is commonly done in a snapshot way, or through the perceptions of the implementer, meaning it is difficult to link social changes in the intervention area back to the intervention in a convincing way. By contrast, the focus in the academic literature has generally been on technically demanding statistical designs that are unrealistic for the majority of interventions and may not answer all the questions that conservationists are interested in.

Although there are some specific tools and methods available for social impact evaluation, there is a lack of practical guidance that is applicable to conservation practitioners, helping them to select and apply these methods. There can be no one methodological toolkit that fits every intervention; so instead of prescribing particular tools and methods, this guidance document aims to help practitioners to navigate through the multitude of decisions they are faced with, taking into account the particular circumstances of their intervention and the purpose of their evaluation, in order to produce useful and credible knowledge about impacts that is fit for purpose.

This guidance was developed from discussions between conservation practitioners, funders and academics during a workshop held in June 2015 in the Anthropology Department of University College London, and funded by the Economic and Social Research Council and the UK Department for International Development.

Who should use this guidance?

- Those managing or carrying out impact evaluations, directly or indirectly, such as non-governmental organisation (NGO) field managers, and monitoring and evaluation advisers
- Those requesting impact evaluations and assessing the results, such as conservation donors
- Those drawing on the evidence generated from evaluations and steering the social impact agenda, such as policymakers.

What this guidance aims to provide

- Understanding of the steps in social impact evaluation and key issues to consider
- Understanding of how decisions that evaluators make depend on the questions the evaluation aims to answer, the characteristics of the intervention, and the organisational context
- Signposts to appropriate methods and tools
- Case studies drawn from real life examples to illustrate a range of situations that practitioners could find themselves in, and potential approaches to take under different circumstances.

2 The components of social impact evaluation

The primary aim of social impact evaluation is to assess what social outcomes (changes in aspects of human wellbeing) have resulted from an intervention. These changes may be intended or unintended, positive or negative, short or long term. It is important here to differentiate between outcomes, outputs and impacts.

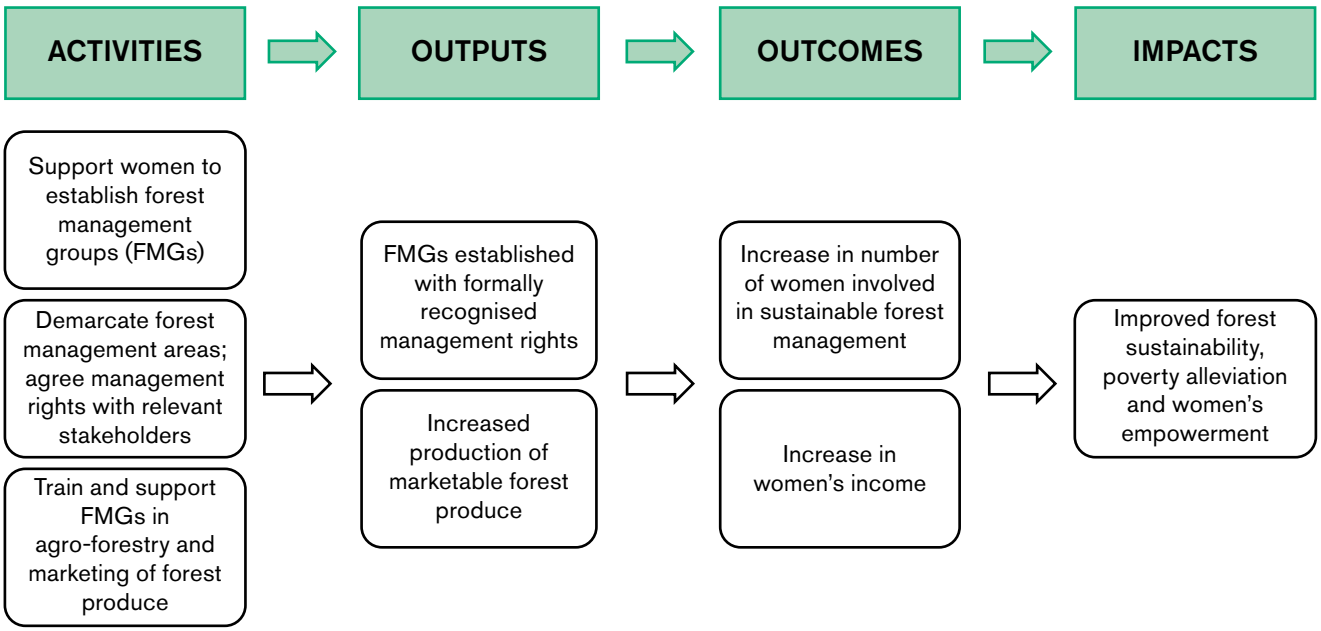
Outputs are the direct and specific products of the activities of an intervention. They provide the conditions necessary to achieve the outcome. For example, households may be producing honey from beehives supplied by a project.

Outcomes are the changes or effects that happen as a result of the intervention. Outcomes are often finite and measurable changes. For example, households may see an increase in their annual income as a result of selling their honey.

Impacts tend to refer to the wider and longer-term effects of an intervention and the contribution the intervention makes to broader goals. The income from the honey may allow children in the household to get an education and thereby move out of poverty.

Evaluations may focus on measuring and understanding changes at different levels of the intervention process (Figure 1). Evaluations always need to measure outputs, but evaluators may decide to stop at this point due to lack of resources or short time frames, or donors may only require reporting at that level. In the longer term, outcomes and impacts need to be evaluated, however, if implementers are to show how their work has led to societal change.

Figure 1. Stages of an intervention where evaluation may take place



Note: Example taken from a community forestry programme focused on women's engagement

The tricky part of an impact evaluation is not so much measuring change in social indicators (such as income, educational attainment or nutritional status) but in **attributing** that change to the intervention. All interventions take place in a social and economic context, operating at a range of scales. For example, just because incomes improve in an area, it is not necessarily possible to claim that this is because of the intervention; perhaps the whole economy has improved and incomes would have increased anyway.

Alongside the focus on attribution, there is increasing emphasis in international development on explanation — the *how* and *why* of impacts. This aspect of evaluation is needed for lesson learning within and between programmes and organisations. For example, we often want to know not just, "Did a community forestry programme improve wellbeing?" but **how** did a community forestry programme improve particular components of wellbeing, such as livelihood and empowerment indicators. Or why did a livelihoods programme not produce the income impacts expected? Why did an environmental education project work in one location but not another? The outcomes may be due to the programme itself, the way it was implemented, or contextual issues.

We will cover four stages of thinking about, and then implementing, a social impact evaluation in Sections 2, 3 and 4:

- Thinking about the aim and purpose of the evaluation (Section 2)
- Defining relevant outcomes, and then indicators which can represent these outcomes (Section 2)
- Designing an evaluation which can link outcomes to the intervention (Section 3)
- Collecting the data required for the evaluation (Section 4).

Below we set out the decisions that must be made within each of these steps, but we start by explaining what we mean by human wellbeing. We do this because the concept of 'wellbeing' is useful as a broad, common framework for understanding the social impacts of interventions.

2.1 What is human wellbeing?

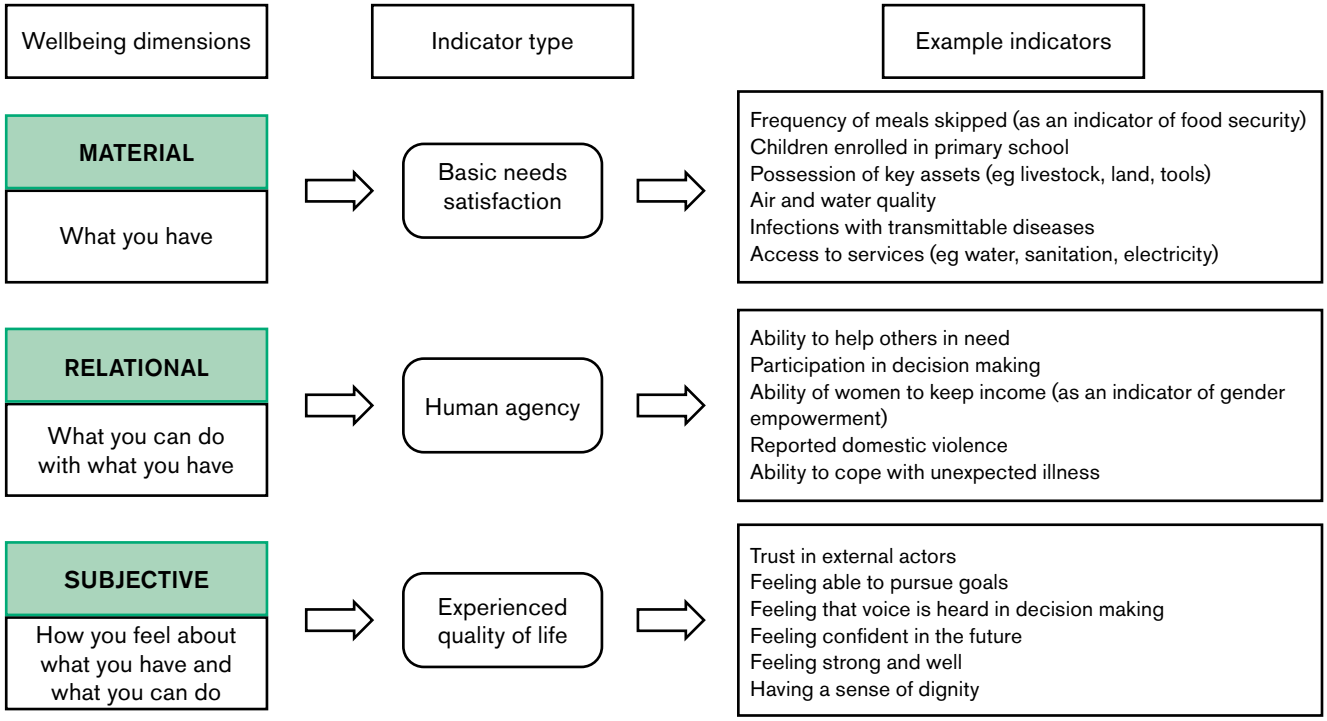
There is an increasing convergence in international policy on the concept of wellbeing as representing a positive physical, social, and mental state (Stiglitz *et al.*, 2009; Summers *et al.*, 2012). Wellbeing is conceptualised in three interacting dimensions (McGregor and Sumner, 2010):

- **Objective, material** circumstances of a person's life and the extent to which their needs are met — for example housing, income, livelihoods, health and the environment
- **Relational** elements, focusing on how people engage with others to meet their needs and achieve goals, acknowledging that wellbeing is not just about what a person has, but what they can do and be. For example, community networks, social institutions and empowerment
- **A subjective** evaluation of one's own life and the meanings and values ascribed to the processes one engages in and the outcomes of those processes. How people feel and their satisfaction with different aspects of their life may be especially relevant for conservation practitioners aiming to gain support from local populations.

In conceptualising wellbeing, there is a need for a universal approach that allows comparisons, while ensuring local relevance. This three-dimensional framework provides a structure with comparable categories, which can be adapted to the local context using locally meaningful specific indicators. The framework (Figure 2) can guide the scope of a wellbeing evaluation, suggesting the types of data that need to be collected (for example objective livelihood data combined with subjective assessment).

This framework is not the only way to sub-divide wellbeing. For example, the empirical research of the World Bank's 'Voices of the poor' project found components that were commonly considered to constitute wellbeing among individuals across 23 countries (Narayan *et al*., 2000). They divided these into: material assets, health, social relations, security, and freedom of choice and action. These provide a useful checklist of themes to consider when starting an evaluation, and illustrative indicators in Figure 2 span these categories. The three dimensions help to define the questions asked and the type of data collected.

Figure 2. Framework for researching human wellbeing



Source: Based on Britton amd Coulthard (2013) and McGregor and Sumner (2010), and drawing on the World Bank's 'Voices of the Poor' research (Narayan *et al*., 2000)

Whichever wellbeing framework is used, its role is to ensure that those designing an intervention in the first place, and then those wanting to evaluate its social impacts and understand how to improve it, think about a broad range of impacts (not just income, for example). And that they think about the individual whose wellbeing might have been affected by an intervention within a social context, and not in isolation.

2.2 Why do you want to do this evaluation?

The first question to address before embarking on an impact evaluation is: why do you want to do it?

A fundamental question is whether you want to **learn internally** within your organisation, or **demonstrate outcomes externally**, for example to funders or to support learning by others. Internal learning is useful for organisations looking to adaptively manage programmes and improve their practice more generally, while external demonstrations may be required to secure funding or support. Often an evaluation will aim to do both.

The next question is whether you are more interested in understanding **how much** change has been produced by your intervention (such as the actual dollar improvement in incomes), **whether** your intervention has had a positive or negative effect; or **why** impacts have occurred. Measurement of effect sizes (that is, how much change) would be useful for the purposes of external demonstrations, as they can provide evidence of the effectiveness of an intervention. Whether the change has occurred, and that the intervention has **contributed** to this change, may be all you need to know, however. Elucidating mechanisms (why the change occurred) can allow lessons to be learnt and successful interventions to be replicated.

Increasingly, projects have a **'theory of change'**, which explains how the project designers assume that a given intervention will lead to the impact that they expect. Some evaluations will want to track progress along this presumed causal chain, from inputs to outputs, outcomes and then impact. But in some cases, it may be enough just to show progress over time towards the long-term goal of the project, without the need to explicitly demonstrate how that progress is being made.

Are you interested in understanding the impacts for this one intervention only, or would you like to know if you can expect it to work elsewhere (the **generalisability** of the approach)? Your organisation may be only interested in the impacts of a specific project in order to report to donors, for instance, or may be testing an innovative approach that it would like to roll out in other areas.

These different needs and priorities will drive all other decisions. We will return to this in Section 5.

2.3 Defining outcomes and indicators

Deciding which outcomes are of interest

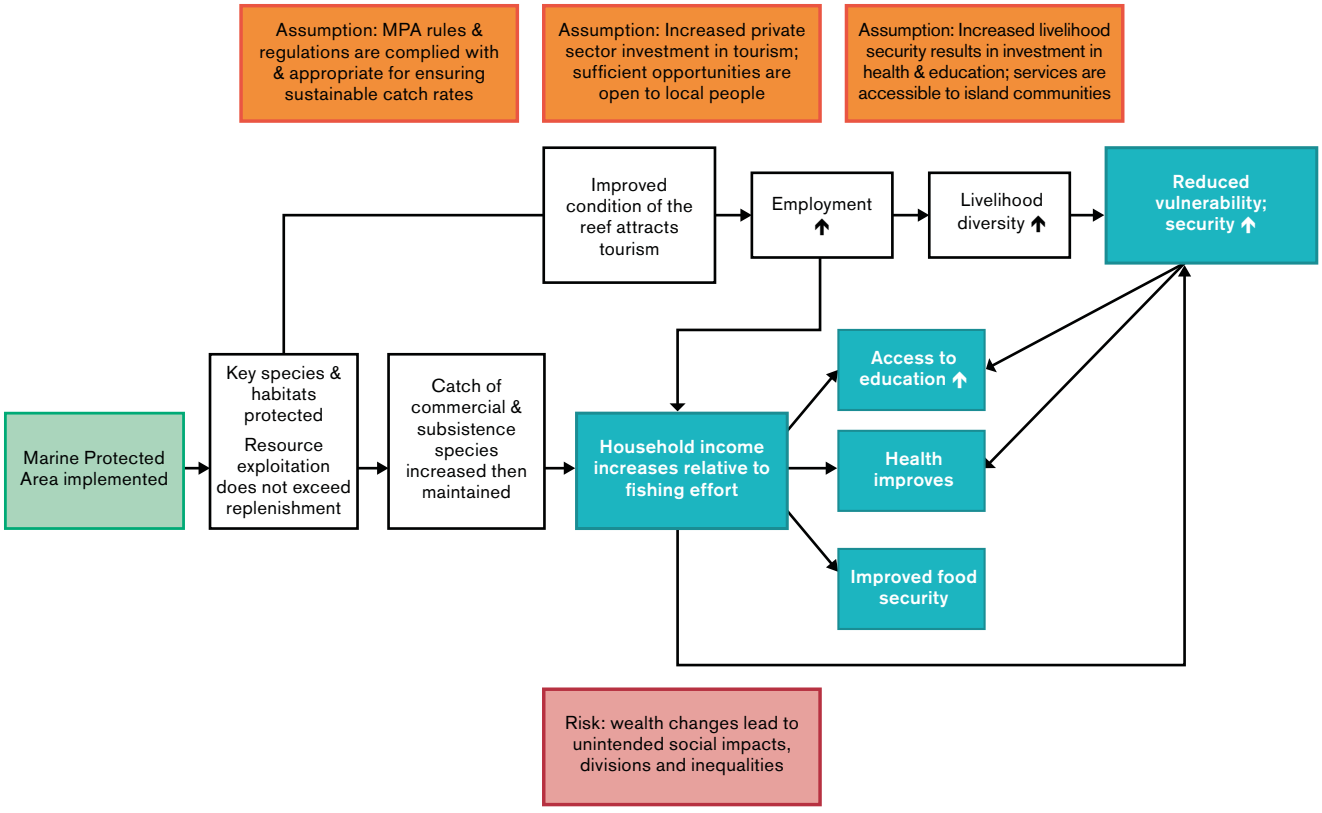
It is important to be clear about what is being evaluated. Positive social outcomes may already be a predefined goal of the project; for example improving levels of women's empowerment through establishing natural resource management community groups, or improving health through watershed protection. In this case, it may make sense to focus the evaluation on whether these hoped-for outcomes have been achieved. For interventions that do not have explicit social aims built into the project goals, but rather only focus on biodiversity and the environment, evaluators of social impact will still need to consider the possible positive and negative impacts on wellbeing of the intervention. The project documentation may then have less guidance on which outcomes should be focused on in the evaluation. In practice, however, the differences in how

impact evaluations are carried out between projects that have an explicit goal to evaluate and those that do not may actually be minimal, because all evaluations will need to consider possible unintended consequences, and therefore need to look at wellbeing change in the round.

Either way — and even if the aim of the evaluation is not to test the causal links between an intervention and its impacts — it is still very helpful to define the outcomes of interest, and the steps presumed to be leading up to them, through a **theory of change** (ToC). This is the explanation of the process by which the intervention is thought to give rise to outcomes. The ToC maps out the causal chain, the underlying logic and assumptions, and other factors which may impact wellbeing (Figure 3). The evaluation is effectively testing this theory and the assumptions on which it is based.

When developing the theory of change, it is important not to be too linear. Often, there are trade-offs between different aspects of wellbeing. For example, a tourism project may increase incomes, but reduce social cohesion. In other cases there may be synergies; for example building a community institution to manage natural resource use may also build a sense of belonging more generally. In order to identify these indirect consequences, synergies and trade-offs, local knowledge is vital. The wellbeing framework (Figure 2) can guide thinking about the different dimensions which might be affected by an intervention.

Figure 3. Theory of change for a marine protected area



Notes: MPA = marine protected area; white = expected material wellbeing outputs; blue = outcomes. Additional relational and subjective wellbeing impacts (eg autonomy and sense of purpose) may be achieved through increased participation in decision making as a result of governance interventions but are not included here.

Source: Adapted from a ToC produced by Fauna & Flora International

Defining indicators for the outcomes and outputs

Those actually impacted by the intervention should have a strong voice in what constitutes impacts, and what the relevant indicators are for the steps along the way. Indicators are specific measurable characteristics that can be used to show progress in meeting specified outputs, outcomes and impacts. An indicator should be defined in precise, unambiguous terms to give a clear idea of the data required (see examples in Table 1). Indicators should be locally relevant, and outcome-based ones should cover all the different aspects of wellbeing (on the three dimensions in Figure 2) that are felt to be important by those impacted.

Different ways to develop indicators include using the team's pre-existing knowledge and data, and using in-depth or rapid participatory methods to develop them with local people. A good approach is to start with a set of indicators based on existing knowledge, but then to refine them in collaboration with local people.

An indicator needs to be simple and specific, or data cannot be collected to demonstrate how it has changed. But it is important to be aware that indicator creation is a value judgement, and information and nuance is inevitably lost in the process. This is particularly true when developing quantitative indicators (ranks, scores, percentages) to represent complex constructs, for example 'sense of empowerment' or 'social cohesion'.

It is important not just to use indicators because they are easy to collect or because the data already exist, or other people have used them in the past. Think about construct validity; does the indicator actually measure what it claims to measure, and does it really relate to the outcome and impact that you are interested in? If not, it may be better to have a more ambiguous qualitative answer than to persist with using an indicator that may be misleading, for the sake of generating a number. For example, if an intervention claims to improve local participation in resource governance processes, a common indicator is attendance at community meetings — but this does not capture whether people had a voice in the meetings or whether their contributions were taken into account in decision making, for which more qualitative data are needed.

Table 1. Quantitative indicators and data sources for wellbeing in a pastoralist region, Tanzania

Wellbeing dimension	Outcome	Indicator	Data type and source
Material	Wealth (and security, status)	Tropical livestock units (standardised measure)	Numeric; household survey
	Wealth	Type of house	Categorical (brick or traditional); observed
	Food security	Agricultural productivity (bags of produce / land planted)	Numeric; household survey
Relational	Participation in development processes	Number of community meetings attended in the last year	Numeric; recall data in individual survey
	Recognition	Feeling that your voice is heard in development processes	Ordinal; Likert scale 1-5; individual survey
	Female autonomy/ agency	Ability of women to keep and use income earned	Binary (yes/no); individual women's survey
Subjective	Security – confidence in the future	Feelings of security regarding access to land in the next 10 years	Ordinal; Likert scale; individual survey
	Trust in external organisations	Trust in specific listed actors, eg conservation NGOs	Ordinal; Likert scale; individual survey
	Continuation of valued cultural traditions	Feelings that beneficial traditions will be sustained in the next 10 years	Ordinal; Likert scale; individual survey

Notes: These examples of specific quantitative indicators and data sources for household and individual wellbeing (men and women), used in a study in a pastoralist region of Tanzania, were generated through and complemented with extensive qualitative inquiry and data. Note that in pastoralist areas livestock may denote aspects of several outcomes.

3 Linking changes in outputs or outcomes to the intervention

3.1 Evaluation design approaches

Research design is the underlying logic of the study that allows the evaluation results to say something about causality.

Although in conservation science the emphasis in the evaluation agenda has been on quasi-experimental designs using controls, there are a range of different research designs which use different bases for making causal inferences (Table 2). These designs have strengths and weaknesses, and allow evaluators to answer different questions (see Section 5 below and Section 2.2 above). Whichever approach is used, it is important to make explicit the basis for causal claims. In this section we go through the different design approaches, and give examples for each.

Experimental designs

Randomised controlled trials (RCTs)

When the treatment is randomised, this gives every unit (such as household or village) an equal probability of being assigned to either a treatment or control group, removing any systematic differences between those involved in the intervention and those who are not. Thus any differences in outcomes can be attributed to the intervention. Randomisation is rarely done in practice in conservation for a number of reasons such as cost, complexity and ethics, but may be more feasible than is often assumed. In development policy, randomised controlled trials are common practice. For example, from a set of villages, a random selection is given an intervention, and the others not; the evaluator then measures how the outcome indicators change in the two sets of villages. In conservation, the concern is often that there is limited scope for controls. But remember that it's the people — not the wildlife — who need to be randomised. For example, if you care about a rare monkey in a forest but your intervention is about stopping land clearance for agriculture, you just need another forest being cleared in similar ways for similar reasons; or even other parts of the same forest, if you can assume independence in terms of social interactions. These areas of forest do not have to have monkeys in them (unless increasing knowledge of the importance of the species is a key part of the conservation strategy).

Quasi-experimental designs

These designs mimic RCTs by identifying observable biases that led to the establishment of a conservation intervention in a particular place, or to certain people participating, or affected its outcomes. These biases are then used as a basis for identifying controls. Instead of randomising treatments, this involves choosing intervention villages/households for your intervention according to your priorities (for example you may want to work in villages that are within a protected area). Then you choose control villages or households to represent the 'counterfactual' (what would have happened in the absence of the intervention), based on similarity to the intervention targets, on a range of baseline variables that are thought to affect participation in the intervention and the

Table 2. Different design approaches for impact evaluation

Design type	Basis for causal inference	Requirements	**Examples** (further details below)
Experimental	'Counterfactual': comparing the change in the outcome variable in the intervention site with what would have happened in the absence of the intervention	Valid control with or without baseline	▪ Randomised controlled trials ▪ Quasi-experimental designs (Before-After-Control-Intervention)
Statistical	Correlations between the outcome indicator and the input indicator, controlling for confounding factors	A large sample size, comparison groups or longitudinal data, and data on confounding factors	▪ Statistical regressions
Theory-based	Identifying the mechanisms that explain changes in the outcome variables, and providing empirical evidence	Strong theory of change (pre-existing or developed)	▪ Process tracing
Case-based	Comparison across and within cases of the outcome of interventions under a combination of presumed causal factors	Strong theory Several different cases are needed for comparison	▪ Qualitative comparative analysis
Participatory	Perceived causation from the point of view of people affected by the intervention	Skilled facilitators	▪ Reflexive counterfactuals ▪ Ranking and scoring ▪ Most significant change (narrative)

Source: Based on Stern (2015)

outcomes. Social change in the control villages is assumed to be happening over time, in similar ways to what would have happened in the intervention villages in the absence of the intervention. The variables that you use to match the controls to the intervention targets are therefore based on an understanding of how the system is changing.

For example, perhaps the assumption is that the main factor determining change in material wealth in the region is remoteness and the amount of agricultural land available around a village. In this case, villages might be matched on distance to city and land cover. Or within a village, you might assume that wealth (among other things) affects whether households are able to participate in the intervention; so matched controls would be other households which were similar in material wealth to your participant households before the intervention took place. The choice of controls can be made statistically (using propensity score matching, for example), or qualitatively.

Collecting baseline information on the indicators before the intervention starts, from both the control and intervention villages, as well as information on the indicators after the intervention, produces a Before-After Control-Intervention (BACI) design. This allows evaluators to isolate effects of the intervention from concurrent changes that are to do with external factors (for an overview, see Gertler *et al*., 2011). These approaches allow firm attribution as long as they are well designed (see Example 1), but they have stringent requirements, and in particular they cannot be used for evaluation after the fact; baseline data are needed for treatment and control sites.

Example 1. Quasi-experimental design

In the Northern Plains of Cambodia, two protected areas (PAs) were supported by the Wildlife Conservation Society from 2005, with the aim of protecting threatened species and limiting deforestation and agri-industrial development. Additionally, three payment for environmental services (PES) schemes were established in some villages within the PAs from 2008 onwards. Clements and Milner-Gulland (2015) used a Before-After Control-Intervention (BACI) design to evaluate the impacts of these interventions on deforestation and human wellbeing. Four measures were used to capture components of wellbeing considered locally important and potentially impacted by the interventions: 1) poverty as measured by the Basic Necessities Survey (BNS), an index-based tool in which necessities are locally defined through a participatory process and scored at the household level (Davies and Smith, 1998); 2) agricultural productivity of rice, as the staple crop in the Cambodian diet as well as being socially and culturally significant; 3) food security, measured as the difference between a household's annual harvest and its rice need for subsistence; and 4) education levels of each household member.

To measure the difference in deforestation rates, PA villages were statistically matched with control villages outside the PA (based on remoteness from roads and markets, village population, and forest cover). For the social impacts, however, it was not possible to match the households statistically, and instead difference-in-difference estimators (which measure rate of change over time) were used to compare household wellbeing within the PA and control villages, and between households who did and did not participate in the PES schemes. However, given that the processes affecting the control and treatment villages and households could be assumed to be the same, this still gave a robust counterfactual. Social data were collected through surveys of 871 households across 20 villages in 2008, and repeated for the same households in 2011. Data were disaggregated according to household livelihoods to elucidate the variable impacts of the intervention on different sectors of society.

Overall, the status of households across the area improved over the study period. Average rates of change in ownership of basic necessities were similar between households in the PA and control villages, but there were significant differences depending on the main livelihood strategies households followed; for resin tappers, and also for households who had less than one hectare of rice paddy (making them among the least food secure members of the community), poverty improvements were faster in the PAs compared to the controls.

Within PA villages, the two higher-paying PES programmes improved BNS scores at a greater rate than in non-PES households, and these households were also able to educate

their children to higher levels. These programmes are the Ibis Rice programme, which pays a premium price for rice to households following participatory land-use plan rules, and a bird-watching ecotourism programme. The Ibis Rice programme also increased rice harvests and improved food security, but was largely limited to those who had sufficient land to produce a rice surplus over their subsistence needs, suggesting elite capture. More educated and well-connected people who could afford to divert labour away from agriculture were more likely to be involved in the ecotourism project. Involvement in the third scheme, which pays individuals for protecting nests of threatened bird species, had no added impact on income because the payments were small; but it was the only intervention which was egalitarian in terms of access. These results and insights are being used by the Wildlife Conservation Society to adapt, target, and structure the interventions for more equitable benefits in the future.

A researcher carries out a household basic necessities survey in the Northern Plains of Cambodia to assess the impacts of PES programmes on human wellbeing. Credit: Emilie Beauchamp.

Statistical designs

Statistical designs give quantitative datasets that can be analysed statistically, often using various types of regression model. This includes longitudinal studies (analysing change over time), and studies which look at differences between groups of people at the same point in time. These models do not have controls, but instead they look for associations between the outcome indicators (response variables in the regressions, for example household income) and variables which are hypothesised to cause these outcomes (such as participation in the intervention; or distance from a protected area — see Example 2 below), taking into account confounding factors such as household size or prior wealth.

Correlation is not causation, so these methods cannot provide firm attribution. For example, just because households on average show improved wellbeing over time, it is not possible to infer that it was because of the intervention; it could be that wellbeing was improving anyway. If there is no information on change over time, but only current status, then the inferences are weaker still; participants in the intervention may be better off not because they were part of the intervention, but because only better off households were able to take part. However, very plausible attributions can be made using these designs if it can be demonstrated that changes in potential confounding factors are either taken into account in the model, or did not happen.

Example 2. Statistical design

This evaluation assessed the development impacts of Tarangire National Park, Tanzania, on surrounding communities. Protected areas can be strong attractors of tourist infrastructure and NGOs concerned with conservation, religious or indigenous rights issues, all potentially acting as constraints on or opportunities for development in local communities. Baird (2014) assessed the impact of Tarangire National Park (TNP) in northern Tanzania on the development outcomes of predominantly Maasai communities living near the park. The assessment used mixed methods: qualitative stakeholder interviews focused on infrastructure development through time at a community level; and structured surveys on educational outcomes at the household level. Six communities east of TNP were selected based on geographical proximity to the park and categorised as near (4) and far (2) from the park.

A structured survey of 216 households was carried out across the communities, and indicators of education were the level of education of the household head, and the percentage of eligible children (aged between six and 15 years) enrolled in school. Regression models estimated the effect of proximity to TNP on these outcomes, controlling for other factors that may influence the outcomes, namely age of household head, household size, membership of a church, and wealth. Proximity was found to have a significant effect on school enrolment, but not on the education of the household head.

Since there were no baseline data, the plausibility of the attribution of educational outcomes to proximity to the park was strengthened through the qualitative data. Interviews with community leaders, government officials and those with knowledge of historical development interventions, confirmed that prior to the formation of TNP, few schools, water points or health clinics existed in the study area, and the distant

communities were as developed as those near the park. After 1970, when TNP was established, development near the park was supported by a range of external organisations such as hunting and tourist companies, foreign donors, and the government body overseeing national park management, TANAPA. This development was initially unsolicited, but was more recently solicited by communities as they formed connections with external organisations. In the distant communities, development was largely supported by local government, and so less progress was made. Opportunities associated with wildlife organisations, however, have presented problems for communities near TNP. There are feelings of unfairness regarding the distribution of benefits from wildlife, and local perceptions of the quality of development projects were generally negative. This suggests that although objective development indicators such as educational enrolment may have improved as a result of the park's presence, this has not necessarily reduced conflict with conservation organisations or improved subjective wellbeing.

Theory-based designs

Theory-based designs compare the results of the intervention to theoretical predictions (such as from the ToC, literature or past experience), presenting alternative hypotheses for change and discounting or confirming them with empirical evidence (quantitative and/or qualitative). For example, a local marine reserve may be expected to protect fish stocks and increase fishers' yields in the long term. Following this theory of change, fish yields would be assessed after a period of time in order to evaluate the reserve's impacts. Other reasons for improved fish yields, such as improved technology or environmental changes, need to be ruled out with empirical evidence.

One benefit of these designs is that they do not require the use of controls, and so are particularly useful where comparative groups are difficult to identify or use. This tends to make them weak on estimating the quantity or extent of impact, but strong on explaining how and why a given intervention or specific activity led to an outcome. They are especially useful where there are multiple, complex causes and effects which require a fine-grained analysis. Although there are several different specific approaches to these designs including 'process tracing' (Example 3), they all take the same general approach of developing and testing a causal model.

Example 3: Theory-based design

This example concerns Oxfam's health advocacy programme in Ghana. Process tracing focuses on the systematic analysis of evidence within a case study to decide between alternative possible explanations. Causal inference is made through the development of hypothesised causal mechanisms, and assessment of the strength of evidence for each part of the chain. The probability is not assessed quantitatively as in statistical methods, but rather has parallels to how evidence is dealt with in a law court — empirical observations are combined with knowledge of contextual factors such as timings, how the evidence has been generated, and the credibility of sources. Oxfam has been using this type of method in an effectiveness review of its programmes, and it has proved particularly

conducive to evaluating policy and advocacy work, where it is nearly impossible to identify control populations. It was used to assess the effectiveness of a campaign to promote universal free health care in Ghana through lobbying, mobilising, media and research (Stedman-Bryce, 2013). The main steps in the process can be demonstrated with the following case study (Punton and Welle, 2015):

1. **Develop a hypothesised causal mechanism**, which involves elaborating a theory of change, including all steps from the cause (the intervention or parts of it) to the outcome of interest including activities. Each part of the mechanism can therefore be framed as a hypothesis. Oxfam used the existing project ToC and revised it with the evaluator and project staff (a simplified version is shown in Figure 4).

2. **Operationalise the causal mechanism.** Each step should be empirically observable in order to determine whether it happened or not, and plausible alternative explanations for each step specified.

For **Step 1** in the chain, the hypothesis is that the campaign increased capacity for civil society organisations (CSOs) to plan and work together. Observable evidence would be examples of CSOs working together (such as events, platforms); and civil society actors attributing coordinated advocacy to the campaign. Alternative explanations for the coordinated activity would be: 1) that CSOs had already planned and worked together on free health care before the campaign; and 2) other campaigns were also promoting coordinated activities.

For **Step 2**, one specific hypothesis was that the government had revised its methodology for calculating active membership of the health insurance schedule because of pressure from the campaign. There was a dramatic drop in reported membership to 34 per cent of the population, and government therefore acknowledged the current insurance scheme as an ineffective vehicle to deliver free universal health care. The two rival explanations for the outcome were: 1) that the revision occurred as a result of campaign pressure, namely a report published by the campaign just before the government revised its methodology; 2) that the revisions were based on the government's own plans and timetable.

3. **Collect evidence for each step of the chain.** Consider reliability, limitations and biases of evidence sources which may be primary and/or secondary. The evaluation drew on 21 key informant interviews with campaign members and government representatives, alongside project documents and data including news stories and reports on social media.

4. **Assess inferential weight.** Evidence is assessed as to whether it gives a reasonable degree of confidence that the step in the mechanism does or does not exist. Ideally evaluators are looking for evidence that has high uniqueness (the evidence is sufficient for the hypothesis to be true), and/or high certainty (the evidence is necessary for the hypothesis to be true).

For **Step 1** — the hypothesis that the campaign strengthened coordination between CSOs — evidence included the fact that results of an initial report published by the campaign were presented to and discussed with members; and that the range and volume of advocacy activities increased thereafter. Both pieces of evidence increase the

plausibility of the hypothesis, without definitively ruling it in or out; they are not completely sufficient or necessary for the hypothesis to be true. Further evidence (from interviews, media checks) however, showed that there was little coordinated advocacy before the publication, which strengthens confidence in the hypothesis, and rules out the alternative hypotheses.

For **Step 2** — the hypothesis that the campaign caused the government to revise its methodology — very compelling evidence was presented which is certainly sufficient. The Ghana delegation at an international meeting on Universal Health Care stated that the campaign's report, "was very helpful and prompted us to revise our figures". It is highly unlikely that the delegation would make this statement if the report had not influenced them, particularly since they had initially dismissed the report.

This kind of 'smoking gun' evidence, which provides near-unequivocal support for a causal story, can be quite unusual. When evidence is less clear, what constitutes credible evidence is inevitably a judgement call. It can help to focus on triangulating evidence sources and methods, and validating findings with key project stakeholders. It also becomes vital to retain transparency in the process, by being clear about the sources and nature of the evidence and tests used. Of course, more than one cause (such as another intervention, economic changes, climate) may contribute to the outcome, so that judgements must also be made on the level of contribution of the intervention (rather than direct attribution). Oxfam's protocol for process tracing involves scoring the relative contribution of the intervention to each outcome based on a qualitative assessment of the strength of evidence (Oxfam, 2011).

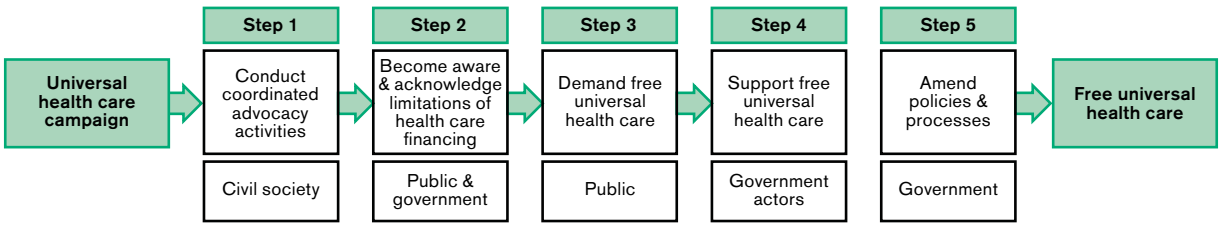


Figure 4. A simplified causal mechanism based on the universal health care campaign, Ghana

Case-based designs

Case-based designs involve comparisons across cases (such as villages) that highlight the combination of causal factors (types of projects, method of implementation, contextual factors) which could explain social impacts. Factors could be identified based on the ToC, in which case they are assumed 'a priori' or based on existing knowledge. In this case this is just a sub-set of theory-based designs. Or they could be identified in a participatory way, based on participants' perceptions or experiences, in which case this is a sub-set of participatory evaluation.

The difference between case-based designs and other approaches is that the evaluator intentionally goes out looking for variation and extremes, a range of different outcomes and circumstances, in order better to understand the underlying mechanisms. Methods are interpretative, relying on in-depth experience and inference. Structured and partially quantitative

methods for comparing cases are available, such as qualitative comparative analysis (QCA; see Example 4). Case-based designs are particularly useful when an intervention is being implemented across multiple contexts, and when an understanding of context is seen as particularly important in understanding success or failure.

Example 4. Case-based design

When looking at cases, there is unlikely to be a single factor that leads to a specific outcome, but a package of causal factors. Qualitative comparative analysis (QCA) is a case-based method that allows the identification of different combinations of factors (often called conditions) that are critical for a given outcome. QCA translates the complexity of in-depth case studies into comparative formulas in order to generate inferences on enabling factors. It therefore combines features of qualitative studies, in that it is interested in holistic cases and complexity, with quantitative approaches that focus on generalisation. It allows the analysis of often complex conservation and development interventions that work differently in different contexts (Schatz and Welle, 2016).

The Center for International Forestry Research (CIFOR) is using QCA in an ongoing process to identify the factors involved in establishing REDD+ (reducing emissions from deforestation and forest degradation) policy at the national level in a way that promotes transformational change likely to lead to successful REDD+ implementation that achieves effectiveness, efficiency, and equity (Sehring *et al*., 2013). This outcome is based on new institutions, procedures and capacity being established, which support concrete policy formulation built on a consensus for change.

Cases were defined as countries, and 12 countries were analysed. The first step in QCA is to develop a set of factors to test, based on a detailed theory of change including contextual conditions. Factors affecting REDD+ were explored thoroughly and indicators developed through an extensive participatory process with expert country teams and reviews of the literature. A preliminary list of 14 factors was reduced to six. Qualitative indicators were developed to provide clear and consistent bench marks for assessing the presence or absence of the condition. For example, one condition was 'national ownership of policy', reliant on presence of indicators such as adequate budget allocations to REDD+, pro-REDD+ statements made by government, and national actors rather than donor agencies dominating policy processes.

The factors were categorised into 'remote' conditions, which are contextual, stable and not easily changed, and 'proximate' conditions, which are closer to the outcome in space and time and easier to change. These categories were analysed in two steps, to allow for inferences to be made regarding which factors play a role if certain contextual conditions exist. This two-step process also reduces the number of possible combination of factors, partly tackling the problem of a lack of diversity in the combinations of conditions across the cases, making it impossible to discount the effects of missing combinations.

In QCA, binary scores for presence or absence are assigned based on analysis of the information (Table 3). This makes it challenging for assessing complex factors, meaning that rather different situations may be given the same value. However, linking the results

with qualitative interpretation allows for nuanced inferences to come through. Some QCA analyses use 'fuzzy sets', which allow a graded assessment of condition fulfilment, scored from 0 to 1 (for instance 0.33). Either way, the tabulation of scores allows patterns of configurations to be explored. The data can be analysed through specialist software, but depending on the amount of data, analysis by eye may be sufficient.

The configuration of successful cases was compared with unsuccessful cases. Only three countries achieved the desired outcome (Brazil, Indonesia and Vietnam). There were two sets of remote enabling conditions. The crucial institutional factor is that policy changes have already been initiated, which provide a smooth path for REDD; for instance, in Vietnam pilot PES projects were initiated in 2008. But this was not sufficient on its own (as shown in the case of Peru), with either effective forest legislation (Brazil and Vietnam) or high pressure from the shortage of forest resources (Indonesia) also necessary for success. National ownership of policy and transformational coalitions of policy actors were also necessary proximate factors. Inclusiveness of policy processes involving key stakeholders such as civil society played only a minor role. Bolivia's outcome showed that although it had the enabling context, it lacked important proximate conditions, which should therefore be the focus of future policy efforts.

Table 3. Configuration of conditions involved in establishing REDD+ policy

	Remote contextual conditions			Proximate conditions			
	Forest resource pressure	Effective forest legislation	Initiative policy change	National ownership	Coalition of policy actors	Inclusion policy processes	Outcome
Brazil	1	1	1	1	1	1	1
Indonesia	1	0	1	1	1	0	1
Vietnam	0	1	1	1	1	0	1
Bolivia	1	1	1	0	0	0	0
Peru	0	0	1	1	1	1	0

Source: Sehring *et al.* (2013)

Participatory designs

Participatory designs focus on garnering the perceptions of intervention beneficiaries to understand impacts and processes of change. These can be done in the place where the intervention took place, without controls or baselines, so this is a popular low-tech choice. But the attributive power of these methods can be questionable, and they need to be very carefully done. It should not be seen as a 'quick and dirty' approach, but as one which answers a specific question about how change is perceived to be operating. A variety of systematic methods can be used to show the relative strength of perceived impacts from intervention and non-intervention

factors, such as ranking and scoring using participatory rural appraisal (PRA) techniques (see Example 5; Catley *et al.*, 2013), or narrative, qualitative methods (Example 6).

The strength of these approaches is that they focus on what is important to the people targeted by the intervention: what they think the impacts were. In many ways this is what matters. These approaches can be used as part of an action research approach, which uses research as part of the intervention itself. More detail about the participatory approach to data collection, and in particular some caveats, is in Section 4.

Example 5: Participatory design

The social assessment of protected areas (SAPA) methodology has been designed as a relatively simple, rapid, low-cost approach to assessing wellbeing impacts of the establishment, existence and management of protected areas (PAs) of any type, based on local perceptions of impacts. Including multiple stakeholders in the process aims to enhance credibility, ensure engagement and meet the information needs of PA managers, government and civil society actors (Franks and Small, 2016). It combines participatory rural appraisal (PRA) tools with a simple household survey to generate both quantitative results and qualitative insights.

The SAPA process is based on a broad framing of wellbeing in three dimensions (material, relational and subjective; Figure 2), and assesses the contribution of the PA (both benefits and costs) across these dimensions. All significant benefits and costs relating to the PA and associated activities are identified, but in-depth assessment of how significant these are for wellbeing is focused on priority areas only. The design is founded on the use of 'reflexive counterfactuals' — meaning that participants compare their current situation to that before the intervention, and define benefits and costs that they feel are attributable to the intervention.

A pilot study was conducted for the Ol Pejeta Conservancy (OPC) in Kenya — a private PA run by a non-profit Kenyan conservation trust (Franks *et al.*, 2014). The facilitation team was composed of the OPC management, an independent civil society group, and representatives from the community-PA advisory group. Crucially, then, this type of participatory process requires a willingness for stakeholders to work together, listen, and take action on issues raised by communities. The process begins with focus group discussions in local communities (targeting a cross-section) to identify and prioritise impacts. In Kenya, groups were split by gender and asked to brainstorm a list of costs and benefits of the PA, and then prioritise the impacts. Afterwards, they came back together, to identify and explore any significant differences between the views of men and women. The groups proposed a total of 18 benefits and seven costs of the OPC. Priorities for more in-depth assessment were identified by asking each participant to allocate beans to impacts that they considered most important from what they had seen in the community during the intervention period (2 beans for high; 1 for medium; 0 for low). Positive priorities included fencing preventing wildlife getting into farms; improved security due to ranger presence; schools and bursaries; and healthcare from the community development programmes. Negative impacts prioritised were crop damage by wildlife; exclusion from jobs; uneven distribution of projects (unfairness); and poor relationships with OPC staff.

The next step in the SAPA process is to design and pilot a household survey based on the impacts identified in the focus groups. In Kenya, households were randomly sampled across villages surrounding the PA, with either men or women respondents depending on who was available. Respondents were asked whether each of the prioritised impacts had high (3), medium (2) or low significance (1) in contributing to wellbeing in their own household. These values were then averaged for each impact, and any differences between social groups and geographic area investigated. For instance, poorer households rated negative impacts of higher significance, and men gave higher priority to uneven job and project distributions.

Other locally important wellbeing indicators, representing the three dimensions of wellbeing, can be inserted into the survey based on community discussions. In Kenya, for the 'material' dimension, food security was indicated by frequency of skipped meals; respondents were also asked how this indicator had changed (improved; remained the same; or deteriorated) and why. At the end of the survey, respondents were asked to summarise the overall impact of the PA on the wellbeing of their household, taking into account all the benefits and costs discussed, on a scale of -2 to +2. In the pilot study, 68 per cent reported that the OPC had made a positive contribution overall, reflecting the unusual amounts of investment in development and relatively high levels of revenue from tourism. However, there were geographical differences in these responses, with areas in the southwest reporting a less positive situation due to crop damage by wildlife and fewer benefits received.

Participatory rural appraisal methods can be used in impact evaluation, including visualisations to rank and score changes (see Catley *et al.*, 2013). Women in Western Tanzania with a completed seasonal matrix of resource use. Credit: Paulo Wilfred.

Example 6. Narrative-based participatory design

This example highlights the 'most significant change' (MSC) approach, which was used for an elephant conservation project in Cambodia. The MSC approach involves collecting personal accounts of change ('stories') from the field, and systematically selecting the most important of these (Davies and Dart, 2005). It is participatory in that project stakeholders at different levels are involved in deciding the kinds of changes that should be recorded, and in analysing the data collected. Significant stories are passed between layers of organisation and feedback is given to stakeholders encouraging upward and downward accountability. Fauna & Flora International piloted the use of MSC in several sites, including for the Cambodian Elephant Conservation Group (CECG). The CECG project focused on reducing human–elephant conflict in farming communities on the forest periphery, which had been impacted by the destruction of crops, property and occasionally harm to people, resulting in negative impacts on livelihoods and retribution killings. Project activities focused on low-tech deterrents, support for growing crops less palatable to elephants, and the formation of guarding groups. The MSC process involves a series of steps (Wilder and Walpole, 2008), as follows:

Defining the domains of change: First, broad categories of change are identified which allow structured analysis without prescribing what constitutes change in the local area. For the CECG project, domains were developed which aligned with the project objectives, including: change in attitudes (to the project, elephants, conservation); change in food security; change in social cohesion and cooperation; and an open category of 'any other change'.

Collecting significant change stories from those most directly involved in the intervention. In Cambodia, stories were recorded onto data sheets by project field teams at the two sites, either through their own observations, or through eliciting stories from community members informally through their regular interactions. Each story was allocated to a domain. The project team decided to collect stories from individuals rather than through focus groups to avoid elite voices dominating. Details on where, when, why and who was involved in the MSC story were recorded. Examples of both positive and negative change were recorded; for example, farmers reported successfully cultivating and selling crops which tended to be raided less by elephants, improved yields, and enhanced wellbeing due to the improved alarm systems. A negative and unexpected change was that a dam system implemented to improve crop productivity resulted in conflict over water supply downstream.

Selecting the most significant stories: the stories elicited are analysed and filtered upwards through layers of the organisation, with panels at each level selecting the most significant story in each domain, and providing reasons for the selection. This results in a smaller number of widely valued stories. One potential disadvantage is the subjectivity involved in the selection, which requires a commitment to transparency about the criteria and values being used in choosing stories. Stories recorded for the CECG project were discussed with team leaders at bi-monthly meetings to debate verity and whether a story was considered 'significant' and why. Stories were then passed up to the national team, and ultimately donors. One issue in this context was the need to translate stories from

Khmer to English, which was resource-demanding, and introduced an additional layer of potential bias.

Verification: stories are brought back to the project site to ensure they are an accurate and honest reflection of events. In Cambodia this was done informally through discussion among the field team, and where necessary with subjects of the story, although more formal methods for external reporting were considered — through follow-up evaluation and triangulation with conflict incident data, for example. Monitoring of who the stories are coming from, and the kinds of stories selected, can allow reflection on how participation and values are affecting outcomes of the process.

MSC is unlikely to be a sole method of monitoring and evaluation, but can augment other methods (such as quantitative or theory-based methods) by informing indicator development, identifying unexpected changes, providing a rich picture of change, and capturing diversity of experiences. It has the potential to increase organisational capacity at different levels. Teams reported that their own capacity for analytical thought and awareness of the wider impacts of their work had increased. It can form the basis for adaptive management, by highlighting unexpected changes and problems; including contextual changes about causal linkages. It can also serve to build trust with local people, through encouraging feedback to them on issues they have raised in their stories.

There are clearly issues of voice, power and trust that need to be addressed in using a methodology such as MSC. There is for instance, a tendency for the process to capture success stories over failure, so that a concerted effort is required to provide the space and generate the trust needed to understand negative experiences and the views of the marginalised. In particular, the reporting of sensitive or negative stories has the potential to cause harm to individuals and relationships. It is vital to consider who conducts the interviews, the composition of the selection panels, training needs, and representation of local people to ensure the process is not co-opted by elites or organisation managers to serve their own interests (see Section 4).

The 'most significant changes' for women from the Cambodian Elephant Conservation Group project included the ability to grow and sell cardamom. Credit: Fauna & Flora International.

3.2 Useful concepts for all evaluation designs

It is useful to consider the following three concepts when carrying out evaluations, even if only as 'thought experiments', if it is not possible to use them formally:

Baselines, to be able to see change over time. What was the situation before the intervention started? It can be possible to construct baselines 'post-hoc', or after the intervention has started, from secondary data or carefully designed surveys. For example this could be based on people's recall of pre-intervention conditions, but caution is needed because this recall is likely to be biased in different ways. Imaginative use of a range of approaches and datasets to get a feel for the pre-intervention status of the system can give support to case-based and theory-based approaches, even if the project is not designed with baselines from the beginning.

Counterfactuals, to be able to see how change differs from what would be expected without the intervention under evaluation. Counterfactuals do not have to be statistical, and a plausible counterfactual can be constructed post-hoc based on observed trends in relevant variables over time, even without exhaustive data (Bull *et al.* 2015). Counterfactuals can also be proxied by spatial differences, such as in matched controls, or including non-intervention cases in case-based evaluations of outcomes under different circumstances. As discussed in Example 5, 'reflexive counterfactuals' — in which people report how their situation differs to that before the intervention, and how the intervention has affected it — are valid in participatory designs.

A theory of change, which gives a hypothesis about how the intervention is supposed to have created change; this will enable you to structure your evaluation to see whether this causal chain was followed. Even if data can only illuminate early steps of the ToC, this can make you more confident that the impact will occur. It also helps in experimental and statistical designs for identifying controls and confounding factors, and contextualising the results of analyses.

3.3 Capturing differential impacts and experiences

Often, especially when using quasi-experimental designs and statistical modelling, it is only possible to capture the **overall, average effect** of the intervention. This is due to the difficulty of getting enough statistical power to fit models appropriately when there are data limitations, caused by the difficulty of getting a large enough sample size. Every split of the dataset (such as by gender, livelihood type, social status) substantially increases the sample size required to reliably detect whether there has been an impact. Overall estimates of effects are not ideal, however, because they can conceal important differences that vary across locations or groups of people — for example by gender, wealth, livelihood type, social status or ethnicity. See Example 2 for an example of livelihood differences impacting on wellbeing outcomes.

In more qualitative approaches (theory-based, case studies and participatory designs), it can be easier to ensure that the experiences of different types of household or individual are captured, such as ensuring that case studies of households or villages include a range of wealth levels, or that men and women are included in focus group discussions. An understanding of social structures is needed in order to ensure that the full range of relevant **variation in experience** is captured.

It is vital to pay attention to the **sampling unit**: do you want to know whether the social impacts of your work differ between villages, households, or individuals? It is common to sample at the household level, because it can be a relatively natural unit, but this can conceal important intra-household variability. For example, women may have very different experiences of the intervention to men, and older people different experiences to younger age groups. It is also crucial to consider what constitutes a household in the local context, which may not be as a straightforward as it first appears, for example due to high mobility of fostered children or other dependents between households as labour.

It can be very difficult to reach marginalised **'invisible' groups** (such as nomadic peoples, migrants, marginalised castes, the poorest households, women). These people may not appear on censuses, may not be invited to join focus groups by the elites, or may not live within the village. However, conservation interventions may often disproportionately affect them. For example, people living out in the forest rather than in the main village may be most reliant on forest products and therefore hit by access restrictions, but not benefit from compensatory benefits like better schools or community facilities. It is very important to seek out these groups and include them in the evaluation.

4 Collecting data

4.1 Selecting appropriate methods and tools

Within particular designs (discussed in Section 3), different methods and tools can be used to collect data — on indicators of outcomes and outputs, context, and processes (definitions in Box 1). Some types of data and data collection methods are linked more closely to particular design types (see Table 4). For instance, statistical designs will require quantitative data, and case studies are more likely to be based on qualitative data, though may include quantitative data. However, it is important that all quantitative surveys are based on qualitative understandings of the cultural, institutional and historical context, to ensure that meaningful questions are being asked (Drury *et al.*, 2011). Identification of variables to use in selecting controls in quasi-experimental designs also needs to be based on qualitative understandings of the system.

Box 1. Defining data collection terms for impact evaluation

Tools: specific data gathering instruments or exercises, such as the Basic Necessities Survey (see Example 1, Section 3.1) or 'most significant change' (Example 6).

Methods: sets of tools of a certain type, such as participatory rural appraisal or qualitative interviews.

Methodologies: overall package of experimental design and information gathering tools, for instance SAPA, which comprises PRA, questionnaire surveys, and a participatory design (Example 5).

Designs: the way that data are collected to allow causal inferences to be made, such as randomised controlled trials, or case-based searches for particularly insightful extremes.

Adapted from Schreckenberg *et al.* (2010)

Table 4. Differences between quantitative, qualitative and mixed data collection

Type of data	Most useful for	Specific tools and methods
Quantitative (numeric)	▪ Measuring magnitude of changes ▪ Collecting objective, measurable data ▪ Statistical analysis ▪ Reporting to external interest groups (funders, governments)	▪ Structured questionnaires (closed-ended questions) — interviews or self-completed ▪ Direct measurements ▪ Participatory methods which involve ranking and scoring (see Mayoux and Chambers, 2005)
Qualitative (text-based)	▪ Contextual knowledge ▪ Processes of change ▪ Providing basis for quantitative questions ▪ Uncertain outcomes, outcomes that are difficult to measure (eg social relations, political change), sensitive issues ▪ Adaptive management and understanding mechanisms	▪ Focus group discussions ▪ Open-ended interviews ▪ Key informant interviews ▪ Participant observation (ethnographic) ▪ Most significant change
Mixed (quantitative and qualitative)	▪ Combining strengths of both approaches for different insights and audiences ▪ Triangulating understandings based on information from different approaches	▪ Combining the collection of different data types within single methods (eg qualitative and quantitative questions within a household survey), or using a range of methods within the evaluation (eg focus groups and key informant interviews prior to a household survey) ▪ Some specific tools use mixed methods, eg Basic Necessities Survey

Data may also be objective or subjective (Table 5):

- **Objective** data are observable and externally verifiable (although they may be based on survey data and may be qualitative); for instance number of livestock, housing quality, school enrolment, livelihood type, participation in meetings.

- **Subjective** data are based on the feelings and perceptions of people, and are not externally verifiable. For example, how content people are with their lives, feelings of empowerment, feelings of exclusion.

Subjective data are needed to understand the subjective dimension of wellbeing; collecting this type of data will give different types of insights to objective data. Subjective data can still be quantitative, for instance Likert scales (such as asking people to rate their satisfaction with life on a scale from -2 to +2), but it is important to ensure that this is a meaningful approach for respondents. It is also important to recognise that changes in objective indicators may not reflect changes in subjective indicators, highlighting the need to capture both dimensions of wellbeing (Box 2).

Table 5. Examples of different data types capturing aspects of food security

	Subjective	Objective
Qualitative	Narrative description of perceptions of change in food availability	Child's nutritional status category based on observation
Quantitative	Quantified anxiety levels regarding food supply on a scale of 0 to 5	Data on food stocks or income spent on food

Box 2. Complex impacts of a marine protected area on objective and subjective wellbeing

Gurney *et al.* (2014) evaluated the short, medium and long-term impacts of integrated marine protected areas (MPAs) in Northern Sulawesi, Indonesia, which were designed to achieve the dual goals of conservation and poverty alleviation. MPAs were implemented through participatory processes in four villages between 1997 and 2002 alongside development activities to improve access to drinking water, livelihood training and environmental education. The study drew on the World Bank's multidimensional approach to poverty encompassing material opportunities, empowerment and security. These three domains were represented by context-specific indicators which were either objective (such as livelihood diversity measured by total number of occupations divided by number of household members) or subjective (perceptions measured on a scale). A quasi-experimental design was used, drawing on longitudinal data from the four villages which were matched (not statistically) to control villages.

Impacts on indicators representing material opportunity (objective wealth based on assets, environmental knowledge, and subjective perceptions of fish catch) were improved in the MPA villages, although these positive impacts were not sustained after the implementation period when funds were withdrawn. In the security domain, livelihood diversity did not decline as much as in the control villages. However, perception of present wellbeing (measured on a 15-point scale) was negatively affected by the MPAs. Insights from qualitative research suggested that subjective wellbeing had been affected by conflict and unmet expectations regarding tourism, despite improvements in livelihoods. Villagers were not happy with the misuse of funds, inequitable sharing of benefits, and confusion over property rights; and there had been failures in punishing poachers due to a lack of multi-level governance structures.

4.2 Applying the chosen methods

Producing valid results is not just about selecting the right tool, but needs consideration as to the way these tools are applied, to take into account culturally sensitive issues, bias, vested interests and ethics. Otherwise the results produced will be compromised, and potentially highly misleading. Best practice needs to be followed so that the methods used are rigorous and appropriate. For example, if you ask a leading question, you will get the answer you expect.

It is important not to assume that qualitative or participatory approaches can be applied in a less rigorous way than quantitative or statistical methods, or that ranking and scoring methods are a quick fix. Whatever the method, amateurs will not get results which are as robust and trustworthy as people who have been trained in these methods and have experience in applying them in a range of contexts.

Another set of questions, therefore, concerns **how** evaluations should be carried out, and **by whom:**

- **Independence:** in most cases, evaluations should not be carried out by the same people who implement the intervention, in order to prevent bias and vested interests influencing the results. It may especially be a better option if there are tensions that mean that people will not open up to those involved in the intervention or organisation. However, if the intention of the evaluation is focused more on collaborative learning, then self-evaluation by project staff alongside communities may be more appropriate (see section on collaborative learning and participation below).

- **Composition of the research team:** the researchers who interact with people should speak the local language and ideally be from the area, or have lived there for a while. This will avoid obscuring meanings; ethnicity and local knowledge is a factor in cultural understanding and building trust. Often, evaluators need to rely on local field researchers who speak the language in order to carry out the research. It is important to be aware that being one step removed from the respondent weakens the evaluation, particularly in qualitative and discussion-based approaches. It is also important to take advice from the local field assistants, and be open to developing the evaluation approach with them, as they are better placed to spot and rectify pitfalls. For example, the local word or phrase for 'wellbeing', and the concepts that surround it, are often difficult to get at without deep local understanding.

- **The identity of the researcher** can affect the equity of the process and lead to subversion or 'capture' of the evaluation by interest groups. For example, outsiders may be duped into thinking they are speaking to the key people, whereas in fact people with different viewpoints are not able to participate. Again, taking advice from a trusted and experienced local field assistant or key informant can help to get a more realistic picture of what is going on.

- **Power dynamics:** be aware of power dynamics that may prevent truthful responses. We cannot escape the relationships between and perceptions of researcher and participant, but we need to consider how these may influence the data. How much are respondents likely to trust the evaluator? How can they relate their experiences to your identity? And how do they relate to your field assistants and research team more generally?

- **Ethics:** go through an ethical review procedure, even if it is internal or informal, in order to make sure you have really thought about the issues that may arise. Make sure you base it on best practice guidelines (see Box 3). Ethical practices are there to protect participants' rights and safety. Document your decisions and the reasons for them. It is vital that your respondents give their free, prior and informed consent (FPIC; see Box 4) before you talk to them. Anonymity and confidentiality should always be offered, unless the interviewee is speaking in an official capacity. People may need to be sure their privacy and reputation is not at stake over activities which may not appear sensitive to the researcher. Confidentiality and anonymity are particularly vital when talking about illegal or socially unacceptable activities.

Box 3. Relevant sources of ethical guidance

Association of Social Anthropologists of the UK and Commonwealth:
http://www.theasa.org/downloads/ASA%20ethics%20guidelines%202011.pdf

British Sociological Association:
https://www.britsoc.co.uk/media/23902/statementofethicalpractice.pdf

Economic and Social Research Council (for a range of case studies highlighting ethical issues):
www.esrc.ac.uk/funding/guidance-for-applicants/research-ethics/ethics-case-studies

Box 4. Free, prior and informed consent

The principle of free, prior and informed consent is enshrined in international law. It is required before any project is undertaken, including research, that could impact on local communities (in particular indigenous groups). The meaning of the term is broken down as follows:

Free: consent for the project should be given voluntarily and without coercion, intimidation or manipulation. Stakeholders must determine the decision-making structures and processes and all community members must be free to participate or not as they choose.

Prior: consent should be sought sufficiently in advance of commencement of the project, in the early stages of development, so that it can be understood and verified. The local community's decision-making timeline must be respected.

Informed: information about the project must be provided that is clear, accessible, objective and independent, and complete. It should be delivered in a culturally appropriate way (for instance in the local language, clearly explaining concepts which may be alien to culturally specific ways of thinking), in sufficient time to be understood. Information must reach the most remote communities and marginalised groups, and be provided on an ongoing basis throughout the process.

Consent: the project must allow communities and individuals to grant or withhold their consent at any stage, according to the decision-making processes of their choice. Consent is not a one-off process and the decision must be respected at any point.

Adapted from: UN-REDD Programme (2013) Guidelines on Free Prior and Informed Consent.

4.3 Using evaluation as an opportunity to learn collaboratively

Evaluations can be independent or carried out by the implementing agency. Independence gives external validity and can give those affected by the intervention the confidence to speak freely. However, they may be expensive or hard to organise on a regular basis. Internal evaluations provide a better chance for joint learning between the implementing team, the project participants and those affected more generally.

One useful approach is to commission an independent evaluation by an expert evaluator, and ask them to develop appropriate approaches and protocols for ongoing monitoring by the implementing agency on a more regular basis, which could be lighter touch or lower tech.

In either case, data collection may be more or less **participatory**. Following criticism of conventional top-down technocratic interventions, participation has emerged since the 1990s as an essential element of people-centred development, which aims to put the poorest first, and in control of the processes that impact on their lives (Chambers, 1997). Both quantitative and qualitative methods for understanding social impacts can be more or less participatory, with researcher attitudes being the crucial factor in listening and learning from people who are targeted for interventions, during their design and evaluation.

A participatory approach, done well, can encourage relationship building between stakeholders, shared learning experiences and commitment to apply learning to adaptively manage projects, and therefore lead to greater sustainability of programme outcomes. Participatory approaches may also be a way of making causal inferences using insider rather than outsider perspectives (see participatory design in Section 3), and again the nature and level of participation in this process can differ.

Participation can take a variety of different forms, from consultative — whereby communities are involved in defining relevant indicators — to a fully locally driven approach, which is instigated, designed and conducted by local people with empowerment and mobilisation being an aim in itself (Table 6). Organisations managing interventions need to decide on the most appropriate and achievable level of participation for their programme. This may depend on the model of conservation being used and the goals of the intervention, but evaluators should always look to maximise integration of local people. The often overused and ambiguous term ‘participation’ can be rather hollow and be used to mask extractive techniques of obtaining information. It is important to be realistic and transparent about the level and type of participation applied in monitoring and evaluation, both with the people whose change in wellbeing is being evaluated, and with the users of the evaluation. Box 5 describes a case study which highlights how levels of participation and local relevance of indicators can be contested between stakeholders involved in project monitoring and evaluation.

Samburu men in Kenya taking part in a focus group as part of a participatory evaluation of the Warrior Watch conservation programme. Credit: Heather Gurd/Ewaso Lions.

Table 6. Different levels of participation in social impact monitoring and evaluation

Category of monitoring and evaluation	Characteristics	Type of participation
Professional	▪ No involvement of local people (except maybe for consent) ▪ Design, data collection, analysis and data use by professional researchers ▪ Evaluation is 'done to' people ▪ Externally defined questions, indicators; all objective ▪ Findings not shared with local people	Manipulative or passive. People's involvement is superficial and they have no influence or power in decision making. Responses and contributions are not listened to
Externally driven	▪ Local people only involved in the data collection stage, with professional researchers designing, analysing and using the data ▪ Local people consulted on indicators	Consultative or incentivised. Project design and info gathering process is controlled externally. Locals are only involved through working for rewards, or consultation where there is no obligation for external people to heed local views
Collaborative – with external design and data analysis	▪ Local people are involved in the data collection and data use in resource management ▪ Design and analysis carried out by professional researchers ▪ Indicators locally defined, subjectivity respected and local voices heard ▪ Results shared and local views feed into decision making	Functional. Local people involved in decision-making processes, though big decisions are often taken externally, and in advance. Participation is a project goal, but seen as a means to achieve evaluation, especially in a cost-effective way

Category of monitoring and evaluation	Characteristics	Type of participation
Devolved, community-based with external advice	▪ Local people involved in all areas of the evaluation process, with professional researchers giving support where needed ▪ Findings owned by local people, and decisions locally driven	Interactive. Local people have control of project design, action plans, resource allocation and activities. They have a stake in the process. Participation is a right, not a means of achieving a goal
Autonomous, locally driven	▪ Very little external involvement ▪ Design, monitoring, analysis and data use by local people ▪ Local people develop responses to evaluation results	Self-mobilisation. Initiative taken locally to address issues. Contact may be made with external institutions for resources, technical advice or advocacy. The process may challenge existing power structures and externally conceived project aims

Source: Adapted from Palmer Fry (2014); with categories of monitoring based on Danielsen *et al.* (2008); and types of participation based on Pretty (1995).

Box 5. Case study of participatory wellbeing monitoring in a REDD+ project in Guyana

In the North Rupununi, Guyana, a community forest monitoring system was established loosely under the banner of REDD+. This project included monitoring human wellbeing alongside natural resources. The design phase for the intervention involved multi-stakeholder focus groups as well as numerous meetings and consultations. Those pursuing internal validity through locally driven processes were the community leaders and the trained community monitors, the local project management team, some of the non-resident NGO advisors, and an external academic. This group were pushing for a mostly qualitative locally defined monitoring system focusing on resource dependent livelihoods, development indicators and community cohesion, with the advisor and the academic stressing the importance of empowerment and local leadership. Internal validity under these headings requires locally relevant indicators such as the possession of key assets, family and community relationships, and farming success. Much of this information cannot be presented with numbers.

The remaining key stakeholders advocated external validity and externally driven monitoring processes: the government of Guyana wanted a community monitoring system that followed national forestry practices and could be rolled out nationally (they were not particularly interested in wellbeing given that REDD+ focuses on forest carbon); and the international NGO's main emphasis was developing a progressive, forest-centred monitoring system that could be used elsewhere in South America and could provide information worthy of carbon payments. External validity under these criteria entails the use of quantitative standardised measures that are broadly recognised and comparable to other societies, such as statistics on education and stable access to food.

As the project reached its conclusion, the emphasis was on quantitative external validity with two thirds of the monitoring system addressing government or investor interests and one third addressing community interests. This imbalance in local involvement and relevance meant that community members had little incentive to continue monitoring aside from external payments, which have ceased. This has ultimately led to the monitoring coming to an end.

Adapted from Palmer Fry (2014)

5 Developing an appropriate evaluation for your needs

Each component of the evaluation involves decisions based on a range of factors related to: 1) the purpose of the evaluation; 2) attributes of the intervention; 3) capacity and resources. There is no blueprint or perfect solution for making these decisions; evaluators and those commissioning the evaluation need to consider their own situation. We provide guidance by highlighting how these three factors can affect or preclude certain options, and point towards solutions.

5.1 Purpose of the evaluation

The purpose of the evaluation is the primary factor that will determine the most appropriate approach to take, in particular regarding the design (Table 7). Some evaluations may focus more on determining whether the intervention caused a particular impact and others more on explaining why an impact happened as it did. Often, and ideally, we will want to do both, but some options may be limited by resources and attributes of the intervention itself. It is important to note that evaluations often take the form of hybrid designs drawing on the strengths of different designs to answer different questions.

Table 7. Linking evaluation questions, purpose, and suitable designs

Evaluation question	Related questions	Purpose of the evaluation	Assumptions	Requirements / suitable designs
To what extent can a specific impact be attributed to the intervention?	What is the net effect of the intervention? How much of the impact can be attributed to the intervention?	▪ Accountability ▪ Contributing to scientific evidence base ▪ Scientific publication ▪ Reporting to donors, advocacy	Expected outcomes clearly understood and specified Focus on primary cause and effect	Controls and/or baselines available Sufficient sample sizes **Quasi-experimental** **Statistical analysis**
Did the intervention make a difference?	What conditions were necessary or sufficient for the effect to happen? Was the intervention needed to produce the effect?	▪ Accountability ▪ Less stringent requirements regarding direct attribution ▪ Focusing on contribution over attribution	There are several causes for change that need to be disentangled	Comparable cases **Theory-based** **Case-based analysis** **Participatory** Also: **Quasi-experimental** **Statistical**

Evaluation question	Related questions	Purpose of the evaluation	Assumptions	Requirements / suitable designs
How has the intervention made a difference?	How and why have the impacts come about? Has the intervention resulted in any unintended impacts? For whom has the intervention made a difference?	▪ Lesson learning within or between programmes ▪ Adaptive management	Possible to develop clear theory of causal processes	A theory of change that identifies proximate, contextual and historical factors **Theory-based** **Participatory** **Case-based if sufficient testing of theory**
Will the intervention work elsewhere?	Can this intervention be transferred elsewhere and scaled up? What generalisable lessons have we learned about impact?	▪ Learning between programmes and organisations – scaling up ▪ Evidence for funders of impact of general approach	There are generalisable similarities between circumstances that can be identified	Understanding of context and how it affects the outcomes **All the methods can potentially shed light on this question but strong theory needed**

Source: Adapted from Stern (2015)

The purpose of the evaluation will also impact on the type of data to collect, and the specific method used. For instance an evaluation focused on adaptive capacity and building relationships with local communities may focus more on the subjective experiences of local people, and take a participatory approach. It may also affect the timescale of data collection; for example for the purposes of adaptive management, a repeated quantitative measurement may be preferred so that improvements in impact can be tracked over time.

5.2 Intervention attributes

Local context

The social, cultural and political context of the intervention and evaluation may affect the method choice; illiteracy clearly rules out self-completed questionnaires, for instance. It will also inform the application of methods. For example, the sensitivity of issues regarding the intervention (such as land use, inequities, or illegal resource use) may mean indirect questioning approaches are needed, such as the 'unmatched count' technique, in which respondents are only asked how many statements they agree with or behaviours they engage with on a list (Nuno and St John, 2014). As discussed in Section 2.3, local context will also inform the locally relevant definition of wellbeing and choice of indicators. The geography and climate may affect logistical decisions and the application of methods; for example the timing of the evaluation may affect responses if there is strong seasonality in livelihoods.

Longevity of work in the local context

How familiar the implementing organisation is with the local context, and the nature of relationships they have with local people, will affect levels of trust, openness, and willingness to participate in an evaluation, all affecting the validity of the data. The level of understanding about local issues will determine the evaluators' ability to make a priori decisions on wellbeing indicators and to develop a ToC without extensive work with local stakeholders. If this is a limitation, expert local knowledge will need to be sought, and evaluators may need to scale back their expectations about the depth of the evaluation and/or the questions they are able to answer.

It is also worth considering the types of knowledge that are deemed legitimate in the implementing organisation, as field staff may have more in-depth local knowledge than managers, and local people more than either. For example, there may be more innovative ways to construct baselines than the use of traditional survey data, through accessing undocumented knowledge of local staff and experts if there is a long history of interaction.

Geographical scale

The scale at which the intervention has been implemented has implications for sampling design. Representing the full range of impacts of an intervention requires sampling across a larger scale than when the main question relates only to its direct impacts on the targets of the intervention; perhaps encompassing more communities, in which case contextual differences between communities will need to be captured. It will be difficult to find controls for large-scale or national-level interventions such as policy changes, in which case, theory-based and case study designs may be most appropriate (as in Examples 3 and 4). It may be possible to do nested designs for activities at different scales, for instance using secondary datasets for national or regional trends (such as government records); quasi-experimental designs to compare impacts at the community level; and case studies on particular villages to study causal pathways. Small-scale interventions implemented in only one or two communities will allow more depth of understanding for a given budget. Where capacity is low, but the scale is relatively large, a few case study villages representing particular contextual characteristics and outcomes may have to be selected and that selection justified. There will also clearly be logistical and budgetary implications for the scale of evaluations, such as the time taken to get between isolated villages.

Timescale of likely impacts and change over time

There may be significant time lags between intervention and effect. For example, improvements in education as an outcome of greater income and livelihood security may take several years to become observable. Post-intervention evaluation is rarely possible, but at the least longer-term impacts (both positive and negative) should be considered in the ToC even if they cannot be evaluated at this stage. For example, evidence that people intend to use funds for education can provide support for this pathway to impact. For projects with short timelines, the focus may have to be on outputs rather than outcomes and impacts, with a strong ToC backing up the intended impacts. Trajectories of change may not be linear, so there could be high initial impact that tails off owing, for example, to improved forest governance arrangements that are eroded through time by pre-existing power structures (also see example in Box 2). Ideally, monitoring throughout the course of an intervention should be carried out to allow for adaptive management, especially when the priority is to learn.

When studying subjective wellbeing, it is important to consider that the reference standards of people impacted may change, potentially as a result of the intervention itself. Increasing material wealth may lead to wealth becoming a more important aspect of people's wellbeing, but result in reduced subjective wellbeing, owing to rising aspirations. Keeping track of this change — for example by understanding who people are comparing themselves to — and what affects subjective experiences, will help in the interpretation of evaluation data as people's situations and aspirations change.

Pre-existing theory of change

If there was no clear theory of change developed for the original intervention, participatory methods can be used to draw up a post-hoc ToC, drawing on the knowledge of a range of stakeholders — especially the targets of the intervention. This is a useful exercise to do at the start of an evaluation even if there was a ToC before the intervention, as it might have changed.

Connections between interventions

Discrete interventions with one main project activity will be easier to analyse, and provide the possibility of setting up a quasi-experimental design to attribute impacts to specific projects. The ToC may be more straightforward in some respects, but evaluators will still need to think about how the intervention interacts with other interventions in the wider landscape and with contextual factors such as the economy and social structures. Sometimes the 'intervention landscape' will be complex with lots of different actors working at different temporal and spatial scales, and with these interventions and actors changing over time. Then it will be more effective to focus on the contribution of the intervention, rather than on attribution. The project may not have been sufficient on its own to lead to wellbeing changes, but may support and be supported by other programmes, policies and actors.

Certainty and complexity of the context and outcomes

If there is one primary cause and one primary effect, (quasi) experimental designs will be more straightforward to carry out well. Often the situation is more complex and uncertain. Outcomes may be uncertain due to the implementer's limited experience of the locality, or they may take place in a context which is politically unstable and experiencing rapid change. Therefore, certain

outcomes may only emerge through the course of the intervention. In these situations, the evaluation needs to be flexible, with participatory work to identify possible unanticipated impacts alongside regular monitoring. External change may mean you have to adjust the counterfactual through the course of the intervention. For example, if there is a change in government and the institutional landscape, retaining any forest at all may be all that can be hoped for by the end of the intervention; when before this change, the expected outcome was forest recovery. The intended impacts themselves may be far from straightforward, intangible and difficult to measure. For example, for interventions where the intended impact is cultural change, empowerment, participation or strengthening of governing structures, qualitative analyses will be more suitable than precise attribution questions and quantifying impacts.

Availability of controls

The presence of comparable, non-intervention communities or households may allow the development of valid controls. A lot of contextual understanding is required in order to identify appropriate covariates for statistical matching or to match qualitatively, and technical expertise is needed to analyse the results. However, setting up controls may not be feasible for a variety of reasons: ethical (such as raising expectations of non-intervention villages), ecological (such as interventions on islands of natural habitat which are not comparable), or a small number of units (such as large-scale, or very circumscribed, policy interventions). In this case, you can look into the other options available, to see which is most appropriate to the situation.

5.3 Capacity issues

Budget

The budget available will often be the ultimate constraint on decisions made about evaluation design and methods, as it will impact on staff, training, expertise, time spent and logistics. The budget may limit the kind of question that can be answered and options for methods. However, valid assessments that provide useful information relevant to the local context and for decision-making and relationship building are still possible on a relatively small budget. The SAPA approach to evaluating the impacts of protected areas described in Example 5 is estimated to cost US$5000–10,000 per assessment. Useful evaluations can be done for a few thousand dollars (such as Case Study 3 below, which was carried out by a masters student). Other tools such as the ranked outcomes method (Example 7) are low tech, rapid and simple but can still provide insights into local priorities and impacts. Although linking impacts to intervention is a key aspect of evaluation, 'indicative' evaluations that show associations, and that the balance of evidence suggests an intervention is having an effect, may be all that is required by donors and programme managers. With the right collaborators supplying expertise for the cost of travel only, these are feasible within a tight budget.

Time available for assessment

Impacts can change over time, at different rates and over years rather than weeks. A decision is needed on how best to capture these impacts within the time available for the evaluation, in terms of the time which the evaluator has to carry it out; how soon the evaluation happens after the start of the intervention; and how often the evaluation should be repeated to capture ongoing impacts. A very short time frame for carrying out the evaluation will obviously limit the quality of data

collection or the scope of the data. It will be better in that case to focus on exploratory qualitative focus group discussions, or a few carefully selected case studies of diverse households, rather than collecting very cursory data through rushed household surveys that are likely to be poorly developed and ultimately meaningless without contextual knowledge. This will then impact on the questions that are sensible to try to answer. A well-developed ToC is the best guide to when an evaluation would be most informative, in relation to the expected timing of the intervention's outputs, outcomes and impacts. A useful approach can be using the first independent evaluation to put ongoing, potentially lighter touch, evaluations in place for monitoring longer-term impacts (see Case Study 4 below).

Technical capacity and type

The type of technical capacity available will limit the choice of method to some extent, and evaluations should cater to strengths as far as appropriate. Quasi-experimental and statistical designs require technical quantitative analytical skills; and participatory approaches require experienced facilitators or at least staff with strong communication skills, and sensitivity to local context. Some level of investment in expertise or training is inevitable, but can vary in cost. Calling on local expertise and language skills will decrease the risk of obscuring local meanings, without the risks of parachuting in external 'experts'. One possible route to take is to engage an independent team of researchers with local knowledge, involving in-country students (although urban students may not have the language skills or cultural sensitivity required), and build links with universities with social science expertise who can provide guidance. The twinning of international and in-country universities can be a great way to build local capacity and take advantage of international expertise — as well as for external researchers to learn about local issues and conditions. It can also be cheap, as a university's training mission may mean they do not charge full economic costs.

Example 7. The ranked outcomes approach to evaluating social impacts of conservation interventions

The 'ranked outcomes' approach is a tool for post-hoc evaluation of interventions involving stakeholders, with a low-tech, relatively simple but robust method. The method involves developing a list of desirable outcomes for a project portfolio which can be identified and agreed by project staff, independent reviews and local people, potentially in a participatory way. These can be categorised into topics so that you are comparing like with like, such as education, livelihoods, legacy, and conservation goals, like species protection. Outcomes are ranked within each category by different stakeholders according to their priorities, and then whether or not they have been met is evaluated separately, for example as a binary assessment based on a review of project documents, interviews with staff, or local people, and direct observations to assess whether the outcome has been met. The priority score is multiplied by the outcome score to reach a category score.

Sainsbury *et al*. (2015) tested the method for a portfolio of small-scale income generating projects such as beekeeping, tree planting and fuel efficient stoves, in communities adjacent to a nature reserve in the Tanzanian Eastern Arc Mountains. Outcomes were developed and prioritised with project staff and independently evaluated. Outcomes were also prioritised in village focus groups, and individuals were surveyed on whether the

outcomes were met and then scored on a binary scale. Outcomes included, for example, 'jobs created directly in the project' and 'improved capacity of people involved'. Although there was broad agreement on the performance of particular projects (for instance tree planting was agreed to be the best-performing by both villagers and the independent assessment), there were some differences between perceptions of implementers and beneficiaries on the most important outcomes and on which projects had delivered most. There were also differences between individual assessors and between villages, highlighting the importance of understanding heterogeneity in perceptions.

Tree planting performed best in the ranked outcomes assessment in the Tanzanian Eastern Arc Mountains although there was heterogeneity in perceptions. Credit: Neil Burgess.

The method converts qualitative statements on planned and realised outcomes into a quantitative score weighted according to priorities, allowing some level of comparison between projects and sites. It provides useful insights, especially in data-poor situations where there are no baselines available, and where there are poorly defined or shifting priorities. The accompanying discussions, both within focus groups and with individuals, can draw out important insights about mechanisms and issues with the project. In the Tanzania case, both project staff and local people felt the structured nature of the process was helpful in prompting them to think about what they wanted the projects to achieve, and why things had worked out in a particular way. As an inclusive process it can feed into adaptive management, highlighting where aspects of the project are not working for people. The method could be applied to wellbeing evaluations to first prioritise aspects of wellbeing which an intervention could target, and then to assess fulfilment of the outcomes — including both observable and subjective aspects.

5.4 Bringing it all together

It is almost impossible to prescribe specific solutions given the multitude of different factors involved in making decisions about the design and methods used in evaluation. However, Figure 5 outlines the interactions between the primary attributes of the intervention, the questions that can be answered and the research design. All evaluations will of course also need to consider the appropriate technical capacity, budget and time needed.

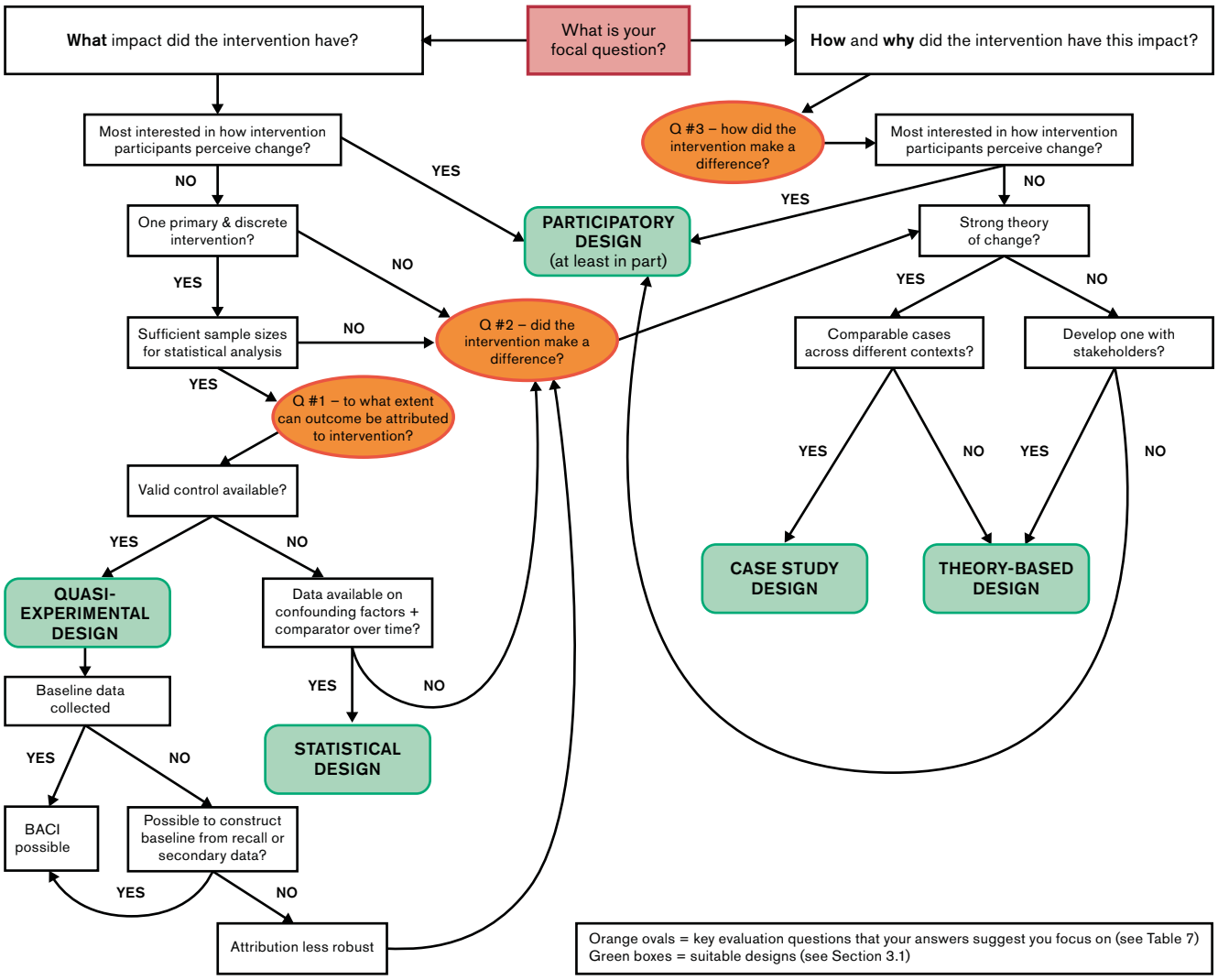


Figure 5. Decision tree: linking questions to evaluation designs

6 Case studies

The range of situations that practitioners can find themselves in is unlimited, and involves navigating a range of issues as outlined above, as well as the messy realities of organisational politics. To better represent the complexity of these situations, aspects of which may chime with readers' own experiences, we outline four case studies which are based on real life situations and offer some potential solutions.

6.1 Case study 1

Short time frame, multiple sites, intangible complex outcomes and a focus on lesson learning

Eli works for a medium-sized European NGO which partners with local NGOs. She is managing the monitoring and evaluation for an integrated livelihoods and conservation programme implemented in six sites across East Africa and South Asia. The programme centres around participatory forest management and aims to produce positive outcomes in sustainable management of forests, livelihoods and income, equity of benefits, participation and empowerment, especially of women and marginalised groups, and the strengthening of civil society groups. There is a common vision and approach to the programme and the sites face similar challenges, but each site has implemented contextually appropriate activities. The first phase of the programme is coming to an end after three years, and the partners wish to establish lessons learned for the next phase, in terms of progress towards aims, successes, and challenges. This mid-term evaluation was not planned for but is now deemed necessary to allow space for learning, refinement of strategies and acknowledgement of errors. There are funds for an external evaluator to lead the process and provide an outsider view of the programme as a whole.

The focus on lesson learning suggests a theory-based approach that can provide evidence for the processes leading to outcomes along a causal pathway — the how and why of change. Given the scale of the project, developing theories of change at different levels is likely to be useful: one overall theory for how the programme works, and then ToCs that are specific to the site context and may make different assumptions and focus on different processes of change: harvesting of non-timber forest products, agroforestry and so on. The latter should be developed by country-level staff alongside local stakeholders and communities who will best know the contextual issues.

Some of the outcomes, such as empowerment, are complex and culturally specific, requiring in-depth knowledge on possible indicators and meanings in the context of the cultural and historical setting. In-depth qualitative work is needed. Although an external evaluator can provide guidance and external legitimacy to the process, the evaluation should rely on local knowledge. The European NGO could possibly draw on expertise from the social science department of a university, either in-country or in Europe, which has worked in these regions and systems before. The large scale of the programme may require that case studies can take place in only a few of the sites, using data from documents, stakeholder and community interviews/focus groups; but could attempt to generate general lessons that can be discussed with site managers to

examine the relevance for their own contexts. The short time that has elapsed from the start of the project means that many outcomes (such as equity of benefits) will not have been achieved, but studies can focus on evidence for progress along the theoretical causal chain — for example the governance structures of community groups and current levels of participation. Some of the outcomes, namely income and those related to livelihoods, could be developed into quantitative indicators. Since there are no baseline data, these could be developed post-hoc from recall data.

The multi-level structure of the programme and the focus on participation and learning would be conducive to a participatory process such as most significant change (Example 6 in Section 3.1), which would complement the theory-based case studies. This would allow stories of change (both positive and negative) to be communicated through the layers of organisation from local people, community groups, field staff, partner offices and the European NGO. Managed well, it could encourage capacity building across these scales, with lessons for all groups shared. Categories of change should align with outputs and outcomes specified in the ToCs, and the process should target women and marginalised groups to ensure that a diversity of experiences is captured.

6.2 Case study 2

Complex intervention landscape, organic development of projects with multiple objectives, no baselines

Susannah works for a small NGO focused on marine conservation and sustainable fisheries in tropical regions. She wants to evaluate the impacts of the NGO's programme of work in one region of a Southeast Asian country. The programme takes an integrated approach to fisheries management, conservation, health (primarily family planning provision and education around water and sanitation), and livelihoods. She is keen to carry out a robust quantitative, statistically valid study to provide proof of concept for the overall approach in the region for advocacy purposes. But she also wants to ensure that the results are relevant to local people, given the community-based approach the organisation takes. There is a complex landscape of interventions, with overlapping projects run by this NGO and others across villages, all with interdependent outcomes. Most of the villages in the area are involved in one or more of the projects, making the identification of control villages for a counterfactual design very difficult. Also, the integrated nature of the projects, in which processes of change and outcomes are connected, makes linking specific causes and effects challenging. There is no baseline data, and although there have been no previous formal evaluations, the NGO has worked in the region for some time and has a good understanding of the social situation and how the projects are working in specific villages.

The attributes of the programme are likely to preclude a quasi-experimental or statistical design. It would be possible to do statistical analyses if there were variation in participation at household level that is thought to affect outcomes. This would involve regression analyses looking at how the outcome indicator is affected by participation, controlling for confounding factors such as wealth levels, household size and positions of power held. The lack of baseline data is a serious constraint on the production of robust results in these designs, however, and would require post-hoc development through available evidence or household recall data. Participation is a particularly tricky variable to use in these circumstances, because it is subject to selection bias

— there may be underlying reasons why people participate in project activities that make them different to others in the village, and invalidates simple comparisons between participants and non-participants. These include tangible things, which could be controlled for in regression analyses, such as higher material wellbeing allowing them the time and ability to participate in activities that take them away from daily activities. But they also include less tangible reasons which are harder to control for, like a proactive mindset and openness to new ideas, which may or may not be linked to factors that can be measured in an indicator. There are also different degrees of participation — these may be fundamental to outcomes but difficult to capture (for instance some people may get a lot more support than others, for various reasons).

A better solution for this kind of complex situation would be a comparative case study design, which could still provide credible results — but with the benefit of being able to untangle the web of factors involved in success and failure in different outcomes, and how outcomes are linked. A strong theory of change for the approach is needed to identify the key outcomes and the set of factors involved in the levels of those outcomes, such as village governance structures, previous interventions, and dominant livelihood types. How are different projects thought to be interacting to affect particular outcomes? Knowledge is held by the organisation and the stakeholders it works with, which could be systematised through workshop discussions with local staff, partners and community-based organisations in order to agree on definitions of the factors involved and prioritise outcomes. Community involvement in these meetings and through further local focus groups would ensure that indicators selected are relevant, and unexpected outcomes captured. Budgetary constraints would inform decisions about how many villages are included. The NGO's knowledge of villages would allow them to make an informed selection of villages to represent different outcome and contextual types. The evaluation would then involve analysis of how different key outcomes (in livelihoods, education, health) are affected by the configuration and interaction of different factors. A ranked outcomes approach (see Example 7) could be used as part of this analysis.

Outcome indicators at the village level could include both: 1) objective, aggregated household measurements based on survey data (such as the percentage over a certain income threshold), which are often more legitimate to external audiences; and 2) subjective, qualitative outcomes based on focus group discussions (feelings about prospects for children's future). It could be possible to construct baselines using recall data on household income, depending on the time frames involved. A simpler method, which would still result in quantitative results, would be to ask whether income has increased on a Likert-type scale, since the implementation of the programme. Other outcomes may be more suitable to measurement at the village level, such as the percentage of children in primary education. Again, baselines could be constructed through expert knowledge in the village. The community-based nature of the programme would require that time is allocated to feedback and discussion of the preliminary results to villages, and to highlight both successful aspects and where efforts need to be concentrated.

In this case, it would be important to manage expectations of what can be achieved in an evaluation, and why a particular approach has been chosen as appropriate. Susannah needs to manage her own expectations and be clear on why this approach is robust, despite not being a statistical or quasi-experimental design, and this understanding needs to be transmitted to external donors and collaborators. This is best done at the planning stage rather than on presentation of the results.

6.3 Case study 3

Small-scale novel project, no baseline or controls, semi-nomadic participants and lack of independence

Joshua is a manager for an East African NGO focused on large predator conservation. Two years ago, the NGO implemented a community-based project that aims to train young local men on wildlife monitoring, conservation and security issues in a conservancy, with the aim of empowering them to participate in conservation decision making, encouraging them to become advocates for wildlife, and to instil a conservation ethic in the wider community. The project is currently at a small scale — at eight village locations in one conservancy. Joshua wants to see how effective the project has been in meeting intended goals, and how improvements can be made, with the intention of investing in expanding the project and leveraging more funding. Monitoring and evaluation had not been incorporated into the project at the design stage so there is no baseline. There is very little funding for the evaluation, which has led to the decision to work with a masters student with a social science background, who is independent of the project. The student is supported by expert supervision from her university. This external evaluation can be used as the basis for future, simpler monitoring processes to be put in place for the future. The only available research assistants who speak the local language have worked for the project previously, and although they have experience of conducting simple questionnaire surveys, are not familiar with other social methods. The men involved in the project and wider community are mobile pastoralists and are largely non-literate.

The focus on learning and participation, and the lack of baseline data, points towards a participatory design as the most appropriate course of action. This would focus on how people attribute changes to the project, and where they see that improvements could be made. Although setting up controls is not entirely necessary with this design, it would be useful to include a site that is a candidate for expansion and is broadly similar, which will provide not only a comparison group but will help identify factors that will be conducive to or constrain scaling up the project. The focus will be to speak to the young men participating in the project, but also the wider community. Eight villages is likely to be too many to include (especially given the scale of the rangelands involved), so Joshua should select fewer (three or four) that are the most diverse, perhaps based on geographical location and exposure to the project. It will be difficult to get around the issue of independence, and attention is needed to the bias this may create in people's responses. The field team should attempt to create an atmosphere of openness and shared learning, by highlighting the intention of the evaluation and showing willingness to act on any identified problems. The process also provides an opportunity to build capacity among local assistants in a wider range of social methods.

To begin, qualitative focus groups with the young men involved, and then wider community members, could aim to discuss social changes that have occurred in the last two years, and how they may be attributed to the project, as well as suggestions for improvements. These discussions can help evaluators to understand the meanings given to, and language used around key intended outcomes such as empowerment and participation, processes of change, and unintended costs and benefits. For instance, is the project improving social cohesion or is its focus on young men alienating other groups? Given the lack of literacy, visual methods such as Venn diagrams in which organisations and projects are represented by circles, with size illustrating relative importance, could support discussions about conservation interventions

and impacts. Based on these discussions, individual questionnaires can be developed that focus in on particular outcomes of interest to understand patterns across different groups. These can be compared to the non-intervention site, which would provide tentative attribution of change if confounding factors were at least considered in the interpretation of results. The mobile nature of the population makes random sampling across households very challenging; instead, local knowledge and communication networks should be drawn on, to create a more purposive sampling frame that captures a range of different people across dimensions of wealth, remoteness, age, and gender.

6.4 Case study 4

Setting up a new project, high levels of funding and capacity, and a determination to make monitoring of impacts central to the project design

Yunus is the in-country technical advisor for a large international NGO, working in a challenging tropical forest landscape with a range of externally driven development pressures underway. He knows the landscape well, as it is where he was brought up, and he has been working in conservation there for over a decade. He has a background in social science, and connections to university researchers in the capital city with both quantitative and qualitative expertise. Over time, Yunus has built relationships with a number of communities, and collected information on their livelihoods, wellbeing and use of natural resources.

The NGO has just won substantial funding for a new landscape-scale conservation project, for an initial five years with the prospect of renewal for a further five years if things are going well. This funding has a specific component for lesson learning and evaluation, and the funder is encouraging innovation and new approaches to working with local people. The project is aimed at supporting community-based participatory land-use planning, the development of premium eco-friendly products from sustainable agriculture, and the setting up of community forest reserves with legal standing. The idea is to work with local people to develop a landscape in which sustainable agriculture and conservation are the sources of ongoing wellbeing, and thereby resistant to external pressures to develop the land, and resilient to social and environmental change.

Yunus is in the enviable position therefore of being able to design an intervention with evaluation in mind, with some baseline information, and with the capacity and funding to produce robust results. He starts by developing a robust theory of change for the intervention, working both with his team and in each of the participating villages, so that everyone understands the goals of the projects, the risks and assumptions underlying it, and the pathway to achieving these goals. He includes key points where evaluation would be useful to track progress, particularly in years four to five in order to report back to the donor; but also interim points where a lighter touch evaluation could check progress and particularly that assumptions are still valid and the project is not being pushed off course.

Yunus decides that since he has a large landscape, and this is a new conservation approach that has not been tried before in this country, he will develop a Before-After-Control-Intervention design for his project. He contemplates a randomised controlled trial, but as the number of villages in the landscape is not that large, and he already has relationships with some villages which would be ideal testing grounds for this rather complex and socially challenging intervention,

he decides that he would rather not randomise the intervention. However, he does match his chosen intervention villages with suitable control villages which are similar in key aspects, such as degree of development pressure and land conversion. He also decides that it is important that his evaluation of project impacts is done by an external and independent assessor, to ensure that the results are credible, and that local people feel able to talk freely to them. This assessor, Sonali, is a researcher at the capital city's university, and Yunus has been careful to work with her in the design phase of the project.

Prior to starting the intervention, Sonali collects detailed baseline information in the intervention and control villages, focused on local understandings of wellbeing, expectations of social impact, and the current livelihoods and wellbeing status of the inhabitants. She is careful to talk to a range of groups, including women and marginalised people. She uses scenario analyses to explore people's aspirations for the future and their expectations of how their behaviour will change over time under different circumstances (including the intervention). This will help to test the theory of change and guide how interventions are structured. She collects a range of data using mixed qualitative and quantitative methods, and explores both the subjective and objective dimensions of wellbeing.

As the intervention is implemented, Yunus maintains close contact with the participants and steers the project based on participatory feedback. Then at the four year evaluation point, Sonali repeats her surveys, with an added focus on experienced change, hopes for the future, and local explanations for why wellbeing has changed over time in their villages and households. Her analysis of this change exploits the BACI design of the project, allowing her to attribute change in wellbeing robustly to the intervention, both statistically and in qualitative terms. She makes recommendations to Yunus about how best to change the project in order to ensure that progress is enhanced over the next five years. She is also careful to go back to the villages where she has worked, to present her general findings to them, and to discuss with them and Yunus how they would like to see things change based on the evaluation, and whether there are things she has missed that they feel are important. The donor is also invited to come and talk to the villagers in their villages, and to the project team, and to hear Sonali's results presented. Satisfied, they agree to extend the project funding for another five years, and to consider funding a similar project in another part of the country.

Acronyms

BACI	Before-After Control-Intervention
BNS	Basic Necessities Survey
CECG	Cambodian Elephant Conservation Group
CSO	civil society organisation
FMG	forest management group
FPIC	free, prior and informed consent
MPA	marine protected area
MSC	most significant change
NGO	non-governmental organisation
OPC	Ol Pejeta Conservancy
PA	protected area
PES	payment for environmental services
PRA	participatory rural appraisal
QCA	qualitative comparative analysis
RCT	randomised controlled trial
REDD	reducing emissions from deforestation and forest degradation
SAPA	social assessment of protected areas
ToC	theory of change
TNP	Tarangire National Park

References

Baird, T D (2014) Conservation and unscripted development: proximity to park associated with development and financial diversity. *Ecology and Society* 19(1) 4. www.ecologyandsociety.org/vol19/iss1/art4

Britton, E and Coulthard, S (2013) Assessing the social wellbeing of Northern Ireland's fishing society using a three-dimensional approach. *Marine Policy* 37 28–36. http://tinyurl.com/hs9atsq

Bull, J W *et al*. (2015) Creating a frame of reference for conservation interventions. *Land Use Policy* 49 273–286. http://www.sciencedirect.com/science/article/pii/S0264837715002409

Catley, A *et al*. (2013) Participatory impact assessment: A design guide. Feinstein International Center, Tufts University, Somerville. http://fic.tufts.edu/assets/PIA-guide_revised-2014-3.pdf

Chambers, R (1997) Whose reality counts? Putting the last first. Intermediate Technology Publications, London.

Clements, T and Milner-Gulland, E J (2015) The impact of payments for environmental services and protected areas on local livelihoods and forest conservation in Northern Cambodia. *Conservation Biology* 29 78–87. www.ncbi.nlm.nih.gov/pubmed/25492724

Danielsen, F *et al*. (2008) Local participation in natural resource monitoring: a characterization of approaches. *Conservation Biology* 23(1) 31–42. www.ncbi.nlm.nih.gov/pubmed/18798859

Davies, R and Dart, J (2005) The 'most significant change' (MSC) technique: a guide to its use. www.mande.co.uk/docs/MSCGuide.pdf

Davies R and Smith, W (1998) The basic necessities survey: the experience of ActionAid Vietnam. ActionAid, Hanoi. www.mande.co.uk/docs/BasicNecessitiesSurveyAAV1998.pdf

Drury, R *et al*. (2011) Less is more: the potential of qualitative approaches in conservation research. *Animal Conservation* 14 18–24. http://onlinelibrary.wiley.com/doi/10.1111/j.1469-1795.2010.00375.x/abstract

Franks, P and Small, R (2016) Social assessment for protected areas (SAPA). Methodology manual for SAPA facilitators. IIED, London. http://pubs.iied.org/pdfs/14659IIED.pdf

Franks, P *et al*. (2014) Social assessment of protected areas: early experiences and results of a participatory, rapid approach. IIED, London. http://pubs.iied.org/pdfs/14643IIED.pdf

Gertler, P J *et al*. (2011) Impact evaluation in practice. The World Bank, Washington DC. http://tinyurl.com/comexto

Gurney, G G *et al*. (2014) Poverty and protected areas: an evaluation of a marine integrated conservation and development project in Indonesia. *Global Environmental Change* 26 98–107. http://tinyurl.com/zg8qln3

Mayoux, L and Chambers, R (2005) Reversing the paradigm: quantification, participatory methods and pro-poor impact assessment. *Journal of International Development* 17 271–298. http://onlinelibrary.wiley.com/doi/10.1002/jid.1214/full Similar content available freely at www.iapad.org/wp-content/uploads/2015/07/Chambers-Mayoux.pdf

McGregor, A and Sumner, A (2010) Beyond business as usual: what might 3-D wellbeing contribute to MDG momentum? *IDS Bulletin* 41 104–112. http://tinyurl.com/zynoo6j

Narayan, D *et al*. (2000) Voices of the poor: crying out for change. Oxford University Press for the World Bank, New York.

Nuno, A and St John, F (2014) How to ask sensitive questions in conservation: A review of specialized questioning techniques. *Biological Conservation* 189 5–15. https://kar.kent.ac.uk/46119

Oxfam (2011) Draft Protocol for Process Tracing. http://tinyurl.com/hcso29f

Palmer Fry, B (2014) Navigating the complexities of community monitoring, reporting and verification (CMRV). PhD Thesis. Imperial College London. www.iccs.org.uk/wp-content/uploads/2015/07/BPF-full-thesis-corrected-1.pdf

Pretty, J N (1995) Participatory learning for sustainable agriculture. *World Development* 23(8) 1247–1263. http://tinyurl.com/gsjub38

Punton, M and Welle, K (2015) Straws-in-the-wind, hoops and smoking guns: what can process tracing offer to impact evaluation? Institute of Development Studies, Brighton. http://tinyurl.com/jdk5syt

Sainsbury, K *et al*. (2015) Exploring stakeholder perceptions of conservation outcomes from alternative income generating activities in Tanzanian villages adjacent to Eastern Arc Mountain forests. *Biological Conservation* 191 20–28. http://tinyurl.com/gpxwnxj

Sehring, J *et al*. (2013) Qualitative Comparative Analysis (QCA): an application to compare national REDD+ policy processes. CIFOR, Bogor, Indonesia. www.cifor.org/publications/pdf_files/WPapers/WP121Sehring.pdf

Schatz, F and Welle, K (2016) Qualitative Comparative Analysis: a valuable approach to add to the evaluator's toolbox? Lessons from recent applications. Institute of Development

Studies, Brighton. http://tinyurl.com/hm6sefd

Schreckenberg, K *et al*. (2010) Social assessment of conservation initiatives: a review of rapid methodologies. IIED, London. http://pubs.iied.org/pdfs/14589IIED.pdf

Stedman-Bryce, G (2013) Health for all: towards free universal health care in Ghana – end of campaign evaluation report. Oxfam GB, Oxford. http://tinyurl.com/jkylo2s

Stern, E (2015) Impact Evaluation: a guide for commissioners and managers. BOND, London. https://www.bond.org.uk/data/files/Impact_Evaluation_Guide_0515.pdf

Stiglitz, J *et al*. (2009) Report by the commission on the measurement of economic performance and social progress. CMESP, Paris. www.stiglitz-sen-fitoussi.fr/documents/rapport_anglais.pdf

Summers, J K *et al*. (2012) A review of the elements of human wellbeing with an emphasis on the contribution of ecosystem services. *Ambio* 41 327–340. www.ncbi.nlm.nih.gov/pubmed/22581385

UN-REDD Programme (2013) Guidelines on free, prior and informed consent. www.uncclearn.org/sites/default/files/inventory/un-redd05.pdf

Wilder, L and Walpole, M (2008) Measuring social impacts in conservation: experience of using the Most Significant Change method. *Oryx* 42(4) 529–538. http://tinyurl.com/glpymff

Conservationists are increasingly seeing the importance of carrying out social impact evaluation to ensure accountability and to learn what works for both biodiversity and human wellbeing. A single toolkit or blueprint method cannot fit the diversity of intervention types and evaluation questions, and conservationists are faced with an array of decisions about the most appropriate methods and research designs to use. This guidance aims to demystify the process of social impact evaluation and support practitioners in navigating through these methodological decisions taking into account the questions the evaluation aims to answer, the characteristics of the intervention, and the organisational capacity and resources available. It takes practitioners through the key steps in an evaluation: 1) thinking through the aims of the evaluation; 2) defining relevant wellbeing outcomes and indicators; 3) designing the evaluation to link outcomes to the intervention; and 4) collecting data including applying methods to account for bias, social dynamics and ethical considerations. The guidance provides a range of real life case studies and ideas for appropriate methods and tools.

iied

International Institute for Environment and Development
80-86 Gray's Inn Road, London WC1X 8NH, UK

www.iied.org
@iied
www.facebook.com/theIIED

Download more publications at http://pubs.iied.org
Printed on recycled paper with vegetable-based inks.

Knowledge Products

Tool

September 2016

Biodiversity

Keywords:
Conservation; protected areas; wellbeing; social impact evaluation

UKaid from the British people

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

We gratefully acknowledge funding from the ESRC-DFID joint fund for poverty alleviation (ES/J018155/1) 'Measuring complex outcomes of environment and development interventions' and an ESRC Impact Acceleration award to Imperial College London.