Introduction: the evolution of evidence-based approaches in biodiversity and poverty research and policy

Over the last decade, donors and decision makers have become increasingly concerned about the likelihood that the policies and projects they support will succeed (Campbell, Benita et al. 2007). Where once received wisdom or a convincing logical case was sufficient, there is now an increasing requirement for robust evidence that an intervention works before it is supported. Decisions taken in the light of such evidence are considered to be evidence-based, and it is now common to hear this terminology applied to policy and practice in a range of different fields.

The concept of evidence-based practice originates from medicine, where the increasing use of evidence to inform decisions is believed to have underpinned a revolutionary improvement in performance (reviewed by Pullin and Knight 2001). Evidence-based medicine has made wide use of random, replicated trials that have blind controls and are analysed with statistics. This approach is intended to provide objective evidence of the effectiveness of an intervention (such as a drug) that can then be used as the basis for decision making. In conjunction with information about cost, this approach is used by the UK National Institute for Clinical Excellence (NICE) when it develops guidance on which medical procedures and treatments should be funded through the National Health Service.

The concept of evidence-based policy and practice is intuitively appealing, and has rapidly gained popularity in a range of different domains – including in international development and biodiversity conservation. For example, the UK Department for International Development (DFID) has renamed its research department “Research and Evidence” and now offers financial support for systematic reviews of available evidence (http://www.dfid.gov.uk/what-we-do/research-and-evidence/). In biodiversity conservation there has, over the last 10 years, been a push for “evidence-based conservation” that has led to a number of new journals and databases intended to help practitioners to access systematic review results and published evidence (Pullin and Knight 2003; Sutherland, Pullin et al. 2004; Pullin and Salafsky 2010; Segan, Bottrill et al. 2011).

Scope and objectives of this discussion paper

It is clearly desirable for important decisions to take account of available information. However, the relationship between ‘evidence’ and good decisions is not always straightforward. Indeed, an emerging literature identifies a range of challenges with the evidence-based approach to policy and practice (Fazey,
Salisbury et al. 2004; Fazey, Fazey et al. 2006; Elgert 2010; Hagen-Zanker, Duvendack et al. 2012. Sandbrook and Adams, submitted). These include:

- How to deal with different sources of evidence? Are some better than others?
- How important are controls / counterfactuals?
- How does evidence get taken up and translated into policy?

This paper is intended to stimulate discussion – and solicit feedback – on these challenges and how to address them. The paper has been produced as a component of the “Biodiversity, Ecosystem Services and Poverty Alleviation: Assessing the Current State of the Evidence” project, funded by an Evidence and Impact Research Grant, under the DFID-NERC-ESRC Ecosystem Services and Poverty Alleviation (ESPA) research programme. The main aim of this project is to conduct a systematic review of the state of knowledge on the relationship between these issues. In conducting the systematic review, the question of how to deal with different forms of evidence will be critical.

This paper provides an overview of the academic literature on evidence, with a particular emphasis on the question of what constitutes good evidence. It begins with a review of the different sources and types of evidence for links between biodiversity, ecosystem services and poverty alleviation. It then reviews different approaches to assessing the quality of evidence, including hierarchies, matrices, and the question of controls. The way evidence is actually taken up into policy is then considered, before a final summary of issues to be considered for this project.

**What is evidence and where does it come from?**

What is evidence? Wikipedia defines evidence as “everything that is used to determine or demonstrate the truth of an assertion” (http://en.wikipedia.org/wiki/Evidence). This leaves a lot of flexibility in determining what sort of information might qualify as evidence for any given assertion. When considering biodiversity, ecosystem services and poverty alleviation, there are many potential sources of information. In this paper, the term ‘ultimate’ sources of evidence refers to the different approaches to knowledge generation that can be used to generate evidence. These range from formalised science through to informal and ‘local’ knowledge (Table 1). The term ‘proximate’ sources refers to the specific ways in which evidence can be accessed (Table 2).

<table>
<thead>
<tr>
<th>Source of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Professional science</td>
</tr>
<tr>
<td>Citizen science</td>
</tr>
<tr>
<td>Expert knowledge</td>
</tr>
<tr>
<td>Local knowledge</td>
</tr>
<tr>
<td>Indigenous knowledge</td>
</tr>
</tbody>
</table>
Table 2: A selection of proximate sources of evidence for biodiversity, ecosystem services and poverty alleviation

<table>
<thead>
<tr>
<th>Source of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Systematic reviews and meta-analyses</td>
</tr>
<tr>
<td>Peer reviewed journal articles</td>
</tr>
<tr>
<td>Peer reviewed books</td>
</tr>
<tr>
<td>International ‘grey’ literature</td>
</tr>
<tr>
<td>National / local ‘grey’ literature</td>
</tr>
<tr>
<td>Internet articles and blogs</td>
</tr>
<tr>
<td>Oral communication</td>
</tr>
<tr>
<td>Personal observation / research</td>
</tr>
</tbody>
</table>

Campbell et al. (2007) offer a typology of evidence that relates to the methods used to collect information. This is based on a survey of UK government officials, and the categories are overlapping and by no means exhaustive. The categories they identified are:

- Quantitative / statistical evidence
- Economic evidence
- Surveys, attitudinal and behavioural evidence
- Qualitative evidence
- Anecdotal evidence
- Scientific evidence
- International evidence
- Social experiments / controlled trials
- Systematic reviews / meta-analyses
- Consultations

Raymond et al. (2010; p. 1769) present a graphical representation of the relationship between different dimensions and types of knowledge (Figure 1). This is a useful approach to show the relationships between the different ultimate sources of knowledge shown in Table 1. By highlighting the differences between types of evidence, it begins to raise important questions about which types of knowledge provide the ‘best’ evidence, which is the focus of the following section.
What constitutes ‘good’ evidence?

The previous section makes it clear that there is a very wide range of ways of generating knowledge about the relationship between biodiversity, ecosystem services and poverty alleviation, and an even wider range of ways to access such knowledge. The challenge then for those attempting to use an evidence-based approach is to decide what kind of evidence is ‘good’ evidence. This is clearly not a black and white issue. In the evidence-based medicine literature, a lot of emphasis in answering this question is put on the type of methodology used to gather evidence. This gives the most weight to evidence derived from quantitative, randomised, replicated, controlled trials, and progressively less weight to different forms of evidence that are qualitative or do not have controls. To date, this approach has also been favoured by those promoting evidence-based conservation (Sutherland, Pullin et al. 2004; Segan, Bottrill et al. 2011).

Margoluis et al. (2009) provide a structured analysis of the strengths and weaknesses of different forms of evidence for evaluating the impact of conservation projects (Table 3). Again, this generally gives more credit to quantitative and experimental approaches, and less to qualitative approaches.

Table 3: Types of evaluation design. After Margoluis et al. (2009)
Quantitative Design

1. **Experimental**: random assignment of subjects to treated (experimental) and untreated (control) groups
   - **Advantages**: approximates counterfactual condition; strong evidence for causality
   - **Limitations**: expensive; often not practical; ethical issues; high expertise
   - **Validity**:
     - Internal: high; random assignments; strongest design for internal validity
     - External: low; artificial setting limits ability to generalize to other settings
   - **Example**: Randomized pre and post: researcher randomly assigns items into control and experimental groups. Measurements taken before and after intervention

2. **Quasi-experimental**: similar to experimental but lacks random assignment
   - **Advantages**: easier to establish than true experimental designs; fairly strong evidence for causality
   - **Limitations**: moderately expensive to expensive
   - **Validity**:
     - Internal: moderate; inability to randomly assign controls, lack of control over variables
     - External: moderate; “natural experiments” allow some generalization
   - **Examples**:
     A. **Matched controls**: intervention group matched with controls selected by researcher
     B. **Regression-discontinuity**: pretest/posttest design in which participants are assigned to program or comparison groups on the basis of a cutoff score on a preprogram measure
     C. **Statistically equated controls**: exposed and unexposed groups or items compared by means of statistical controls
     D. **Generic controls**: exposed group or items compared with outcome measures available on general population

3. **Noneperimental**: draws inferences about the effect of a treatment on subjects, where assignment of subjects into a treated versus control group is outside the researcher’s control
   - **Advantages**: least expensive quantitative design; easier to implement
   - **Limitations**: observe state of world without manipulating it, so less power to detect causal relationships
   - **Validity**:
     - Internal: low; no randomization, no controls
     - External: moderate; natural settings make generalizability stronger
   - **Examples**:
     A. **Pretest/posttest**: subjects measured before and after intervention
The natural conclusion of the decision to give greater weight to particular forms of evidence is that there exists an ‘evidence hierarchy’. This is explicitly the case in much writing on evidence-based medicine. For example, Petticrew and Roberts (2003) consider the following to be a standard evidence hierarchy for medicine\(^1\) (p.527):

1. Systematic reviews and meta-analyses
2. Randomised controlled trials with definitive results
3. Randomised controlled trials with non-definitive results
4. Cohort studies
5. Case-control studies
6. Cross sectional surveys
7. Case reports

Whilst the hierarchy approach is appealing, it has been noted that the most appropriate form of evidence will vary depending on the question that is being asked (Petticrew and Roberts 2003). So for example, a randomised replicated trial might be an appropriate method for answering a question about the fundamental ecology of an ecosystem, but a qualitative, case study approach might be more appropriate.

\(^1\) It should be noted that this is not their own hierarchy, and that Petticrew and Roberts’ paper is critical of the hierarchy approach
for a question about the influence of political incentives on a system of governance (Adams and Sandbrook Submitted). One approach that combines the underlying notion of a hierarchy of evidence with the recognition that this hierarchy will vary with the nature of the question is the ‘evidence matrix’ or ‘evidence typology’ (Petticrew and Roberts, 2003; Figure 2).

![Figure 2: An example of a typology of evidence, after Petticrew & Roberts 2003](image)

It is often argued that to be convincing, evidence must be based on the comparison of controls to test cases. This follows the logic that without a control it is not possible to know whether an observed effect is due to an intervention or some other third-party factor. The matrix approach introduces the possibility that the need for evidence with controls will be greater for some questions than others. So for example, a scientific question about ecology might be answered through an experimental manipulation involving controls, whereas a question about the influence of political incentives may not be amenable to an experiment-with-controls approach.

Another factor that can be used to assess the strength of evidence is the degree of consistency or agreement between different sources. This approach is used by the IPCC in its assessments of the evidence for climate change (Mastrandrea, Field et al. 2010). Figure 3 shows the way in which the IPCC combines information on the type, amount, quality and consistency of each source of evidence with the agreement between sources to identify the level of confidence provided by the overall evidence available.
Figure 3: A depiction of evidence and agreement statements and their relationship to confidence. Confidence increases towards the top-right corner as suggested by the increasing strength of shading. Generally, evidence is most robust when there are multiple, consistent independent lines of high-quality evidence (after Mastrandrea et al. 2010, p. 3)

**Integrating evidence from different sources**

A particular challenge when assessing evidence is how to incorporate information from a range of different proximate and ultimate sources, as described in Tables 1 and 2. The hierarchy, matrix and agreement approaches described in this section are mostly targeted at evaluating scientific evidence that has been generated using different methodologies. This is challenging, but nowhere near as challenging as deciding how to incorporate indigenous knowledge that may be based on an entirely different worldview (e.g. West 2005; Raymond, Fazey et al. 2010). This is a problem that is currently being addressed by the new Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services (IPBES), which has set out to incorporate local and indigenous knowledge into its assessments (Tengö, Kvarnström et al. 2011). Raymond et al. (2010) offer a framework to follow for integrating different forms of knowledge into environmental management projects (Figure 4). This is useful, although not entirely relevant to the particular challenge of evaluating evidence in isolation from a particular project.
The most common approach used to integrate evidence for evidence-based policy or practice is the systematic review. These are highly structured reviews of existing evidence that follow a defined methodology, making them replicable and reducing the risk of researcher bias influencing the findings of the review. Before conducting a systematic review, the reviewer(s) must make a decision about what kind of evidence to include. This decision will inevitably be based on two factors. First, there are theoretical questions regarding the most appropriate forms of evidence for the question in hand (as reviewed above). Second, there are pragmatic constraints placed upon the reviewer by the time and resources they have available to them (Hagen-Zanker, Duvendack et al. 2012). Often it is this second factor that determines the extent to which more difficult sources of evidence, such as indigenous knowledge, are incorporated into reviews, because it simply isn’t possible to go and ask indigenous people for their views given the resources available. Similarly, some scientific sources, such as peer-reviewed journal articles, may not be available for review because the reviewer does not have access to the full text. As a result, systematic reviews can give the impression of being comprehensive and objective, when in fact they may be rather limited. They also face problems in determining how to evaluate qualitative and quantitative data (Hagen-Zanker, Duvendack et al. 2012).

From evidence to policy and practice

In order to establish what constitutes good evidence, it is necessary to consider not only where evidence comes from and how likely it is to be accurate, but also how likely it is to be translated into policy and practice (Adams and Sandbrook Submitted). This process is affected by several factors. First, some
evidence may be particularly amenable to uptake by decision makers, whereas other evidence may be much less so. This may be for the same reasons of accessibility outlined above, or it might be because evidence that supports a decision-makers pre-existing ideas is more likely to be favoured than contradictory evidence (Roe 1991). This undermines the concept of objective evidence-based decision making. Second, it has been noted by many scholars (e.g. Keeley and Scoones 2003; Jones 2009) that there is not a linear process leading from good evidence to good policy or practice, as some of those producing the evidence seem to expect. In practice, most decisions are taken on the basis of multiple factors, and these might include resource constraints, political consequences and historical factors, as well as evidence of what works or does not work. Many decisions are also deliberative, and taken on the basis of group debate or votes rather than a direct evaluation of evidence alone. This is certainly the case in many situations relevant to biodiversity, ecosystem services and poverty alleviation.

The research community often calls for decision makers to improve their understanding of science and to make more use of evidence from scientific sources (e.g. Pullin and Knight 2003). On the other hand, there have been calls for researchers themselves to gain a better understanding of how the decision-making process works, in order to help them to deliver more useful and policy-relevant research. For example, Jenkins et al. (2012) call for conservation researchers to gain “embedded experiences”. They argue that spending “an intensive period enmeshed in the culture and operations of other work communities allows scientists to bridge the gaps between research outputs and policy change, and research outputs and conservation impact.” (p. 740).

In light of the recognition that there is no simple linear process leading from evidence to good decisions, some commentators have called for ‘evidence-informed’ policy and practice (Nevo and Slonim-Nevo 2011; Adams and Sandbrook Submitted). This approach continues to support the use of evidence in decision making, but argues that it is often undesirable, and unrealistic for decisions to be based on evidence alone. In general the critical literature on the use of evidence in international development is better developed than for biodiversity conservation, where there are relatively few critical pieces (e.g. Fazey, Salisbury et al. 2004; Fazey, Fazey et al. 2006; Adams and Sandbrook Submitted).

Conclusions and lessons learned

This discussion paper has considered the question of what evidence is in the context of the linkages between biodiversity, ecosystem services and poverty alleviation, has reviewed a range of approaches to determining the quality of different sources of evidence, and has considered the uptake of evidence into policy and practice. What lessons can be learned for the treatment of evidence in this context?

First, a very broad definition of evidence should be applied when considering biodiversity-poverty linkages, which does not automatically exclude information from particular ultimate sources like local knowledge. Putting this into practice may require making an effort to identify evidence from an unusually wide range of proximate sources, including the grey literature. This increases the likelihood of capturing evidence in the broadest sense, but it is also likely to increase costs.
Second, the diversity of sources and types of evidence for biodiversity – poverty linkages suggests that it is not sensible to apply a strict evidence hierarchy in this context, as might be appropriate in certain aspects of medicine. However, that does not mean that all forms of evidence are equal, for two reasons. First, for certain specific questions that might be asked about the biodiversity – poverty relationship, there may be forms of evidence that are more appropriate than others, and therefore an evidence hierarchy will exist for that question. Second, two sources of the same kind of evidence may differ in their quality, for example because of differences in the quality of execution of a particular methodology.

Third, it is clear that despite efforts to systematise the collation and analysis of evidence, the process retains an element of subjectivity. This is particularly true in the context of biodiversity – poverty linkages where a reviewer may be required to compare quantitative evidence from western science with qualitative evidence derived from an indigenous knowledge system. This requires the good judgement of the reviewer. Placing too much attention on evaluating evidence in a way that is systematic and replicable inevitably leads to exclusion criteria that would discard the latter source of evidence in favour of the former, a process that may ultimately undermine the quality of the review.

Finally, the complexity and diversity of evidence makes the process of translating evidence into policy and practice equally complex. There is no simple linear pathway between the two, and in fact the relationship is two-way and influenced by all kinds of external factors. It is unrealistic to expect evidence to be taken up directly by decision makers, particularly in contexts where decisions are likely to be deliberative. On this basis the concept of ‘evidence-informed’ rather than ‘evidence-based’ policy and practice has merit.

**Your Help is Needed!**

This paper is intended to provoke discussion and we would be very interested in your feedback so that we can improve it. In particular we are interested in your thoughts on the following questions:

1. Are studies based on randomized control trials the “gold standard” when reviewing linkages between biodiversity and poverty?
2. What is your interpretation of what constitutes good and bad evidence relevant to biodiversity-poverty linkages?
3. What is your understanding of what kinds of studies/methods are and are not rigorous?
4. How can evidence from different sources best be integrated?

We are also interested in your feedback on the specific contents of the paper. Please send any comments to Chris Sandbrook (cgsandbrook@gmail.com) and Dilys Roe (dilys.roe@iied.org).

**References**


