USAID
FROM THE AMERICAN PEOPLE

EVIDENCE IN ACTION

USING AND GENERATING EVIDENCE ABOUT EFFECTIVENESS IN BIODIVERSITY PROGRAMMING

Unit 2: Using Evidence

APRIL 2018

Prepared by: Natalie Dubois, Environmental Incentives; Andrés Gómez, ICF

Front Cover: La Reserva Bosque Nuboso Santa Elena is one of the first community reserves in Costa Rica. Photo credit: Ulf Rydin.

Back Cover: Habitat loss is a secondary threat to rhino species already threatened by poaching. Photo credit: Michelle Gadd/USFWS

The authors' views expressed in this publication do not necessarily reflect the views of the United States Agency for International Development or the United States Government.

Acknowledgements: The authors would like to acknowledge Diane Russell for her vision, guidance, and contribution to *Evidence in Action.* Elizabeth Daut and Stacey Young provided helpful comments on a draft of this unit..

About Measuring Impact: This work is made possible by the generous support of the American people through the United States Agency for International Development (USAID) under the terms of its requisition number REQ-EGAT-12-000014 (Measuring Impact) implemented by Environmental Incentives, LLC; Foundations of Success; and ICF International. Measuring Impact has been issued under contract number AID-OAA-C-12-00078 and supports the same program objectives as described in RFP number SOL-OAA-000050. Measuring Impact is funded and managed by the USAID Bureau for Economic Growth, Education, and Environment Office of Forestry and Biodiversity.

TABLE OF CONTENTS

1. OVERVIEW 5

2. INTRODUCTION 6

3. ASKING QUESTIONS OF THE EVIDENCE BASE 12

4. GATHERING EVIDENCE 20

5. APPRAISING THE QUALITY AND STRENGTH OF EVIDENCE 28

6. USING EVIDENCE IN USAID BIODIVERSITY PROGRAMS 33

7. SUMMARY OF KEY CONCEPTS 39

8. FURTHER READING 40

ENDNOTES 41

GLOSSARY 42

LITERATURE CITED 45

LIST OF BOXES

Box 1: What Is Evidence? 7

Box 2: Sources and Types of Evidence Used in Biodiversity Programming 8

Box 3: Understanding Assumptions 14

Box 4: Understanding Risk and Hazard 16

Box 5: Consulting the Literature 20

Box 6: Evaluating Scholarly Journals 25

Box 7: Implementing a Strategic Approach When the Team Knows the Evidence Base is Weak 38

LIST OF EXAMPLES

Example 1: Using Evidence in a Situation Model for the Grand River Example 10

Example 2: Key Assumptions in the Design of a Cookstove Program 17

Example 3: Using Evidence to Improve Local Governance in Community-Based Natural Resource Management 34

LIST OF TABLES

Table 1: Types of publications teams may encounter when searching for evidence 22

LIST OF FIGURES

Figure 1: An extract from a situation model identifying threats and drivers affecting river fish populations in the fictitious Grand River basin 10

Figure 2: An illustration of a fictitious data set that would support the assumption that overfishing of juveniles is associated with the stock status of fished populations 11

Figure 3: Using a risk framework to identify priorities for gathering evidence about the validity of program assumptions 15

Figure 4: Depiction of the problem analysis for a cookstove program 17

Figure 5: Common cognitive biases 21

Figure 6: Results chain for a community-based natural resources management strategic approach 34

Figure 7: Four possible program outcomes that a team should consider when the validity of key assumptions is unknown 37

ACRONYMS

ADS Automated Directives System

USAID United States Agency for International Development

1. OVERVIEW

Healthy rivers, forests, and oceans are essential to development, as they support and sustain livelihoods and human well-being. Conservation protects the biological resources that people depend on and that are a critical component of good development outcomes. To this end, the United States Agency for International Development (USAID) has made significant investments in mitigating threats to biodiversity in key ecosystems and landscapes.

Faced with finite resources and great demand, it makes sense to ask tough questions about the effectiveness of biodiversity programs. It is not only important to know if a program achieved its expected outcomes; it is also important to understand how and why a program achieves success. Using and generating evidence about what works, what doesn't, and in which contexts can help teams make better programming decisions. *Evidence in Action* helps mission staff and implementing partners use and generate evidence about the effectiveness of biodiversity programs. The resource is presented in four units that can be used alone or as a series. A glossary defining key terms is included with each unit.

- *Unit 1: Understanding an Evidence-Based Approach* provides an introduction to evidence and evidence-based approaches to biodiversity programming in the context of the USAID Program Cycle.
- **This second unit: *Using Evidence* focuses on the critical review and use of evidence to increase the effectiveness of biodiversity programs.**
- *Unit 3: Generating Evidence* identifies Program Cycle processes that teams can use to generate credible evidence about the effectiveness of biodiversity programs.
- *Unit 4: Building the Evidence Base* highlights ways in which evidence can be shared and applied to strengthen biodiversity programs across USAID.

2. INTRODUCTION

USAID program managers and implementing partners make decisions that need to be supported with evidence (see Box 1 on page 7) throughout implementation of the USAID Biodiversity Policy and Program Cycle. Program effectiveness relies on addressing the need for evidence at three critical stages of design and implementation: the problem analysis, the development solution, and implementation.[1]

- ***When analyzing the development problem***, teams need evidence to support assumptions about the influence of threats and drivers on the status of biodiversity focal interests. If the wrong threats and drivers have been identified, implementing a strategic approach designed to change one or more of those threats or drivers will fail to generate desired outcomes.
- ***When identifying the development solution***, teams need evidence to support assumptions underpinning the relationships between results along the theory of change for a particular strategic approach. Validating these assumptions gives teams more certainty that achieving intermediate results will lead to conservation outcomes.
- ***When developing implementation plans***, teams need evidence that the actions they are planning will achieve desired results in their program context.

This unit of *Evidence in Action* helps teams understand how and when to use evidence throughout the Program Cycle.

Hatcheries can be an effective approach for protecting endangered sea turtles. Photo credit: Arpandhar

BOX 1: WHAT IS EVIDENCE?

Automated Directives System (ADS) Chapter 201 defines evidence as the "[b]ody of facts or information that serve as the basis for programmatic and strategic decision-making in the Program Cycle. ... [Evidence] can be sourced from within USAID or externally and should result from systematic and analytic methodologies or from observations that are shared and analyzed" (page 145).

The term "evidence" refers to both (1) individual findings or pieces of information used to help make a decision or support a conclusion; and (2) the body of findings or information providing support for (or countering) a belief or claim.

Evidence can be generated through primary research, literature reviews, case studies, assessments, evaluations, and performance monitoring. Evidence for program effectiveness comes from real-world observations and documentation of program outcomes. Observations are not considered evidence unless they are used to investigate whether a belief or claim is true.

Evidence can be gleaned from a number of sources (see Box 2 on page 8):

- Peer-reviewed published research, especially literature reviews and systematic reviews, which synthesize evidence across multiple data sources
- Grey literature such as donor and non-governmental organization reports
- Evaluations and monitoring reports including summaries of evaluation findings (e.g., USAID Bureau evaluation summaries)
- Tools and syntheses that collect practitioner experience (e.g., "stocktaking" guides and reports)

Field visits like this one to a Feed the Future activity in Kenya are a source of evidence drawn from practitioner experience. Photo credit: Muthoni Njiru/AVCD

Within the Program Cycle, different actors are responsible for gathering evidence. During project or activity design, mission staff are more likely to be directly involved in gathering evidence – through literature searches and other means – than they will be post-procurement. During and directly following procurement, the role of mission staff will shift to the review and critical appraisal of the use of evidence in proposals, implementation plans, and other documents submitted by partners.

The same rigor that USAID staff apply to using evidence in project and activity design should be expected from implementing partners. During implementation, evidence can be acquired through monitoring and evaluation of the activity (see Section 4 in *Unit 3: Generating Evidence*) as well as through targeted assessments and site visits. Adaptive management – refining and revising strategies to adjust to new understanding and realities – is built on a sound evidence base.

BOX 2: SOURCES OF EVIDENCE USED IN BIODIVERSITY PROGRAMMING

Evidence can come from many sources including peer-reviewed scientific journal articles, grey literature reports and notes, program evaluations, and expert understanding. An evidence-based approach should not limit program managers to using evidence from quantitative studies. Both qualitative and quantitative studies can be appropriate sources of evidence in biodiversity programming.

The advantage of peer-reviewed literature is that study designs and conclusions have been reviewed by independent observers. However, the focus of available studies may or may not provide appropriate evidence for the information need or question being asked. Users should be critical consumers of evidence by recognizing the factors that contribute to the quality and strength of evidence. Different research designs may provide stronger or weaker support depending on the type of question being asked.

Reports and evaluations are also useful resources on program effectiveness, and they are increasingly included in systematic reviews on conservation topics. While such grey literature can contain valuable information for program managers, it is important to remember that the study designs and conclusions have not been through the same review process as findings from the published scientific literature. However, being peer-reviewed does not by itself guarantee that evidence is robust or trustworthy, and critical appraisal of the evidence is merited.

WHY USE EVIDENCE ABOUT EFFECTIVENESS?

When teams design programs for biodiversity conservation, they make a series of decisions about what the problem is, the solutions available to address it, and how to implement those solutions. Without evidence about what works and what doesn't work, teams may miss opportunities to replicate successes and risk continuing to invest in programs with a low track record of success.

Missions that implement evidence-based approaches to biodiversity programming recognize that ignoring evidence can have costs. Implementing ineffective strategic approaches where biodiversity focal interests are at risk – even if that strategic approach does not make matters worse – is an inefficient use of resources and may be harmful if it results in a missed opportunity to effect positive development and conservation outcomes.

Using an evidence-based approach is a means to improving program outcomes, not an end in itself. Adopters of evidence-based approaches employ a critical thinking mindset to uncover practices that may be based on unfounded beliefs. They are willing to seek out evidence that will support or refute program assumptions.

Unit 2: Using Evidence covers topics related to using the evidence base to increase the effectiveness of biodiversity programs. The unit is organized around three topics:

1. Identifying information needs and setting priorities for gathering evidence
2. Gathering and appraising evidence related to program effectiveness
3. Applying evidence to programming decisions

After completing this unit, teams will be able to:

- Critically evaluate the use of evidence in program design and implementation decisions
- Access resources that expedite retrieval of existing evidence from internal and external sources
- Use an evidence-based mindset to make programming decisions even when the existing evidence base is limited

EXAMPLE 1: USING EVIDENCE IN A SITUATION MODEL FOR THE GRAND RIVER EXAMPLE

Biodiversity How-To Guide 1: Developing Situation Models in USAID Biodiversity Programming illustrates the development of a situation model as a part of a problem analysis undertaken by a fictitious design team working in the Grand River basin. The team uses the situation model to convey the most important direct threats and drivers thought to affect biodiversity focal interests. A portion of the situation model identifying suspected drivers leading to overfishing of river fish populations is depicted in Figure 1.

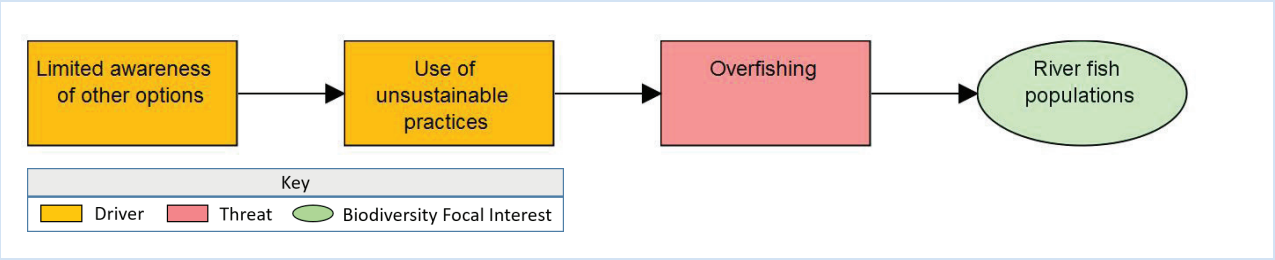

Figure 1: An extract from a situation model identifying threats and drivers affecting river fish populations in the fictitious Grand River basin

During situation model development, the team started with a question, "Is the use of unsustainable fishing practices a primary driver of overfishing in the Grand River basin?" Interviews with fishers in the Grand River basin and fishing boat logs indicated that local fishers were using fishing methods that resulted in bycatch of large numbers of juvenile fish. The team suspects that overfishing of juveniles is a primary driver of changes in the population structure of slow-growing fish species in the Grand River basin.

A search of USAID documents reveals that a previous activity monitored the size of fish stocks in relation to exploitation pattern (the proportion of immature fish in the total catch) in watersheds in the Grand River basin (see Figure 2 on page 11).[2] These findings support the assumption that stock status is lower in watersheds with higher fishing mortality of immature fish compared to watersheds with lower fishing mortality of immature fish, but does this mean that take of immature fish is driving current stock status? At a minimum, the team would want to look for evidence confirming that the use of the fishing methods in question coincided with the onset of declines in stock status. They should also consider the extent to which the study is able to address other plausible explanations for the observed variation in stock status across the watersheds that are unrelated to fishing method.

How do assumptions in the problem analysis influence program effectiveness? The theory of change is linked to the threats and drivers identified in the problem analysis. In the Grand River example, the team identified the use of unsustainable fishing practices as leading to overfishing. This assumption is important to test because if overfishing is the result of other drivers such as lack of economic opportunity, then implementing a strategic approach that is successful in getting fishers to adopt new practices will be unlikely to achieve conservation outcomes.

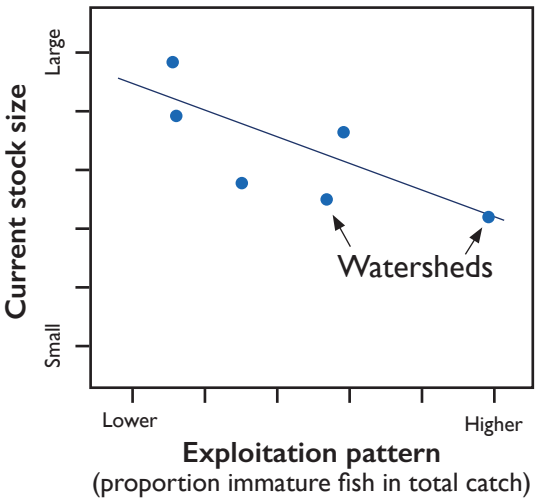

Figure 2: An illustration of a fictitious data set that would support the assumption that overfishing of juveniles is associated with the stock status of fished populations

3. ASKING QUESTIONS OF THE EVIDENCE BASE

Before a team starts searching for evidence, they need to know the right questions to ask. Gathering evidence requires an investment in time and resources, so it is important that teams prioritize searching for evidence that is useful to program decisions such as site selection, focal interests, threat identification and ranking, and choice of strategic approaches. Each of these decisions requires teams to review evidence and determine how to fill information needs as they design and implement.

An information need (Case 2012) is a recognized disparity or gap between what is known and what would ideally be known for a program design or adaptive management decision. The information need articulates what is unknown about a particular problem.

Recognizing how different information needs relate to different types of questions is an important part of an evidence-based approach (Booth 2006). In the early stages of program design, teams are likely to ask more information-gathering questions – or "background" questions – about the problem, the problem context, and the types of solutions that have been applied. They should ask background questions before articulating their assumptions (see Box 3 on page 14) about what will work in their context. They should also not assume that previously used approaches will be effective and that their understanding of the local context still applies.

Background questions could include:

- What political, social, and cultural factors influence the use of forest resources in the Amazon rainforest?
- What strategic approaches have been used to address overexploitation of forest resources in the Amazon rainforest?

As teams increase their understanding of the problem context and viable programming options, their information

needs move toward seeking evidence to verify their assumptions about how the program will work.[3] These different information needs change the type of question being asked. "Foreground" questions elicit evidence that is intended to address assumptions in the problem analysis, development solution, or implementation plan – all of which underlie program effectiveness. For instance, a project design team might pose the following foreground questions in the:

- **Problem analysis:** "Is overexploitation of forest resources in Amazonian forests associated with a lack of community rights?" to address the assumption that overexploitation is driven by lack of community rights.
- **Development solution:** "Does receiving financial benefits generated by community forestry strategic approaches affect individual attitudes about protection of forest resources?" to address the assumption that attitudes towards conservation will improve if benefits are generated for local communities.
- **Development of the implementation plan:** "Does training stockholders in community-based approaches increase community support for sustainable resource management?" to address the appropriateness of investing in training in community-based approaches as a means of engaging stakeholders in natural resource management.

The Center for International Forestry Research conducts social dimension analysis and forestry livelihood framework assessments of community-based commercial forestry in Indonesia. Photo credit: Aris Sanjaya/CIFOR

BOX 3: UNDERSTANDING ASSUMPTIONS

In USAID programming, the term "assumption" can have different connotations in different contexts.[4] As used in this document, an assumption refers to the logical connections between drivers, threats, and the status of biodiversity focal interests in a problem analysis or those that underlie anticipated results articulated in a program's theory of change.

This conceptualization of an assumption goes beyond describing what is expected to happen (i.e., a result) by articulating the team's understanding of why and how something will happen. The assumption can be expressed as a declarative statement that can be substantiated or refuted with evidence. Gathering evidence is facilitated by the framing of questions that can be used to search the evidence base.

Result: *Increased vigilance and control over external actors responsible for illegal harvest of forest resources is achieved.*

Assumption: *Community involvement in enforcement results in fewer external actors illegally harvesting forest resources than government enforcement alone.*

SETTING PRIORITIES FOR GATHERING EVIDENCE

The use of evidence can help teams increase the likelihood that their investments will achieve desired conservation and development goals. However, teams are often faced with a broad array of information needs that they must trade off with the reality of time and resource constraints. For this reason, it is important that teams have ways to focus their search on their most critical needs for evidence.

A risk framework (similar to that used to identify risky assumptions in the Country Development Cooperation Strategy) can help teams identify a smaller set of assumptions that are priorities for gathering evidence. A risky assumption is one that both has a high likelihood of being invalid and would jeopardize the program's success if invalid (see Figure 3 on page 15). Because of their potential negative effects on program success, teams should prioritize gathering evidence about assumptions with serious consequences (see Box 4 about risk and hazards on page 16).

Certain assumptions are particularly important for ensuring program success

(i.e., they almost certainly jeopardize program success if invalid). These assumptions are high priorities for gathering evidence[5] about their validity and include:

1. Assumptions in the problem analysis that identify the driver that a strategic approach is designed to directly influence
2. Assumptions in the problem analysis that identify the immediate cause of the threat
3. Assumptions with doubtful causality in the theory of change
4. Assumptions about the effectiveness of actions intended to influence key drivers in the program context

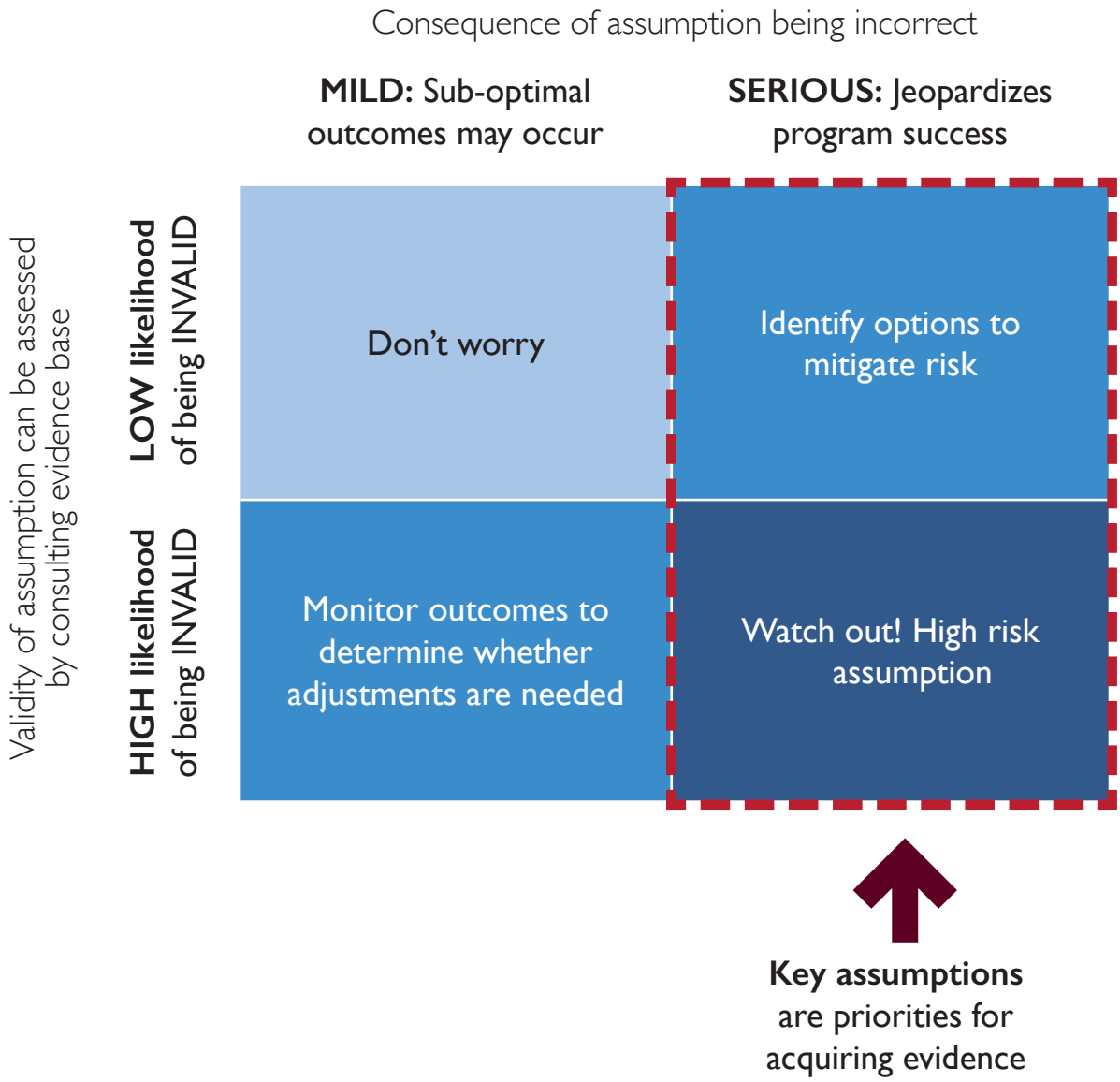

Figure 3: Using a risk framework to identify priorities for gathering evidence about the validity of program assumptions (adapted from Guijt (2013))

BOX 4: UNDERSTANDING RISK AND HAZARD

An assumption can be low risk even if it has the potential to jeopardize program success, as long as there is reasonable certainty that it will hold true. A skydiver assumes that pulling the ripcord will release her parachute. If invalid, this assumption poses a significant hazard (or source of harm) to the success of the dive, but it is not a high-risk assumption because the evidence suggests there is a low likelihood that the assumption will be invalid. If the evidence suggested that pulling the ripcord often fails to release the parachute (i.e., the evidence suggests that the assumption was likely to be invalid), the skydiver would certainly want to reconsider the activity.

Flood hazards present a risk to low income households along riverbanks in Jakarta, Indonesia. Photo credit: Farhana Asnap/ World Bank

EXAMPLE 2: KEY ASSUMPTIONS IN THE DESIGN OF A COOKSTOVE PROGRAM

A design team is developing an activity in a province where fuelwood collection is thought to be a primary driver of deforestation. The team has identified a lack of alternatives as a reason why inefficient cookstoves are currently in use (see Figure 4). They are considering a strategic approach that would distribute free efficient cooking devices to adults who attend literacy classes as an incentive to increase literacy rates among adult men while decreasing rates of deforestation. The team identifies several important assumptions that will influence program success.

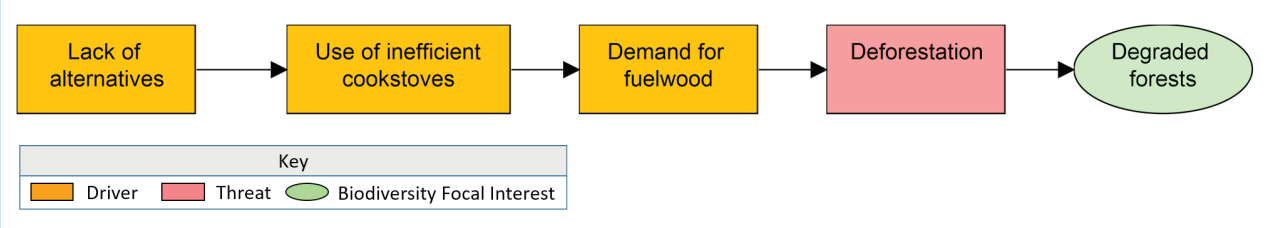

Figure 4: Depiction of the problem analysis for a cookstove program

KEY ASSUMPTION 1: Lack of access drives low usage of efficient cookstoves among households where adults have low literacy.

Review of a political economy analysis conducted early in the design process confirms that households where the head of household is illiterate report lower access to efficient cookstoves than households where the head of household is literate. In this instance, the team concludes that they have identified an appropriate set of stakeholders and a valid driver for the strategic approach to influence.

This remains a key assumption because program success would be jeopardized if the strategic approach targets the wrong set of stakeholders. However, the findings from the political economy analysis suggest that the assumption is not likely to be particularly risky.

KEY ASSUMPTION 2: Demand for fuelwood drives deforestation.

The team has assumed that fuelwood use is the most immediate cause of deforestation. A review of recent scientific literature reporting regional trends in forest status and threats reveals that development of oil palm plantations is contributing much more significantly to deforestation rates in the region.

This is a high-risk assumption because the evidence base suggests the assumption may be invalid (and it poses high consequences to program success). Incorrectly assuming that fuelwood consumption is the primary driver of deforestation in the local area could lead to over-allocating resources to address a relatively minor threat. To increase their confidence in this assumption, the team partners with researchers at a local university to analyze recent satellite imagery and assess whether patterns of deforestation are consistent with reported fuelwood consumption. The team may need to update their problem analysis based on the findings.

KEY ASSUMPTION 3: Switching from traditional to efficient cookstoves causes households to use less fuelwood.

The team has assumed that changing the type of cookstove will cause households to use less fuelwood. A review of recent scientific literature suggests that the assumption holds true in some program contexts, but that in other contexts, households continue to collect the same amount of fuelwood in order to sell the surplus in local markets.

In this case, the existing evidence base does not resolve the team's uncertainty about the validity of the assumption. They know that if the assumption is incorrect, it will jeopardize program success by breaking a critical linkage in the theory of change. To minimize risk, the team decides to modify their program design to include incentives that deter fuelwood collection for non-personal use.

KEY ASSUMPTION 4: Distributing efficient cookstoves free of charge is an effective approach for increasing their use.

The team has assumed that high use of efficient cookstoves can be achieved by distributing cookstoves free of charge. However, the evidence base suggests that this may not be the case and identifies some of the barriers to widespread adoption of the new cookstoves (Jeuland and Pattanayak 2012, Lewis and Pattanayak 2012).

The team remains uncertain whether the distribution of cookstoves will lead to adoption of the new technology. They review evaluations from activities that have implemented the same strategic approach and identify several low-cost community programs that have helped households learn about new cooking methods and how to maintain their cookstoves. They modify their work plan to include these additional actions.

PRACTICAL TIPS FOR PROGRAM MANAGERS

A key opportunity for critical review of the use of evidence in biodiversity programming is during activity procurement.

When reviewing proposals, managers can evaluate whether and how effectively implementing partners are using evidence to support their technical solutions. Managers can ask the following questions to assess the use of evidence during proposal review:

1. Does the proposal reference a review of existing evidence about what has worked and not worked in similar contexts? Is this information consistent with the proposed activities? When proposed actions are supported by a weak evidence base, the proposal should provide a rationale for using the proposed actions instead of any known alternatives.

2. Have key assumptions (i.e., those with the potential to be high-risk) been identified and has the uncertainty in the validity of those assumptions been appraised? A strong proposal may contain information gaps or rely on assumptions that lack strong evidence, but it will identify those gaps as high priorities for monitoring, evaluation, and learning.

A water project in the Mau Forest in Kenya draws on evidence that forests, land use, and water quality are interconnected. Photo credit: Patrick Shepherd/CIFOR

4. GATHERING EVIDENCE

A core activity of an evidence-based approach is gathering objective external information (Pullin and Knight 2001, Sutherland et al. 2004). For program managers, this necessitates a commitment to checking information gained from personal experience, and even long-held beliefs, against other sources (see Box 5). These external checks are important because cognitive biases can cause teams to misinterpret information and reach erroneous conclusions. Simply understanding the different types of cognitive biases that influence decisions (Kinderman et al. 2016) can help teams to understand when it may be important for them to revisit the information that is being used to make a decision.

BOX 5: CONSULTING THE LITERATURE

There is a difference between simply consulting the literature and adopting an evidence-based approach. Teams adopting an evidence-based approach gather evidence in a systematic and transparent way in order to minimize the influence of common biases on decisions. Their search should aim to uncover evidence that supports their assumptions as well as evidence that would refute those assumptions. In addition, they pay careful attention to the quality and strength of the evidence they are using to support their assumptions and rule out alternative explanations. An ad hoc search of the literature with the purpose of finding a few pertinent references to support a program assumption is not the same as taking an evidence-based approach, and it may in fact have the opposite effect of reinforcing confirmation bias.

Two common cognitive biases (see Figure 5) that are particularly relevant to programming decisions are:

- **Confirmation bias:** The tendency to selectively favor information that confirms existing beliefs and ignore contradictory information.
- **Groupthink:** The tendency to adopt a belief that is held by other members of a group without critical evaluation of alternative viewpoints.

When designing evidence-based biodiversity programs, teams can counter these common cognitive biases by looking for evidence both for and against a suspected cause and by looking for evidence for and against other important possible causes.

Figure 5: Common cognitive biases

EVIDENCE IN THE USAID CONTEXT

Program managers rightly use many kinds of information when making program decisions, not all of which would be considered objective evidence. This does not mean that this information should not be used or considered, but instead that it should be critically evaluated if it is helping the team assess the validity of key assumptions.

Most teams are well-versed in using information from many sources, such as the opinions of stakeholders, the sharing of success stories, and reflecting on previous experiences, but they may have less experience with the different types of publications (see Table 1) they may encounter when searching the evidence base.

Table 1: Types of publications teams may encounter when searching for evidence (adapted from Briner and Barends (2016))

Publication type	Typical pros	Typical cons	Where to find it
Systematic reviews[6]	May be very relevant. Focused on a specific question and are designed to be transparent and comprehensive.	Still relatively rare in biodiversity programming. Practitioners may find it difficult to apply inconclusive findings to program decisions.	The journal Environmental Evidence focuses entirely on systematic reviews. Other conservation journals periodically publish systematic reviews.
Meta-analyses	May be very relevant. Statistically summarize previous research on a particular research question.	May be more likely to address questions about biological and social processes and may be less likely to provide actionable information for program decisions.	Subject matter journals periodically publish meta-analyses (e.g., Conservation Biology, Journal of International Development).
Literature reviews	May be very relevant. Usually quite readable. Provide some details about reviewed studies. Often point to further research needs.	Usually focused on a topic rather than a question. Author(s) may select studies that support a particular conclusion.	Subject matter journals and edited volumes.

Publication type	Typical pros	Typical cons	Where to find it
Evidence syntheses	Readily available and easily digestible. May be very relevant and tailored to USAID's needs. May provide a rich bibliography to search for more evidence on a particular topic.	May provide information that is too general to be applied to a specific context.	USAID Biodiversity Conservation Gateway and Learning Lab; Conservation Evidence publishes evidence synopses about conservation strategic approaches.
Individual scientific studies	May be highly relevant. Users should critically evaluate methods and results to assess quality.	Drawing inferences from single studies can be problematic. Single studies matter most in terms of their contribution to the larger body of evidence.	Subject matter journals
Grey literature (e.g., white papers, reports and practitioner guidance)	May be very relevant. Can include case studies and examples that can be compared and contrasted to the team's programming context.	Not peer-reviewed. May vary in quality and the extent to which they are based on sound empirical evidence.	Organizational websites, clearinghouses, and online communities of practice.
Organizational data	May be highly relevant. Can identify successes and failures in similar programming contexts.	Focus tends to be on evidence about outcomes and confirming (rather than testing) program assumptions. Conclusions likely to be context-dependent. May be limited in the extent to which they examine evidence against suspected causes and for and against other important possible causes.	Documents and databases within organizations including evaluations, progress reports, case studies, etc.

Beginning the search. Program managers at USAID have access to a number of resources through the USAID Knowledge Services Center that can support them in gathering evidence.

- Library and reference services: Desktop access to electronic journals and the USAID library catalog; expert searches for journal articles, news stories, and USAID documents

- Research services: Quick updates or in-depth research and analysis on topics identified by development practitioners; individualized expert help with research tools

- Access to the Development Experience Clearinghouse: The largest online resource for USAID-funded technical and program materials, including research reports, evaluations and assessments, and other policy and planning documents. Resources are searchable by topic, geographic area, and document type

Biodiversity-specific resources are available through the USAID Biodiversity Conservation Gateway. USAID's Office of Forestry and Biodiversity maintains the Gateway as a publicly available information portal with a rich repository of USAID experience implementing commonly used strategic approaches. The Gateway also hosts USAID's Biodiversity Cross-Mission Learning Program, which provides an opportunity for USAID staff to share evidence and learning.

Self-directed searches. Through the Knowledge Services Center USAID staff have access to Web of Science, an online citation indexing service that covers topics including science, social science, and the humanities. Indexed content includes peer-reviewed journals, scholarly books, reviews, editorials, and abstracts. In addition to topic searches, Web of Science supports citation searches that track the number of times papers have been cited and their cited references. The Knowledge Services Center provides training on the use of Web of Science.

Design teams can review evidence for key assumptions when developing theories of change. Photo credit: International Initiative for Impact Evaluation

Google Scholar provides another simple way to broadly search for scholarly literature. Google Scholar includes both peer-reviewed articles and grey literature, and can be particularly helpful when alternative terms are used by practitioners and academics. Google Scholar is also a useful tool for locating full-text articles that have been made available online.

The search process. As decision makers, the information needs of program managers are distinct from researchers and those tasked with carrying out on-the-ground activities. Program managers need to stay current on new developments and find answers to program-related questions. Filtering out useful information from a large volume of published materials can seem a daunting task. While it may seem counterintuitive, the goal of the search-and-filter strategy is not to focus on comprehensiveness, but to find the best answer as quickly as possible (Grandage et al. 2002). The process involves: (1) identifying relevant search terms, (2) filtering results for high quality sources of evidence in the secondary literature (e.g., systematic reviews), and (3) expanding the search to sources of evidence in the primary literature as needed. Teams should pay attention to the credibility of the sources they find (see Box 6). A good search is systematic, transparent, and verifiable.

Useful information is judged on its validity, relevance, and value relative to the effort it takes to acquire.

BOX 6: EVALUATING SCHOLARLY JOURNALS

Over the past decade the problem of "predatory journals" has become more prevalent. Predatory journals may claim to conduct peer review, but the articles they publish fail to meet scholarly publishing standards and should be regarded as suspect. The World Association of Medical Editors provides a helpful resource on *Identifying Predatory or Pseudo-Journals* (Laine and Winker 2017) that applies more broadly to scientific publishing.

PRACTICAL TIPS FOR AN EFFECTIVE SEARCH

An effective search starts with the right question. A team paves the way for an effective search when they take the time to construct a well-structured question.

1. The team should assess whether they are ready to pose a foreground question (see Section 3). If the team does not yet have enough information to formulate assumptions, their information needs are likely best addressed by background questions. Background questions typically lead to topical searches and rely on a broader set of resources, which may include books, guidance documents, reviews and syntheses, and even trusted websites. These are the types of literature searches with which most teams have experience from their academic training and professional activities. A team asking, "What types of strategic approaches have been used to combat wildlife trafficking?" is asking a background question.

2. Once a team has articulated specific assumptions, they are ready to pose foreground questions. Making the question as specific as possible helps to narrow the search terms. This reduces the time spent searching for relevant evidence and increases the likelihood of finding it. Most foreground questions can be divided into three parts:

   i. The subject or population – Consider who or what the subject of the question is. What are the most important characteristics of the subjects or population? How would a similar group be described? The descriptors may include the threat, biodiversity focal interest, or context.

   ii. The driver or change agent – Identify the strategic approach, change, or driver that is under consideration. What factors influence the current status of the biodiversity focal interest? What factors lead to desired change?

   iii. The consequence – Identify the end result or outcome that is produced as a consequence of the change agent or driver. What are the effects of the driver? What changes are desired?

Example:

The team has identified a lack of awareness about alternatives as a potential driver for the use of unsustainable fishing practices.

Initial question: To what extent does a lack of awareness about alternative fishing practices lead to unsustainable fishing?

The initial question did not identify a group of subjects or population. The team describes the subject/population as subsistence (small-scale) fishers using unsustainable fishing practices. Specifically, they are working in the Danajon Bank coral reef system in the Philippines.

The team reexamines the driver and consequence in their initial question. The team thinks through their beliefs about how they expect lack of awareness about alternatives to influence the practices used by fishers in the activity area. They refocus their understanding of the driver on the lack of awareness about the costs and benefits of alternative fishing practices relative to current practices.

3. The team realizes that their underlying assumption is that this lack of knowledge poses a barrier to the use of sustainable practices. They revise their question to directly test this assumption.

Searchable Question: Among subsistence fishers, is lack of knowledge about the costs and benefits of alternative fishing practices a significant barrier to the use of sustainable fishing practices?

4. The team uses the components of the question to identify search terms. Here they might combine "fishing knowledge" and "fishing methods" as an initial set of search terms. The team may use more or fewer search terms, for example including related terms describing the subject (destructive fishing practices, bottom trawl, dredging) to broaden or refine their search as needed.

5. The team screens the available results quickly to determine whether there are any relevant systematic reviews, meta-analyses, or literature reviews that could answer their question. If they found none, they would turn their search to individual studies.

5. APPRAISING THE QUALITY AND STRENGTH OF EVIDENCE

Identifying the "best" evidence depends on what kind of information is most valid and reliable for addressing a particular question or information need. Teams should assess both the quality and the strength of the evidence they are using to inform their decisions (DFID 2014). The issues of quality and strength of evidence are especially important whenever a program assumption describes a causal relationship.

Quality of evidence reflects the extent to which a study's design is able to rule out alternative explanations for observed outcomes. Criteria for rating the quality of the evidence from a particular study include the rigor of a study's design and methods of implementation, the precision of measurements and analyses, and the coherence between the results and the study's conclusions.

Strength of evidence refers to the level of confidence in conclusions drawn from aggregated findings from the studies making up the evidence base. Criteria for rating the strength of the evidence base include the quality of the studies comprising the evidence base, as well as the quantity and consistency of findings across multiple studies.

An evidence base that includes multiple studies that repeatedly confirm the same "plausible association," even if the individual study designs cannot rule out alternative explanations, is stronger than one that includes a smaller number of such studies. For this reason, secondary sources of evidence that systematically review the quality and strength of evidence for specific questions are particularly valuable resources. Both systematic reviews and meta-analyses are designed to comprehensively search for, appraise, and synthesize research evidence. Although these types of evidence syntheses are becoming more common, realistically, a team may not be able to find a systematic review or meta-analysis relevant to their specific questions.

STUDY DESIGN

Study design is particularly important to consider when using evidence to support a program assumption that goes beyond a description of "what is" to explaining how and why something is expected to happen. When teams are looking for evidence to support or refute assumptions about causal relationships they need to pay particular attention to the quality of evidence generated by different study designs (Trochim 2006). Stronger designs are those that do a better job of ruling out alternative or "rival" explanations.

For example, if a study measures change in fishing income among fishers adopting sustainable fishing methods, knowing that these fishers increased their income does not necessarily mean that the use of sustainable fishing methods was the cause. A study design that includes fishers who continue to use traditional fishing practices as a comparison group allows differences between groups to be assessed. This design provides stronger evidence for the use of sustainable fishing practices as a cause of higher incomes among fishers. The team must still consider and rule out other factors that may differ between the groups before they can conclude that the type of fishing method was the cause of the observed differences in income.

RECOGNIZING BIAS

Bias distorts or confounds the findings in a study, making it difficult to interpret the study's results (Smith and Noble 2014). Because it is usually impossible to know the extent to which biases have affected the results of a particular study, it is important to consider the risk of bias whenever a team uses evidence to support a particular assumption. Four issues that program managers should be aware of when presented with evidence in support of causal assumptions are: aggregation bias, selection bias, confounding, and external validity.

Aggregation bias (Jargowsky 2005) should be considered whenever associations at the population level are used as evidence to support causal relationships that operate at the individual level. Teams should pay particular attention to the potential for aggregation bias when indicators are used to draw conclusions about causal relationships. For example, if a team finds that the percentage of fishers using sustainable fishing practices and average household income both increase over the same time period, the association

USAID's Office of Forestry and Biodiversity can provide technical assistance to teams seeking support in critical appraisal of evidence.

observed at the village level may or may not reflect a positive relationship between increased income and the use of sustainable fishing practices *among individual fishers.*

Confounding (Skelly et al. 2012) occurs whenever some outside factor co-varies with the presumed cause, making it difficult to establish the true cause of an observed effect. When confounders are present, it may be difficult to rule out alternative explanations for observations that are consistent with program assumptions. In the example above, researchers may have failed to take into account that all villages in the province experienced higher average incomes due to oil royalties. The adoption of sustainable fishing practices coincidentally increased in parallel with average income, but did not cause the increases.

Selection bias (Collier and Mahoney 1996) can introduce confounders into study designs that use comparison groups. Selection bias is a risk whenever groups differ systematically in some way other than the presumed cause that influences measured outcomes. That is, the groups are not truly "comparable." For example, selection bias could be introduced into a study design if fishers adopting sustainable fishing practices have access to more productive fishing grounds, or fish different species, than those using traditional fishing practices. Randomization, restriction, and matching[7] are examples of methods that are used to minimize or exclude selection bias in study design.

External validity (Steckler and McLeroy 2008), or transferability, is the extent to which the findings from one study can be applied to other contexts. It is not a bias in the strict sense, but studies that have problems with bias automatically have problems with external validity. External validity is also compromised when the subjects included in the study are not representative of the broader population to which the findings are being applied. For example, a study showing that seasonal fishing closures increases fish stocks without reducing annual incomes of fishers in small-scale fishing communities is unlikely to apply to communities dependent on commercial fisheries for their livelihoods. Assessing external validity is particularly important when a team uses results from one context or setting to infer that an assumption will be valid in other contexts (Ferguson 2004).

The Research Methods Knowledge Base provides an overview of social research methods for teams looking to refresh their knowledge.

REPRODUCIBILITY

Reproducibility refers to the size of the evidence base in support of an assumption (Casadevall and Fang 2010). The basic idea behind reproducibility is that a relationship that is observed repeatedly, particularly if observed consistently in different contexts and under different conditions, is more likely to be true. For program managers, reproducibility is an important consideration when the evidence supporting an assumption is drawn from less robust designs. In these cases, observing the same pattern across multiple sites and contexts can increase confidence that an assumption is valid.

TRIANGULATION

Triangulation refers to using multiple forms of evidence to support a conclusion (Thurmond 2001). It differs from reproducibility in that the emphasis is on corroborating findings through diverse and redundant types of evidence, rather than the quantity of studies that provide similar findings. Triangulation is generally used to strengthen the confidence of findings within a study's context.

Triangulation works well for addressing program assumptions about behavior and institutional change because the change agents (stakeholders) can be asked directly about their motivations, beliefs, and perceptions of change. For example, a study using a single group design[8] that measures an increase in income among fishers who have adopted sustainable fishing methods provides fairly weak evidence that changes in fishing methods cause changes in fisher income. However, a mixed methods approach – where quantitative findings are corroborated with interviews that elicit fishers' perceptions of the factors limiting and enhancing their own income potential – can make the quantitative data on change in income more meaningful.

The point of triangulation is not to eliminate biases from individual study designs, but to bring in cross checks on the accuracy of conclusions drawn from studies that may be susceptible to bias (Jick 1979). Triangulation is most effective when corroborating evidence comes from methods that have different weaknesses and potential for bias. If these different methods all produce findings that lead to the same conclusions, credibility of those findings is increased.

PRACTICAL TIPS FOR PROGRAM MANAGERS

When reviewing the use of evidence in program design, program managers should pay close attention to the quality of evidence cited in support of causal assumptions. A frequent misapplication of evidence is concluding that X causes Y without fully considering whether that evidence excludes alternative or "rival" explanations. In scientific studies, authors often report "associations" rather than suggest causation, because they cannot be certain that X caused Y.

Program managers should recognize the limitations to the evidence base when appraising evidence. Similar considerations apply when teams interpret findings from commissioned research and other Program Cycle processes used to generate evidence (see Section 5 in *Unit 3: Generating Evidence*).

When appraising evidence in support of causal explanations, program managers should be able to determine:

1. **Whether the evidence comes from studies that measured or observed outcomes or conditions.** Soliciting opinions about whether something *should* work is not evidence that it does work.[9] Outcomes do not always have to be measured directly. They can also be assessed indirectly by asking stakeholders about outcomes that have been achieved and the conditions that motivated or prevented change.

2. **Whether time order has been addressed.** A cause must be in place before a change in outcome occurs. A study that shows an association between fishing method (e.g., use of traditional vs. sustainable fishing practices) and income provides weaker evidence than a study that shows that average fisher income increased *after* the adoption of sustainable fishing methods.

3. **The methods used to establish correlation between cause and effect.** Studies that use before-and-after comparisons or appropriate control groups can establish co-variation between the cause and effect. Careful attention should be paid to the alignment between the presence or absence of the presumed cause and any comparisons made in the study design.

4. **Whether plausible explanations have been considered and reasonable efforts have been made to rule them out.** Not all studies are equally robust for establishing relationships between presumed causes and effects, but there is no single "best" design that applies to all situations.

6. USING EVIDENCE IN USAID BIODIVERSITY PROGRAMS[10]

Teams using an evidence-based approach understand that their assumptions can – and should – evolve as they progress through the Program Cycle. As teams build their understanding of the evidence base, they must decide whether and how to apply these findings to program decisions. It is at this point that program managers and implementers integrate knowledge from the evidence base with their own understanding of the socio-political and biological context.

The USAID Program Cycle offers several junctures where use of evidence is very valuable. National- and regional-level Biodiversity and Tropical Forestry (Foreign Assistance Act 118/119) Assessments and targeted Biodiversity Threats Analyses identify threats and drivers of biodiversity loss, biodiversity focal interests, enabling conditions, and key actors. Additionally, they may highlight the relationships between biodiversity and other development sectors. These analyses can be used by USAID managers to gather evidence supporting the use of particular strategic approaches before and during program design.

Evidence may also come from outside the biodiversity sector. For instance, it can be instructive to learn about the effectiveness of capacity-building approaches, women's empowerment strategies, media campaigns, or livelihood alternatives that may be integrated into conservation programming.

Evidence relevant to biodiversity and integrated programming can help teams meet the requirements of the Biodiversity Policy and Code. For example:

- Reviews and syntheses relating to biodiversity focal interests, threats, drivers, social context, and enabling conditions can guide the scope and scale of a program.
- Reviews and syntheses on common theories of change that describe what is known about the effectiveness, appropriateness, scale, risks, and benefits of the proposed strategic approaches can help determine actions.
- Reviews of cross-sectoral analyses (e.g., gender, sustainability, and environmental compliance) can inform design and monitoring.

EXAMPLE 3: USING EVIDENCE TO IMPROVE LOCAL GOVERNANCE IN COMMUNITY-BASED NATURAL RESOURCE MANAGEMENT

An implementing partner is in the process of developing an activity-level theory of change as part of the activity start-up process. The project had identified unsustainable logging as a threat to dipterocarp forest, the biodiversity focal interest. According to the theory of change outlined in the Project Appraisal Document (see results chain in Figure 6), empowering the local communities to manage their natural resources through community-based natural resource management will lead to better resource governance, which in turn will lead to better resource management and result in better conservation outcomes. The team applies an evidence-based approach to their theory of change.

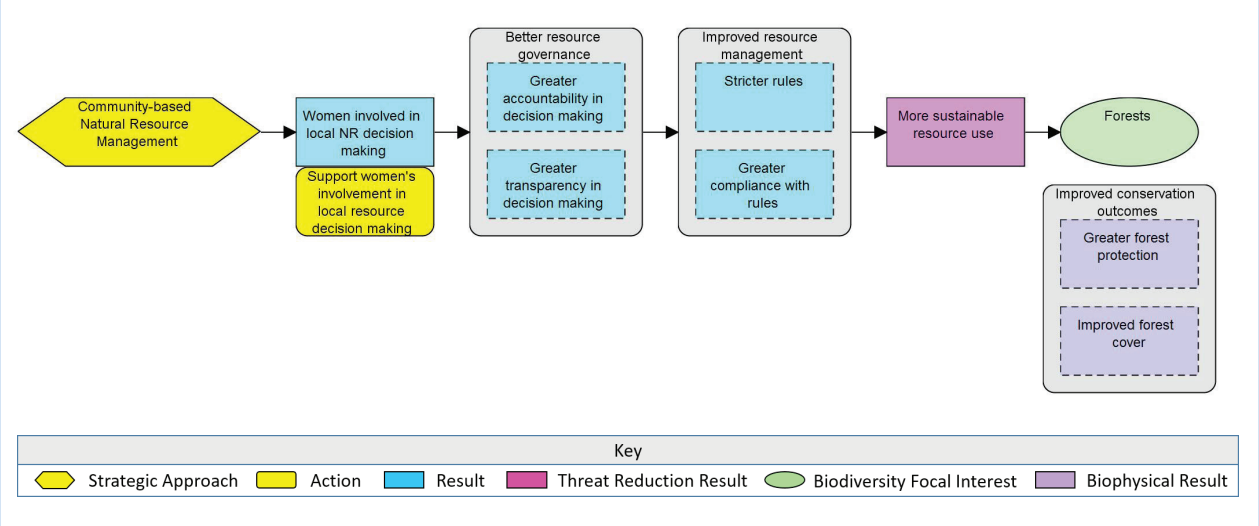

Figure 6: Results chain for a community-based natural resources management strategic approach

Identifying the Information Need:

Based on their experience working in the area, the implementing partner knows that men and women use forest resources differently. Men tend to focus on extraction of profitable timber products, while women are more likely to collect firewood and fodder for animals. The team suspects that the composition of the decision-making group may be an important factor affecting the governance of natural resource management decisions, but they are uncertain about the extent to which including women in local resource governance is likely to increase sustainable resource use.

Asking the Question:

The validity of the assumption that gender balance influences the success of community-based management groups influences program decisions about how to manage actions to maximize conservation outcomes. The team wants to make sure that they are investing their resources appropriately. Should they focus their efforts on empowering women to participate in local resource decision-making? Or would similar outcomes be achieved if they mobilize traditional male-dominated community groups already active in the community? To make these decisions they need to answer the following question about the effectiveness of the strategic approach, "Does the gender composition of forest management groups affect conservation outcomes?"

Gathering Evidence:

The team reviews a USAID-supported systematic map, or overview of the quantity and quality of evidence on a policy- or management-relevant topic, that appraised the available evidence linking the gender composition of community groups managing natural resources to resource governance and conservation outcomes in forestry and fisheries (Leisher et al. 2016). The study identifies 14 articles about gender composition of community-based forest management that met the inclusion criteria for analysis. All 14 studies looking at forest management found improvements in local natural resource governance when women participated in the management of the resources.

Appraising the Evidence:

The team notices that not all of the 14 studies were able to exclude alternative explanations for the observed outcomes, and that most of them were about community-based natural resource management activities in India and Nepal (which limits the external validity of the conclusions). Although the evidence base is incomplete, the team is encouraged that none of the studies found negative outcomes associated with more equitable gender representation in community-based natural resource management. Thus, they conclude that the risks to governance and conservation outcomes are low, and the benefits may be high. The team decides to include actions that target participation by women in local resource governance as part of their strategic approach.

Activities addressing threats to endangered species, such as this scalloped hammerhead shark, may have a lower risk tolerance for missed opportunities. Photo credit: Kevin Lino NOAA/NMFS/PIFSC

CONSIDERATIONS WHEN THE EVIDENCE BASE IS LIMITED

One benefit to using evidence-based approaches is that teams develop a better understanding of what is not known as well as what is known about key assumptions influencing program success. Knowing where there is uncertainty in the validity of key assumptions allows teams to be more transparent in their decision-making process and identify where to focus efforts on filling knowledge gaps.

A weak or mixed evidence base can limit the team's ability to assess the level of risk that a given decision poses to program success. When this occurs, teams can base their decision-making strategy on their risk tolerance for alternative program outcomes (Keeney 2004, Salafsky and Redford 2013).

Four program outcomes are possible when the validity of key assumptions is unknown (see Figure 7 on page 37):

1. The strategic approach is implemented and key assumptions are valid. The investment of money and resources leads to conservation outcomes (good return on investment).

2. The strategic approach is implemented but key assumptions are not valid, which results in suboptimal outcomes. The return on investment is poor (misplaced investment).

3. The strategic approach is not implemented but key assumptions are valid. The program would have been effective but instead the status of the biodiversity focal interest continues to decline (missed opportunity).

4. The strategic approach is not implemented and key assumptions are not valid. The program would not have been effective and the money and resources remain available for other purposes (prudent use of resources).

If the consequences of a missed opportunity are relatively low (e.g., because those resources can be allocated to other less risky investments) relative to the consequences of a misplaced investment, the team may decide to delay implementation of that particular strategic approach until better information is available. When the consequences of inaction are high, the team may decide to move forward with implementation and focus on generating evidence about the effectiveness of the strategic approach (see Box 7 on page 38).

A team that implements a strategic approach with incomplete knowledge about the validity of important program assumptions can use the learning sections in the program's Monitoring, Evaluation, and Learning plan[11] to identify important evidence gaps and formulate their plan for addressing them. Generating evidence to address such gaps is the topic of *Unit 3: Generating Evidence*.

	Validity of key program assumptions	
Decision	Valid	Invalid
Implement the strategic approach with incomplete information	**Good return on investment**	**Misplaced investment**
Wait for better information or pursue alternatives	**Missed opportunity**	**Prudent use of resources**

Figure 7: Four possible program outcomes that a team should consider when the validity of key assumptions is unknown

BOX 7: IMPLEMENTING A STRATEGIC APPROACH WHEN THE TEAM KNOWS THE EVIDENCE BASE IS WEAK

The following considerations may help teams decide when to employ a strategic approach with less than perfect information about the validity of key assumptions.

Consider the alternatives to implementing the proposed strategic approach

Program managers are accountable for responsible allocation of funds in addition to achieving conservation outcomes. When managers have low risk tolerance for misplaced investments, forgoing action is one alternative – but it may also mean that the status of the biodiversity focal interest is likely to continue to decline.

When the costs of a particular strategic approach are high, many program managers may be less comfortable implementing a strategic approach with uncertain outcomes. They may be able to mitigate the consequences of delaying action by pursuing alternatives with more certain conservation outcomes – even if they are of potentially lesser impact – while waiting for more information.

Consider the consequences of a missed opportunity

If there is immediate risk that the biodiversity focal interests will be lost without action, program managers may decide that the consequences of doing nothing (which may be fairly certain) favor implementing the strategic approach even if the evidence base is weak. In this situation, program managers should work with implementing partners to ensure that robust monitoring and evaluation activities are in place to monitor change in biodiversity focal interests alongside other performance indicators.

A systematic review on alternative livelihood projects found limited evidence supporting the effectiveness of this strategic approach (Roe et al. 2015). Among 21 studies that measured conservation effectiveness, there was significant heterogeneity in program outcomes. Failure to review what is known about the enabling conditions associated with key results (see Building a Conservation Enterprise: Keys for Success) or to consider possible alternative approaches that have stronger track records of success, can lead to poor program investments.

7. SUMMARY OF KEY CONCEPTS

- Program managers and implementing partners use evidence to increase the effectiveness of biodiversity programs. Without evidence about what works and what doesn't work, teams are susceptible to missing opportunities to replicate successes, and they risk continuing to invest in programs with a low track record of success.
- Mission staff and implementing partners make decisions that can be supported with evidence throughout the Program Cycle. Teams use evidence to support program assumptions when (1) analyzing the development problem, (2) articulating the theory of change for a strategic approach, and (3) developing implementation plans.
- Evidence can come from many sources including peer-reviewed scientific journal articles, grey literature reports and notes, program evaluations, and expert understanding.
- Gathering evidence requires an investment of time and resources, so it is important that teams have ways to focus their search on their most critical information needs. One way to set priorities for gathering evidence is by focusing on high-risk assumptions, i.e., those that have serious consequences (including program failure) if they are invalid.
- Identifying the "best" evidence depends on what kind of information is most valid and reliable for addressing a particular question or information need. Teams should assess both the quality and the strength of the available evidence they are using to inform their decisions.
- A weak evidence base should not automatically bar a team from deciding to implement a strategic approach. But the team should carefully consider the costs and benefits of action versus inaction and their risk tolerance for alternative outcomes when making decisions about whether to proceed.

8. FURTHER READING

Assessing quality and strength of evidence:

USAID Degrees of Evidence Framework (USAID 2008). This practical framework can be used by program managers to match evaluation methodologies to program purpose, level of credibility required, and level of resources available.

How to note on assessing strength of evidence (DFID 2014). This document provides guidance for assessing the strength and appropriateness of evidence. It argues that different research designs and methods are appropriate for different kinds of questions.

Biodiversity, ecosystem services and poverty alleviation: What constitutes good evidence? (Sandbrook 2013). This paper gives a brief introduction to evidence-based conservation and summarizes key messages from the literature with the intention of stimulating debate about what constitutes good evidence, as well as how to address challenges to using an evidence-based approach.

Twenty tips for interpreting scientific claims (Sutherland et al. 2013). This commentary reviews concepts that can help decision makers improve their interpretive scientific skills in order to better evaluate scientific findings.

Gathering evidence:

Why discrepancies in searching the conservation biology literature matter (Calver et al. 2017). This paper compares the efficiency of literature searches for biological conservation across multiple databases.

Selecting appropriate methods of knowledge synthesis to inform biodiversity policy (Pullin et al. 2016). This paper presents a range of different methods used to conduct knowledge syntheses addressing information needs and questions arising in biodiversity policy and management.

Making decisions under uncertainty:

Turning uncertainty into useful information for conservation decisions (USGS 2016). This fact sheet describes common socio-ecological uncertainties, suggests how to apply a decision-making perspective to uncertainty and risk, and lays out four principles for robust conservation decision-making.

ENDNOTES

1 Box 2 in *Unit 1: Understanding an Evidence-Based Approach* includes a representation of three components of program success.

2 A real-word example of how data about fishing practices and stock status might be analyzed comes from Vasilakopoulos et al. (2011) who conducted a meta-analysis to quantify independent effects of fishing intensity and fishing selection on current stock status of 13 species in the Northeast Atlantic.

3 During the design phase, teams articulate assumptions that provide the rationale for how and why the program is expected to work. *Unit 3: Generating Evidence* provides further information on how an assumption is verified (or refuted) with evidence generated by testing one or more hypotheses derived from the assumption.

4 ADS Chapter 201 requires teams to identify critical assumptions and risks which refer to the most critical uncertainties and risk factors beyond USAID's influence or control that could affect achievement of the program's planned results. These differ from the assumptions representing causal (if-then) relationships in a results chain (see Box 4 in ***Biodiversity How-To Guide 2: Using Results Chains to Depict Theories of Change in USAID Biodiversity Programming***).

5 Assumptions identified as high priorities for gathering evidence may also be high priorities for generating evidence when the existing evidence base is weak (see Section 3 in *Unit 3: Generating Evidence*).

6 In *Unit 1: Understanding an Evidence-Based Approach*, the use of a systematic review in program design is illustrated in Example 1: Applying the Evidence Base to Marine Protected Areas (page 11).

7 Restriction refers to the practice of narrowing down the subjects of a study to eliminate the influence of a confounder. In the fisher example presented on pages 30-31, the observations could have been restricted to fishing villages not receiving oil royalties. Matching refers to creating matched comparisons between subjects that differ in one relevant attribute. In the examples above, researchers can choose to compare fishing villages receiving training in sustainable fishing practices with similar fishing villages that did not receive the training.

8 Limitations of different designs for establishing causal relationships are discussed in Box 5 in *Unit 3: Generating Evidence.*

9 TNC (2016) provides a discussion of minimum standards of evidence for effectiveness.

10 Section 6 is adapted from ***Identifying and Using Evidence in Biodiversity Programming***.

11 ADS Chapter 201 includes "Identifying knowledge gaps during strategy development or project design and implementing plans to address them through evaluations, use of monitoring data, assessments, or other means" as an essential component of learning plans (page 129).

GLOSSARY

Adaptive management: An intentional approach to making decisions and adjustment in response to new information and changes in context (ADS 201, page 140). Evidence is an important source of information for adaptive management.

Aggregation bias: Occurs when effects measured at the level of a group do not hold for individuals belonging to the group.

Assumption: Used in *Evidence in Action* to refer to the logical connections between drivers, threats, and the status of biodiversity focal interests in a problem analysis or those that underlie anticipated results articulated in a program's theory of change.

Biodiversity focal interests: The species, habitats, and/or ecosystems that a program is working to conserve.

Confirmation bias: The tendency to selectively favor information that confirms existing beliefs and ignore contradictory information.

Confounding: Occurs when an outside factor co-varies with a presumed cause, making it difficult to establish the true cause of an observed effect.

Effectiveness: The degree to which an implemented project or activity achieves intended outcomes. Understanding the effectiveness of a strategic approach involves testing the assumptions that underlie a program's design.

Critical appraisal: The process of assessing findings from scientific research to judge their value and relevance in a particular decision-making context.

Evidence: The body of facts or information that serve as the basis for programmatic and strategic decision making in the Program Cycle (ADS Chapter 201, page 145). Used in *Evidence in Action* to refer to (1) individual findings or pieces of information used to help make a decision or support a conclusion; and (2) the body of findings or information providing support for (or countering) a belief or claim related to effectiveness or attribution.

Evidence-based approach: The conscientious, explicit, and judicious use of current, best evidence in program decisions. An evidence-based approach encompasses identification, use, and generation of evidence to increase program effectiveness.

External validity: The extent to which findings from one study can be applied to other contexts.

Grey literature: Documents and other materials produced outside of commercial or academic publishing and distribution channels, including government agencies, universities, corporations, non-governmental organizations, societies, and other professional organizations.

Groupthink: The tendency to adopt a belief that is held by other members of a group without critical evaluation of alternative viewpoints.

Meta-analysis: A research approach that combines the results from multiple studies.

Organizational data: General term for information that is collected and analyzed by organizations and institutions on their internal programs, processes, and performance.

Primary literature: Publications produced from original research or observations.

Program (and Programming): Used in *Evidence in Action* as a general term to encompass USAID project and activity levels.

Reproducibility: A relationship or result that is observed repeatedly.

Risky assumption: An assumption that has both a high likelihood of being invalid and would jeopardize program outcomes if invalid.

Quality of evidence: A measure of the extent to which a study's design is able to rule out alternative explanations for observed outcomes.

Secondary literature: Published materials that compile, interpret, or analyze other published sources.

Selection bias: An artificial skewing of the results caused by non-random selection of individuals, groups, or data.

Situation model: A graphic representation of a context or problem analysis (often called a conceptual model).

Strategic approach: A set of actions with a common focus that work together to address specific threats, drivers, and/or opportunities in order to achieve a set of desired results.

Strength of evidence: A measure of confidence in the evidence base (i.e., the aggregated findings from available studies) supporting or refuting a belief or claim.

Systematic map: An overview of the quantity and quality of evidence on a policy- or management-relevant topic. In contrast to a systematic review, a systematic map does not attempt to synthesize evidence.

Systematic review: A type of literature review using established protocols to acquire, critically appraise, and synthesize all available evidence relevant to a specific question.

Triangulation: Using multiple sources of evidence drawn from diverse research methodologies to support a conclusion.

LITERATURE CITED

Booth, A. 2006. Clear and present questions: formulating questions for evidence based practice. Library Hi Tech 24(3): 355–368. https://doi.org/10.1108/07378830610692127

Briner, R. B., and E. Barends. 2016. The Role of Scientific Findings in Evidence-Based HR. People + Strategy 39(2): 16–20.

Calver, M. C., B. Goldman, P. A. Hutchings, and R. T. Kingsford. 2017. Why discrepancies in searching the conservation biology literature matter. Biological Conservation 213: 19–26. https://doi.org/10.1016/j.biocon.2017.06.028

Casadevall, A., and F. C. Fang. 2010. Reproducible science. Infection and Immunity 78(12): 4972–4975. https://doi.org/10.1128/IAI.00908-10

Case, D. 2012. Looking for Information: A Survey of Research on Information Seeking, Needs and Behavior. Third edition. Emerald Group Publishing, Bingley, UK.

Collier, D., and J. Mahoney. 1996. Insights and pitfalls: Selection bias in qualitative research. World Politics 49(1): 56–91. https://doi.org/10.1353/wp.1996.0023

Department for International Development (DFID). 2014. Assessing the Strength of Evidence. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/291982/HTN-strength-evidence-march2014.pdf

Ferguson, L. 2004. External validity, generalizability, and knowledge utilization. Journal of Nursing Scholarship 36(1): 16–22. https://doi.org/10.1111/j.1547-5069.2004.04006.x

Grandage, K. K., D. C. Slawson, and A. F. Shaughnessy. 2002. When less is more: A practical approach to searching for evidence-based answers. Journal of the Medical Library Association 90(3): 298–304.

Guijt, I. 2013. ToC Reflection Notes 3: Working with Assumptions in a Theory of Change Process. http://p-shift.care2share.wikispaces.net/Theory+of+Change+Guidance

Jargowsky, P. A. 2005. Ecological Fallacy. Pages 715–722 in K. Kempf-Leonard, editor. Encyclopedia of Social Measurement. Elsevier, New York. https://doi.org/10.1016/B0-12-369398-5/00112-2

Jeuland, M. A., and S. K. Pattanayak. 2012. Benefits and costs of improved cookstoves: Assessing the implications of variability in health, forest and climate impacts. PLoS ONE 7: e30338. https://doi.org/10.1371/journal.pone.0030338

Jick, T. D. 1979. Mixing qualitative and quantitative methods: Triangulation in action. Administrative Science Quarterly 24(4): 602–611. https://doi.org/10.2307/2392366

Keeney, R. L. 2004. Making better decision makers. Decision Analysis 1(4): 193–204. https://doi.org/10.1287/deca.1040.0009

Kinderman, P., F. Murden, and K. Scott. 2016. Making Better Decisions: How Understanding Our Psychology Can Stop Us Falling Into the Bias Trap. British Psychological Society. https://www1.bps.org.uk/system/files/Public files/Comms-media/Making better decisions.pdf

Laine, C., and M. A. Winker. 2017. Identifying Predatory or Pseudo-Journals. World Association of Medical Editors. http://www.wame.org/identifying-predatory-or-pseudo-journals

Leisher, C., G. Temsah, F. Booker, M. Day, L. Samberg, D. Prosnitz, B. Agarwal, E. Matthews, D. Roe, D. Russell, T. Sunderland, and D. Wilkie. 2016. Does the gender composition of forest and fishery management groups affect resource governance and conservation outcomes? A systematic map. Environmental Evidence 5: 6. https://doi.org/10.1186/s13750-016-0057-8

Lewis, J. J., and S. K. Pattanayak. 2012. Who adopts improved fuels and cookstoves? A systematic review. Environmental Health Perspectives 120(5): 637–645. https://doi.org/10.1289/ehp.1104194

Pullin, A. S., and T. M. Knight. 2001. Effectiveness in conservation practice: Pointers from medicine and public health. Conservation Biology 15(1): 50–54. https://doi.org/10.1111/j.1523-1739.2001.99499.x

Pullin, A., G. Frampton, R. Jongman, C. Kohl, B. Livoreil, A. Lux, G. Pataki, G. Petrokofsky, A. Podhora, H. Saarikoski, L. Santamaria, S. Schindler, I. Sousa-Pinto, M. Vandewalle, and H. Wittmer. 2016. Selecting appropriate methods of knowledge synthesis to inform biodiversity policy. Biodiversity and Conservation 25(7): 1285–1300. https://doi.org/10.1007/s10531-016-1131-9

Roe, D., F. Booker, M. Day, W. Zhou, S. Allebone-Webb, N. A. O. Hill, N. Kumpel, G. Petrokofsky, K. Redford, D. Russell, G. Shepherd, J. Wright, and T. C. H. Sunderland. 2015. Are alternative livelihood projects effective at reducing local threats to specified elements of biodiversity and/or improving or maintaining the conservation status of those elements? Environmental Evidence 4: 22. https://doi.org/10.1186/s13750-015-0048-1

Salafsky, N., and K. H. Redford. 2013. Defining the burden of proof in conservation. Biological Conservation 166: 247–253. https://doi.org/10.1016/j.biocon.2013.07.002

Sandbrook, C. 2013. Biodiversity, Ecosystem Services and Poverty Alleviation: What Constitutes Good Evidence? International Institute for Environment and Development, London. http://pubs.iied.org/pdfs/G03715.pdf

Skelly, A. C., J. R. Dettori, and E. D. Brodt. 2012. Assessing bias: The importance of considering confounding. Evidence-Based Spine-Care Journal 3(1): 9–12. https://doi.org/10.1055/s-0031-1298595

Smith, J., and H. Noble. 2014. Bias in research. Evidence-Based Nursing 17(4): 100–101. https://doi.org/10.1136/eb-2014-101946

Steckler, A., and K. R. McLeroy. 2008. The importance of external validity. American Journal of Public Health 98(1): 9–10. https://doi.org/10.2105/AJPH.2007.126847

Sutherland, W. J., A. S. Pullin, P. M. Dolman, and T. M. Knight. 2004. The need for evidence-based conservation. Trends in Ecology & Evolution 19(6): 305–308. https://doi.org/10.1016/j.tree.2004.03.018

Sutherland, W. J., D. Spiegelhalter, and M. A. Burgman. 2013. Twenty tips for interpreting scientific claims. Nature 503(7476): 335–337. https://doi.org/10.1038/503335a

The Nature Conservancy (TNC). 2016. Conservation by Design 2.0 Guidance Document. https://www.conservationgateway.org/ConservationPlanning/cbd/Documents/CbD2.0 Guidance Doc Version 01.pdf

Thurmond, V. A. 2001. The point of triangulation. Journal of Nursing Scholarship 33(3): 253–258. https://doi.org/10.1111/j.1547-5069.2001.00253.x

Trochim, W.M. 2006. The Research Methods Knowledge Base, 2nd Edition. http://www.socialresearchmethods.net/kb/ (Version current as of October 20, 2006)

U.S. Geologic Survey (USGS). 2016. Turning Uncertainty into Useful Information for Conservation Decisions. SE CSC Fact Sheet 2016-02. https://globalchange.ncsu.edu/secsc/wp-content/uploads/SECSC_FactSheet_2016_02_Final.pdf

USAID. 2008. The Degrees of Evidence Framework. https://www.microlinks.org/library/program-evaluation-methodology-degrees-evidence-framework

Vasilakopoulos, P., F. G. O'Neill, and C. T. Marshall. 2011. Misspent youth: does catching immature fish affect fisheries sustainability? ICES Journal of Marine Science 68(7): 1525–1534. https://doi.org/10.1093/icesjms/fsr075

U.S. Agency for International Development
1300 Pennsylvania Avenue, NW
Washington, D.C. 20523
Tel: (202) 712-0000
Fax: (202) 216-3524
www.usaid.gov/biodiversity