An interdisciplinary approach to evidence for managing social-ecological systems

Authors: Edward T. Game[a,b], Heather Tallis[c,r], Lydia Olander[d], Steven Alexander[e], Jonah Busch[f], Nancy Cartwright[g], Seema Jayachandran[h], Elizabeth L. Kalies[i], Yuta J. Masuda[j], Anne-Christine Mupepele[k,q], Jiangxiao Qiu[l], Andrew Rooney[m], Erin Sills[n], William J. Sutherland[o]

[a]The Nature Conservancy, South Brisbane, QLD 4101, Australia

[b]Center for Biodiversity and Conservation Science, University of Queensland, St Lucia, QLD 4067, Australia

[c]The Nature Conservancy, Santa Cruz, CA, USA

[d] Nicholas Institute for Environmental Policy Solutions, Duke University, Durham, NC 27705, USA

[e] National Socio-Environmental Synthesis Center, University of Maryland, Annapolis, MD 21401, USA

[f] Center for Global Development, 2055 L Street NW, Fifth Floor, Washington, DC, 20036 USA

[g] Philosophy Departments, Durham University, DH1 3HN, UK & UCSD, La Jolla, CA 92093-0119, USA

[h] Department of Economics, Northwestern University, 2211 Campus Dr., Evanston, IL 60208

[i] The Nature Conservancy, Durham, NC, USA

[j] The Nature Conservancy, Seattle, WA, USA

[k] Department of Biometry and Environmental System Analysis; University of Freiburg, Tennenbacher Str. 4, 79106 Freiburg, Germany

[q] Chair of Nature Conservation and Landscape Ecology; University of Freiburg, Tennenbacher Str. 4, 79106 Freiburg, Germany

[l] School of Forest Resources & Conservation, Fort Lauderdale Research and Education Center, University of Florida, FL, USA 33314.

[m] Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA

[n] Department of Forestry and Environmental Resources, NC State University, Raleigh, NC 27695-8008

[o] Conservation Science Group, Department of Zoology, University of Cambridge, The David Attenborough Building, Cambridge, CB2 3QZ, UK

[q] Stockholm Resilience Centre, Stockholm University, 106 91 Stockholm, Sweden

[r] Department of Environmental Studies, University of California, Santa Cruz, CA, USA

Abstract

Discipline-specific approaches to evidence represent a barrier to the interdisciplinary understanding of social-ecological systems required to tackle sustainability challenges like those embodied in the UN Sustainable Development Goals. Here we propose an approach to evidence for interdisciplinary challenges where interventions are linked across environment, health, development, or other domains. This approach is centered around a logic model of hypothesized cause-effect linkages, an inclusive view of candidate evidence, and an assessment of evidence against four key principles; support from multiple types of evidence, consistency of effect in the candidate evidence, credibility of the sources, and applicability to the question of interest. Evidence-based policy in the sustainability space relies on an ability to engage evidence from across different fields. We see a strong need and opportunity to shape what evidence-based policy looks like in this space, and provide the foundations for a reasonable and practicable understanding of evidence for addressing social-ecological problems.

Introduction

Social and environmental systems are linked and as this relationship becomes ever more apparent, governments, communities, and organizations are increasingly faced with, and focused on, interdisciplinary problems. Indicative of this focus, 12 of the 17 United Nations Sustainable Development Goals (SDGs) directly reference linkages between human development challenges and environmental health [1]. The environment community increasingly argues that conserving nature can help address challenges in the health and development domains [2-5], and the health community is growing more aware of how environmental degradation is driving negative health outcomes [6, 7]. Given this growing understanding of the cross-sectoral nature of these challenges, the environment, development, and health communities are investing more in shared endeavors. The use of fire to clear tropical peatland forests illustrates these connections and shared interests across the environment, health, and development communities [8]; fire is an important tool for agricultural production, but it leads to significant carbon emissions, the loss of forests and associated biodiversity, and human respiratory illness and mortality linked to smoke.

Decisions about interventions and strategies to solve interdisciplinary problems should be supported by evidence on their causal effects. This is widely agreed, and there is momentum towards evidence-based decisions in a range of fields [9-11]. However, there is no universal agreement about what constitutes strong evidence in support of pursuing a specific intervention or strategy [12, 13]. With the appearance of terms like “alternative facts” that may impact clear communication of opinion vs scientific judgment based in facts, we are arguably in a period when a standard and generally accepted approach to evidence is needed more than ever [14, 15].

Standards for assessing and communicating the strength of evidence are desirable for three principal reasons. First, choices about interventions and strategies require assessing both positive and negative effects, and conclusions on the strength of evidence are directly related to our confidence in the evidence for these effects [16]. Second, there can be a wide range of potentially relevant candidate evidence available to inform policy choices, but it is important to assess and communicate how well that evidence supports an intervention or strategy. Third, in the absence of a well-established framework for assessing the strength of evidence, an understanding about whether candidate evidence provides strong support or not, and why, will remain a source of confusion rather than providing clarity.

Although there is no universal approach to assessing evidence, there is emerging consensus within some disciplines, such as in clinical medicine where the GRADE approach [16] has been widely adopted. In addition, there are other fields where structured approaches, including GRADE with some topic-specific considerations, have been consistently used to address questions of evidence concerning toxicology and public health [17]. These discipline-specific approaches have grown out of calls for evidence-based practice and systematic reviews under the Cochrane Collaboration for health interventions and the Campbell Collaboration for social interventions since the 1990s. Similar calls for evidence-based practice in environmental management and conservation [18, 19] have more recently encouraged the growth of systematic reviews in that field; however, a broader consensus on an approach to evaluating and determining the strength of evidence remains elusive [20]. The largely ad hoc approach to evidence assessment practiced in the environment and conservation space [11] is increasingly juxtaposed with the more consistent approach to evidence in other fields.

There is a strong temptation to simply apply standards that have gained traction in other disciplines to the environmental domain [e.g., 21]. However, because of differences in the types of interventions being considered, the types of candidate evidence available, and the types of questions being asked of the evidence, this translation of evidence assessment approaches is unlikely to reflect important considerations that are critical to environmental sciences. One much debated example is the medical sector's randomized controlled trials (RCTs) with large sample sizes as the standard for robust evidence about casual relationships. The environment and conservation community seldom uses RCTs because they are so often interested in interventions that are difficult to randomize and for which it is impossible to blind those delivering the intervention, those receiving it, and those who analyze the outcomes sufficiently to guard against statistical bias. Much of the available evidence comes from studies that are observational rather than experimental, involve little replication, and often do not even include observations on the baseline or controls [22]. Use of an unadjusted GRADE framework would result in all of the evidence being judged as inadequate to support decision making and would not effectively stratify stronger evidence from weaker evidence when a decision is nevertheless required based on the available evidence for the environmental and conservation community.

There are also differences in the focus of evidence assessment across fields. Medical evidence assessment schemes are designed primarily to evaluate support for general claims: the causal pathways or treatments investigated are anticipated to work in similar ways even across different populations, or at least to have roughly the same effects on average. In contrast, the environment and conservation fields often deal with complex social-ecological interactions [23], where context is critical and evidence assessment has tended to focus on the strength of support in particular social, political, economic, and biophysical contexts.

These discipline-specific approaches to evidence, driven in large part by the types of evidence historically available in different disciplines, represent a barrier to the sort of interdisciplinary understanding of social-ecological systems embodied in the SDGs. Assessing the strength of evidence in order to decide whether to pursue an intervention (or which of many interventions to pursue) to solve these complex socio-ecological problems requires considering impacts that span different domains (e.g., environment, health, development). For instance, in the case of managing fire on tropical peatlands the outcomes of interest extend into development and health. The impact of smoke from peat fires on human health is a medical, economic, and meteorological consideration; the impact of stopping the use of fires on agricultural productivity is a development and agricultural question; the impact of escaped fires on forests and biodiversity is an ecological and meteorological question; and the individual or organizational response to a proposed intervention is an economic and behavioural question. There is clear logic to these linkages, but evidence-based policy in this case will require drawing on a wide variety of evidence from multiple fields. Using a medical standard to assess evidence, for example, may exclude much of the candidate evidence (e.g., modelling studies), and likely offer little ability to discriminate the relative strength between the available evidence and to assess the overall degree to which the total body of evidnce suposrts (or not) the intervention . Alternatively, assessing evidence in a way that fails to evaluate validity to the same standard as medicine could undermine credibility and collaboration across disciplines.

We aimed to develop a reasonable and practicable approach to evidence assessment [24] for these interdisciplinary challenges where interventions are linked across environment, health, development, or other domains. We see a strong need and opportunity to shape what evidence-based policy and practice look like in this space. We interpret policy as any institutionalized set of behaviors or practices, such that management interventions are a form of policy [25]. The intended audience for this contribution is both individual researchers and organizations working on management interventions for natural resources and/or social-ecological systems. The need to draw on a broad set of evidence from across disciplines while maintaining credibility within these disciplines means that the approach to evidence we propose is less strict but more reasonable and practicable than a hierarchy of evidence based solely around study designs. It is important to clearly state (and to the extent practical, to agree on) a set of principles to guide the assessment of evidence in this interdisciplinary space, and here we offer a proposal for what these principles could be.

Evidence assessment approach

An assessment of evidence must deliver an understanding of what is sound and defensible from a scientific perspective. Although decisions will be influenced by preferences and values, we often expect a scientific basis for why we are undertaking the actions we are where we are. To this end, we propose an approach and set of core principles that draw from existing standards in health, development and environment domains, and constitute a reasonable and practical way to begin evaluating the level of support for a hypothesis regarding environmental interventions in light of all available pieces of evidence. The key steps of the approach are:

1. Create a logic model to identify key **causal pathways** from intervention to outcomes.

2. For each causal link, state a clear question using **PI/ECO framing** (Population, Intervention/Exposure, Comparator, Outcome).
3. Gather **candidate evidence** (including different types of evidence) relevant to the question.
4. Apply a set of **evidence principles** regarding the types, consistency, credibility and applicability of evidence to the process of assessing and synthesizing a body of candidate evidence.

Each of these evidence assessment steps is described in more detail below.

Causal pathways (Step 1)

Our principal context for evidence assessment begins with a model that identifies the key hypothesized causal pathways from interventions to outcomes. These models typically begin as graphical representations of mental models linking interventions, results, and outcomes (Figure 1). Causal models have become familiar tools in environmental management and conservation [e.g., results chains in the Open Standards, 26, 27], and have been identified as valuable frameworks with which to understand evidence [28]. Importantly, causal models (known by various terms) are useful to communicate evidence relationships within the field of environmental conservation, but also across disciplines because similar models are also widely used as tools in numerous other fields [29-32]. As such, these causal pathway models support the interdisciplinary nature of contemporary environmental problems, both through their familiar structure and the ability to represent hypothesized relationships in multiple domains (e.g., ecology, health, development) in the same causal model.

When done well, causal models clarify the assumptions and conditions for which evidence is sought and the logical link between any two nodes. For example, in Figure 1, the link between smoke events and human exposure to particulate matter assumes that smoke events, had they occurred, would have resulted in particulate matter exposure to people. The linked node structure of casual models makes them a natural frame for treatment-outcome (T-O) evidence between adjacent nodes and across nodes, but they can also help to clearly identify other kinds of evidence that are important as well as the relationships among a broad range of assumptions, results, and outcomes of interest. For instance, the assumption that reducing or stopping the loss of peat forest will lead to increased viability of tiger populations depends on high-quality measurement of the presence of tigers, which, in turn, is typically dependent on evidence for why a particular survey method will deliver an estimate indicative of the true tiger population. The evidence assessment approach described here is focused on T-O evidence, but we acknowledge there will be a lot of evidence associated with a project that does not relate directly to the relationship between a treatment and outcome. For example, evidence that establishes confidence in the approach used to map forest cover does not directly relate to treatment or outcome, but rather to the measurement of key results or variables, which are similarly important in terms of confidence in the science-basis of a particular project [33].

Importantly, causal pathway models constitute candidate evidence in and of themselves. Causal models are not random collections of nodes and linkages, but instead are generally built based on expert knowledge, and the plausible mechanistic reasoning they describe provides support for the hypotheses they outline [34].

PI/ECO framing (Step 2)

Each link in the causal chain represents a treatment-outcome (T-O) relationship. This relationship can be framed with the PI/ECO approach, which was developed in the field of medicine to clearly define the factors related to a specific research question, but can be applied to various disciplines with minor modification [19]. The basic structure to address the specific question is clarified by the general structure of what effect does intervention "I" have on the outcome(s) "O" of interest for on population "P", where:

- **P**opulation is the unit of study, which may be species, ecosystems, or people in a given area and on which the intervention is focused,
- **I**ntervention/**E**xposure is the treatment, including management or policy actions, or other environmental variables to which the population is exposed, and recalling that 'treatment' now designates any node in the causal chain from proposed intervention to targeted outcome,
- **C**omparator or Control is a counterfactual comparison based on either no intervention, an alternative intervention, or some other well-defined baseline (e.g., lower exposure level),
- **O**utcome(s) are the outcomes of interest, positive and negative, resulting from the proposed intervention, again recalling that all nodes in the causal chain except the first, proposed intervention, are 'outcomes' of previous nodes.

While the entire causal chain represents a PI/ECO from intervention (i.e., incentives to reduce use of fire) to outcome (i.e., increased jobs, reduced illness, increased animal populations, etc.), each node-to-node relationship also represents an intervention (treatment) and outcome, with the control implicit in the assumption of that pathway or represented by a separate pathway. For example, the causal chain predicts that reduced loss of peat swamp forests due to uncontrolled fire will result in increased populations of forest-dependent species. In this case, the relationship would be framed with the subject of interest of peat swamp forest habitat (population) which is exposed to fire (treatment) in comparison to habitat without fire (control), and the differences in numbers of forest-dependent animals (outcome) would be compared. Evidence that fits this framework, for each link, would be gathered in the next step. It is also important to note that the model of the chain connecting intervention and outcome is a reduction of a fuller causal map that depicts the other features that have to be in place for one node to lead to the next (the interactive or moderator variables that epidemiologists graph in causal pies). Evidence about features of the larger causal map can also be important for reliable conclusions about the chain between intervention and outcome.

Candidate evidence (Step 3)

Candidate evidence can come from a wide range of sources (e.g., qualitative studies, quantitative studies, models, etc., see Table 1). Because of the complex and context-dependent nature of social-environmental problems, we require an approach to evidence assessment that can make use of multiple types of candidate evidence. Doing so requires deliberately considering multiple types of candidate evidence, and having an assessment approach that is able to combine different types of candidate evidence. Existing evidence assessment or grading schemes commonly exclude much candidate evidence by not combining different types of evidence in reaching a conclusion about evidence strength. Rather, they make an assessment based on the “strongest” evidence available, for example, discarding qualitative evidence in favour of quantitative evidence. This effectively excludes much of the evidence for social-environmental problems, where rigorous, quantitative studies [e.g., 35] are often unavailable. It also is inapplicable for evaluating evidence for matters other than causal relationships, for instance about the reliability of measurements or the plausibility of assumptions made in modelling. While this highlights the need to generate more research on complex social-environmental interactions and problems, it also identifies the need and challenge for researchers and decision-makers to apply a broader lens to what constitutes candidate evidence.

We also require an evidence assessment approach that ensures that the evidence considered addresses the specific question in view. For instance, evidence that a particular cause-effect relationship exists in situations of a given type needs to be combined with evidence that the situation being evaluated is indeed of this type. The PI/ECO framework described in the previous step is designed to get clarity around the question in view, and therefore guide the search for candidate evidence. In addition, evidence that an intervention can produce the desired outcome via a particular causal pathway given the current state of the system is only relevant when coupled with reasons to assume that the state of the system will not shift enough to invalidate the causal pathway. These typically ancillary, but equally important pieces of evidence, are often based on expert judgement and should not be excluded from our assessment of evidence if our goal is to evaluate the level of support for a hypothesis.

We acknowledge that the set of evidence types described in Table 1 is not the only possible taxonomy of evidence types. However, it effectively characterizes the underlying source and important differences between the range of candidate evidence used, both explicitly and implicitly, in the management of social-ecological systems. These types are not mutually exclusive and a candidate piece of evidence will, in many cases, reflect more than one of these types of evidence.

In the ideal, quantitative studies have the potential to produce stronger inference than the other types of evidence, and some evidence grading schemes only admit quantitative studies. In practice, however, there is no inherent rank order in terms of the strength of evidence these six types produce for answering the suite of questions present in a casual pathway, nor is there general agreement that one entire type is better than another. For instance, the methods and design used to create a type of evidence is also important. A well done expert elicitation is likely to be superior to a poorly done quantitative study for many of questions [36], and may be the only type of evidence available in some cases. Some forms of evidence are more appropriate for particular questions than others, for example, whether increased knowledge is likely to result in a change in behavior may be better informed by qualitative evidence that describes context than quantitative.

However, there is a general agreement that some modes of information collection or generation produce less bias than others, and some give more precise or more reliable answers than others, and that “methodological standards” exist within each information type, and can be used to determine the quality of evidence provided [37].

Evidence principles (Step 4)

We propose a set of principles and rules that can act as an initial screen and synthesis tool for identifying strong evidence across fields (Figure 2):

a) Support for a causal link comes from **multiple types of evidence**
b) There is **consistency of effect** in the candidate evidence
c) The candidate evidence comes from **credible sources**
d) The candidate evidence is **applicable** to the question of interest

Where the body of evidence for a specific question, or set of questions, about the causal pathway is consistent with all four principles, that evidence in support or in contrast to the question is considered strong. Where the body of evidence for a specific question is not in keeping with one or more of the evidence principles, then the evidence is considered weaker and the basis for the lack of confidence should be clearly stated. In cases where evidence is weaker, expertise from the relevant field is needed to help judge the quality of the individual evidence and to transparently describe the strengths, weaknesses, and applicability to the specific research question. If the judgment about the overall conclusion based on the evidence remains unclear, policy planning should reflect that lack of certainty.

These principles reflect characteristics of a body of candidate evidence that would provide strong support for a hypothesized causal link or for any of the other matters that affect the assessment of effectiveness, while retaining the ability to incorporate a variety of different types of evidence from multiple disciplines. Recognizing the existence of well-established evidence standards in some fields, there remains a need for a practical and reasonable way to interpret evidence with respect to interdisciplinary social-environmental problems. Our evidence principles provide an entry point that, if met, allow researchers and practitioners to assess whether a set of evidence is relatively strong, without applying the standards of each relevant field. If one or more of the principles is not met it does not necessarily follow that evidence is weak, just that it will require expertise familiar with methodological standards in the relevant field in order to assess the quality of individual pieces of evidence, always keeping in view the question that the evidence is supposed to address. These are not novel principles, rather they are widely agreed on across disciplines, and we believe could be useful when interdisciplinary evidence evaluation is necessary.

For causal links or pathways of interest, an initial assessment of candidate evidence should evaluate this evidence against the four principles described below. This does not mean a systematic review, rather an effort to seek and gather candidate evidence across the different types of evidence. All four of these principles require individual judgement and will be implemented variably by different individuals. This element of subjective judgment is a near universal feature of evidence grading schemes. In Supplementary Information S1 we illustrate these evidence principles with an example related to the use of fire in clearing tropical peatlands. Here, we discuss further each of the proposed characteristics of strong evidence.

Multiple types of evidence. Each type of evidence described in Table 1 has its own strengths and weaknesses. Where a hypothesized causal linkage is supported by more than one type of evidence, we conclude that it confers greater confidence that the association exists than where evidence is only available from a single type. The basic premise for this principle is that methodological variation between evidence types reduces the likelihood that the reported link is due solely to how a study was conducted and that it is unlikely that the limitations of different types of evidence would each bias the findings in the same direction. From Figure 1, if an assessment of the linkage between fire and agricultural productivity drew on evidence from theory linking the chemistry of forest ash to a neutralization of acid soils, expert knowledge linking the use of fire to increased agricultural productivity of peat soils, and measurement results confirming the acidity of peat soil in the area of interest, it would be reasonable to consider evidence for that linkage strong. Evidence assessment schemes commonly recognize that multiple, unrelated lines of evidence provide stronger overall evidence assessment [22, 33], but do not generally equate multiple lines of evidence with multiple types of evidence as we do here.

Consistency of effect. Where the body of candidate evidence is consistent in its findings, this increased confidence of the claim in question. *Consistency* implicitly places value on having a larger amount of evidence and has multiple dimensions to it. Consistency can be considered for the direction (or sign) of the effect, its size, and the range and variance of the effect. Consistency of effect across studies is considered a central tenet of many evidence assessment schemes. This is particularly true of those schemes looking to assess general claims of relationships between a treatment and outcome for medical interventions [16] or exposure and hazard for public health questions [38]. While it is reasonable to consider a consistent effect as indicative of strong evidence, variation in findings across studies does not preclude strong evidence because there can be plausible explanations for this variation relating to the context and approach of different studies.

Credible sources. Where candidate evidence is available from sources widely seen as credible, it provides confidence in characterizing evidence as strong. Confidence is instilled by the fact that these sources have standards and checks in place to ensure methodologies are appropriately matched to the study question and that the impact of motivational bias on findings is minimized. Although there is no objective rank of sources by credibility, and no source provides an unequivocal guarantee of study quality, the process of publication in peer reviewed journals is designed explicitly to improve and support the credibility of findings. It would be remiss not to take advantage of this process to provide an indication of the credibility of different pieces of candidate evidence. There is, of course, a spectrum of credibility within the peer reviewed journal system, but one that defies easy characterization. Candidate evidence that comes from sources outside the peer reviewed journal system may still be widely seen as highly credible, for example, reports that come from United Nations' organizations. However, where all the candidate evidence comes from sources with the potential for perceived motivational bias (e.g., because of an organization's agenda or funding), a thorough assessment of study design would be a necessary part of evidence assessment.

Applicability. Where there is a good fit between the body of candidate evidence and the question of interest (as, for instance, described using the PI/ECO framing), it is reasonable to assume that this evidence is relevant and therefore has greater potential for being considered strong. How applicable candidate evidence is to the question at hand is dependent on both the context for the evidence (for instance, the presence or absence of similar moderator variables and ability to account for their effect on the outcome) and the methods used, such as the type of treatment or the outcome measures (e.g., was the same outcome of interest measured or was a related outcome measured). The applicability principle is akin to the assessment of evidence directness found in systems like GRADE [16], but with greater flexibility as to what constitutes fit with the question and context in focus.

Across previously published evidence grading schemes, a large number of factors have been proposed as relevant to an assessment of evidence strength. However, among the authors of this paper (experts from a range of disciplines), the four principles above were the only ones considered unequivocal indicators of strong evidence. As such there was agreement that these principles should be the initial considerations in any evidence evaluation of social-ecological problems.

Discussion

We propose an approach to evidence that is centered around a logic model of hypothesized cause-effect linkages, an inclusive view of candidate evidence, and an assessment of the body of candidate evidence against four key principles. This approach was developed in response to the need for an interdisciplinary understanding of evidence where environmental outcomes are linked to outcomes in health, development, or other sectors. The approach described here is not intended to be a formal evidence grading standard that classifies one piece of evidence as better than another, or a set of strict criteria for whether a body of evidence is or is not strong evidence. Rather, our aim is to provide the foundations for a reasonable and practicable understanding of evidence for addressing social-ecological problems.

The extensive linkages between social and environmental systems, and the reflection of these linkages in project design and objectives, increasingly place environmental problems firmly in an interdisciplinary space. Understanding these linkages is also central to sustainability research more generally. As such, evidence-based policy in this sustainability space relies on the ability to engage evidence from across different fields. For instance, in the case of fires on tropical peatlands introduced at the start of this manuscript, a comprehensive understanding of the cause-effect relationships linked to the strategy would require input from ecology, agriculture, development, economics, public health, and meteorology. The methodological standards within each of these domains is different, so a consistent understanding of evidence to assess different interventions demands an approach to evidence that can embrace and integrate across these different standards. Without this, it will be challenging, if not impossible, to reach a shared understanding of how different policies and interventions support the sustainability of peatland agriculture, an issue of significant global importance [39, 40].

It is unlikely that an organization with an environmental (or any other single disciplinary) focus aiming to assess the evidence for a strategy that crosses environmental and social domains will have easy access to experts from all the relevant domains. In developing the approach described here, we had in mind organizations or researchers aiming to use evidence under clear resource constraints in terms of expertise and time. The proposed approach to framing the search for evidence (causal pathways and inclusive definitions of candidate evidence) and the four principles by which to screen the candidate body of evidence are intended to be feasible with the resources generally available to large organizations. This includes staff with masters or doctoral degrees and extensive literature access. Such staffs have the training and resources to glean the main findings (e.g., effect size and statistical significance for a quantitative study) from a publication reporting evidence, but not necessarily the training (or time) to independently assess the quality of each piece of evidence produced by researchers from many different fields. We recognize that even this level of resource commitment is substantial and will make interdisciplinary evidence assessment described here difficult for many organizations. However, the approach here forms the foundation for an understanding of evidence that can bridge disciplinary barriers in the case of projects with linked environmental and social outcomes and improve what evidence-based policy looks like in this interdisciplinary space.

The approach and evidence principles described here allow for a rapid appraisal of whether evidence can be broadly interpreted as strong or whether a more detailed evaluation of candidate evidence is required. This contrasts with evidence rubrics that often grade evidence into 4 categories (e.g., strong, moderate, fair, weak). We argue that once a body of candidate evidence has been evaluated against the four evidence principles, an overall conclusion that evidence is moderate or fair is of limited utility without additional information. In making strategy choices, we want to know where we have strong support for a hypothesis, and if there is not strong support, in what way it is not strong, or in other words, which of the four evidence principles above is not currently satisfied. Grading evidence into multiple categories requires a more detailed, often more subjective, and possibly more inconsistent assessment process, that in many cases may not be worth the extra effort given the remaining uncertainty. For instance, it is unclear whether and how decisions would differ when made under a finding of fair versus a finding of moderate that would not require understanding in what way exactly support for a hypothesis is currently limited. This argument is similar to the approach to evidence found in Norris et al. [22] who identify where evidence supports a hypothesis, or if not, how and in what way it does not. By simply looking at cases where strong evidence is readily available, we also avoid problems associated with scoring systems that require weighting different types of data. In the context of complex social-ecological problems this is positive as it allows the evidence assessment to draw on a broad set of evidence types as appropriate for each hypothesis.

We believe that the four principles of strong evidence proposed here can help bridge across disciplines because, at least amongst the authors of this paper, there was agreement that they are indicative of strong evidence and do not require an expert knowledge of a particular discipline to assess. Where a body of evidence is not consistent with one or more of these principles, it does not preclude evidence being strong. Rather, it requires a more detailed evaluation of individual pieces of evidence against the methodological standards of the relevant discipline and the relevance to the question at hand. This requires engaging expertise in the related discipline. The need for environmental organisations and agencies to seek out and engage experts from other disciplines, such as health and development, poses a challenge for evaluating evidence in these interdisciplinary projects. However, it may be necessary given the current significant differences in methodological approaches in different disciplines, and is also likely to foster stronger interdisciplinary engagement to tackle linked and complex social-ecological problems that are at the heart of sustainability.

Acknowledgements

This collaboration was supported by a grant from the David and Lucile Packard Foundation to HT, ETG, and LO. WJS is funded by Arcadia.

References

1. Griggs, D., et al., *Policy: Sustainable development goals for people and planet.* Nature, 2013. **495**(7441): p. 305-307.
2. Sachs, J.D., et al., *Biodiversity conservation and the millennium development goals.* Science, 2009. **325**(5947): p. 1502-1503.
3. Tallis, H., et al., *An ecosystem services framework to support both practical conservation and economic development.* Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**: p. 9457-9464.
4. Sandifer, P.A., A.E. Sutton-Grier, and B.P. Ward, *Exploring connections among nature, biodiversity, ecosystem services, and human health and well-being: Opportunities to enhance health and biodiversity conservation.* Ecosystem Services, 2015. **12**: p. 1-15.
5. Bottrill, M., et al., *What are the impacts of nature conservation interventions on human well-being: a systematic map protocol.* Environmental Evidence, 2014. **3**(1): p. 16.
6. Almada, A.A., et al., *A case for Planetary Health/GeoHealth.* GeoHealth, 2017. **1**(2): p. 75-78.
7. Whitmee, S., et al., *Safeguarding human health in the Anthropocene epoch: report of The Rockefeller Foundation–Lancet Commission on planetary health.* The Lancet, 2015. **386**(10007): p. 1973-2028.
8. Marlier, M.E., et al., *El Niño and health risks from landscape fire emissions in southeast Asia.* Nature climate change, 2013. **3**(2): p. 131-136.
9. Slavin, R.E., *Evidence-based education policies: Transforming educational practice and research.* Educational researcher, 2002. **31**(7): p. 15-21.
10. Sackett, D.L., et al., *Evidence based medicine: what it is and what it isn't*. 1996, British Medical Journal Publishing Group.
11. Sutherland, W.J. and C.F.R. Wordley, *Evidence complacency hampers conservation.* Nature Ecology & Evolution, 2017. **1**.
12. Cartwright, N., *Are RCTs the gold standard?* BioSocieties, 2007. **2**(1): p. 11-20.
13. Deaton, A. and N. Cartwright, *Understanding and misunderstanding randomized controlled trials*. 2016, National Bureau of Economic Research.
14. Goodman, S.N., *Aligning statistical and scientific reasoning.* Science, 2016. **352**(6290): p. 1180-1181.
15. Goldman, G.T., et al., *Ensuring scientific integrity in the Age of Trump.* Science, 2017. **355**(6326): p. 696-698.
16. Guyatt, G.H., et al., *GRADE: an emerging consensus on rating quality of evidence and strength of recommendations.* BMJ (Clinical research ed.), 2008.
17. Morgan, R.L., et al., *GRADE: Assessing the quality of evidence in environmental and occupational health.* Environment international, 2016. **92**: p. 611-616.
18. Sutherland, W.J., et al., *The need for evidence-based conservation.* Trends in Ecology & Evolution, 2004. **19**(6): p. 305-308.

19. Pullin, A.S. and G.B. Stewart, *Guidelines for systematic review in conservation and environmental management.* Conservation biology, 2006. **20**(6): p. 1647-1656.

20. Collaboration for Environmental Evidence, *Guidelines for Systematic Review and Evidence Synthesis in Environmental Management. Version 4.2.* 2013, Collaboration for Environmental Evidence.

21. Wilkie, D. and J. Ginsberg, *Why adoption of the medical model would cure conservation impact evaluation.* Conservation: the source for environmental intelligence. University of Washington, Seattle, 2014.

22. Norris, R., et al., *Analyzing cause and effect in environmental assessments: using weighted evidence from the literature.* Freshwater Science, 2011. **31**(1): p. 5-21.

23. Game, E.T., et al., *Conservation in a Wicked Complex World; Challenges and Solutions.* Conservation Letters, 2014. **7**(3): p. 271-277.

24. Cartwright, N., A. Goldfinch, and J. Howick, *Evidence-based policy: where is our theory of evidence?* Journal of Children's Services, 2010. **4**(4): p. 6-14.

25. Game, E.T., M.W. Schwartz, and A.T. Knight, *Policy relevant conservation science.* Conservation Letters, 2015. **8**(5): p. 309-311.

26. Conservation Measures Partnership, *Open Standards for the Practice of Conservation, Version 3.0*. 2013.

27. Fargione, J., et al., *Conservation by Design 2.0 Guidance Document, Version 1.* 2016, The Nature Conservancy: Washington D.C.

28. Norton, S.B. and K.A. Schofield, *Conceptual model diagrams as evidence scaffolds for environmental assessment and management.* Freshwater Science, 2017. **36**(1): p. 231-239.

29. Roberts, D. and N. Khattri, *Designing a results framework for achieving results: a how-to guide.* Washington, DC: Independent Evaluation Group, World Bank, 2012.

30. Brest, P., *The power of theories of change.* Stanford Social Innovation Review, 2010. **8**(2): p. 47-51.

31. World Health Organisation, *WHO Evaluation Practice Handbook*. 2013, World Health Organization: Geneva, Switzerland.

32. Gugerty, M.K., D. Karlan, and D. Welsh, *Innovations for Poverty Action. Guiding Your Program to Build a Theory of Change*. 2016, Innovations for Poverty Action: Washington D.C.

33. Mupepele, A.C., et al., *An evidence assessment tool for ecosystem services and conservation studies.* Ecological Applications, 2016. **26**(5): p. 1295-1301.

34. Cartwright, N. and J. Hardie, *Evidence-based policy: a practical guide to doing it better*. 2012: Oxford University Press.

35. Jayachandran, S., et al., *Cash for carbon: A randomized trial of payments for ecosystem services to reduce deforestation.* Science, 2017. **357**(6348): p. 267-273.

36. Martin, T.G., et al., *Eliciting expert knowledge in conservation science.* Conservation Biology, 2012. **26**(1): p. 29-38.

37. Khagram, S. and C.W. Thomas, *Toward a Platinum Standard for Evidence-Based Assessment by 2020.* Public Administration Review, 2010. **70**(s1).

38. Rooney, A.A., et al., *Systematic review and evidence integration for literature-based environmental health science assessments.* Environmental health perspectives, 2014. **122**(7): p. 711.

39. Page, S.E., J.O. Rieley, and C.J. Banks, *Global and regional importance of the tropical peatland carbon pool.* Global Change Biology, 2011. **17**(2): p. 798-818.

40. Tacconi, L., *Preventing fires and haze in Southeast Asia.* Nature Climate Change, 2016. **6**(7): p. 640-643.

Table 1 Types of candidate evidence

Evidence type	**Description**
Quantitative studies	Studies based on inference through numerical data and analysis that describe the relationship between parts of a system. Quantitative studies may be experimental, quasi-experimental or observational.
Qualitative studies	Studies based on inference through a thorough understanding of a case(s) under study, but unable to characterize an absolute numerical relationship between parts of a system.
Models	Representation of how a system (or part of a system) functions. Potentially a tool for prediction, e.g. working out where smoke from a particular peat fire is likely to end up. Models can be conceptual or mathematical, and are typically, but not always, used in conjunction with the results of quantitative studies, theory, or expert knowledge.
Expert knowledge	The judgement of those with specialized knowledge obtained through training or experience. This includes local knowledge, traditional knowledge, and subject matter expertise.
Theory	A scientifically accepted general principle or body of principles offered to explain phenomena.
Interpretation of measurement results	Information gained from measurement which may or may not be part of study. We require confidence in the veracity of the measurement approach.

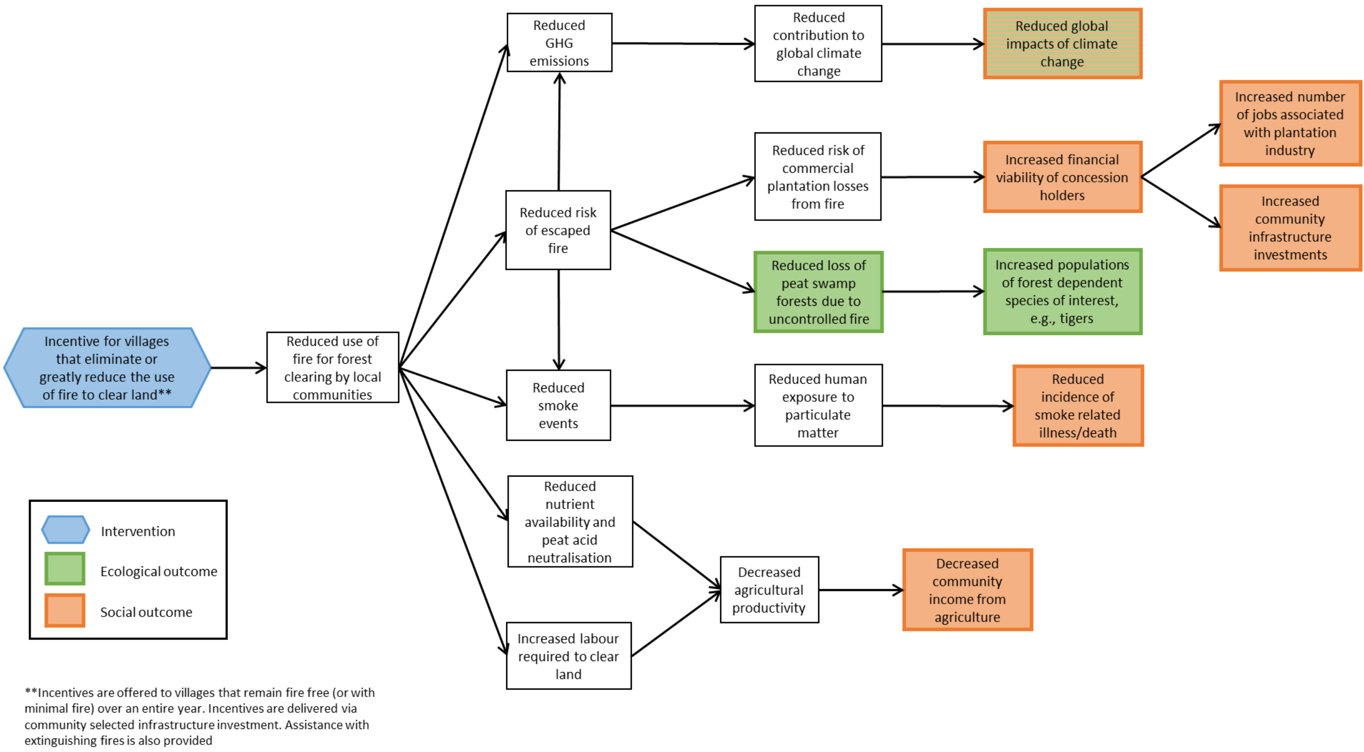

Figure 1 Causal pathway graphical model for a project in the peatlands of Sumatra, Indonesia, where an intervention designed to reduce the use of fire for land clearing is linked to a set of ecological and social outcomes.

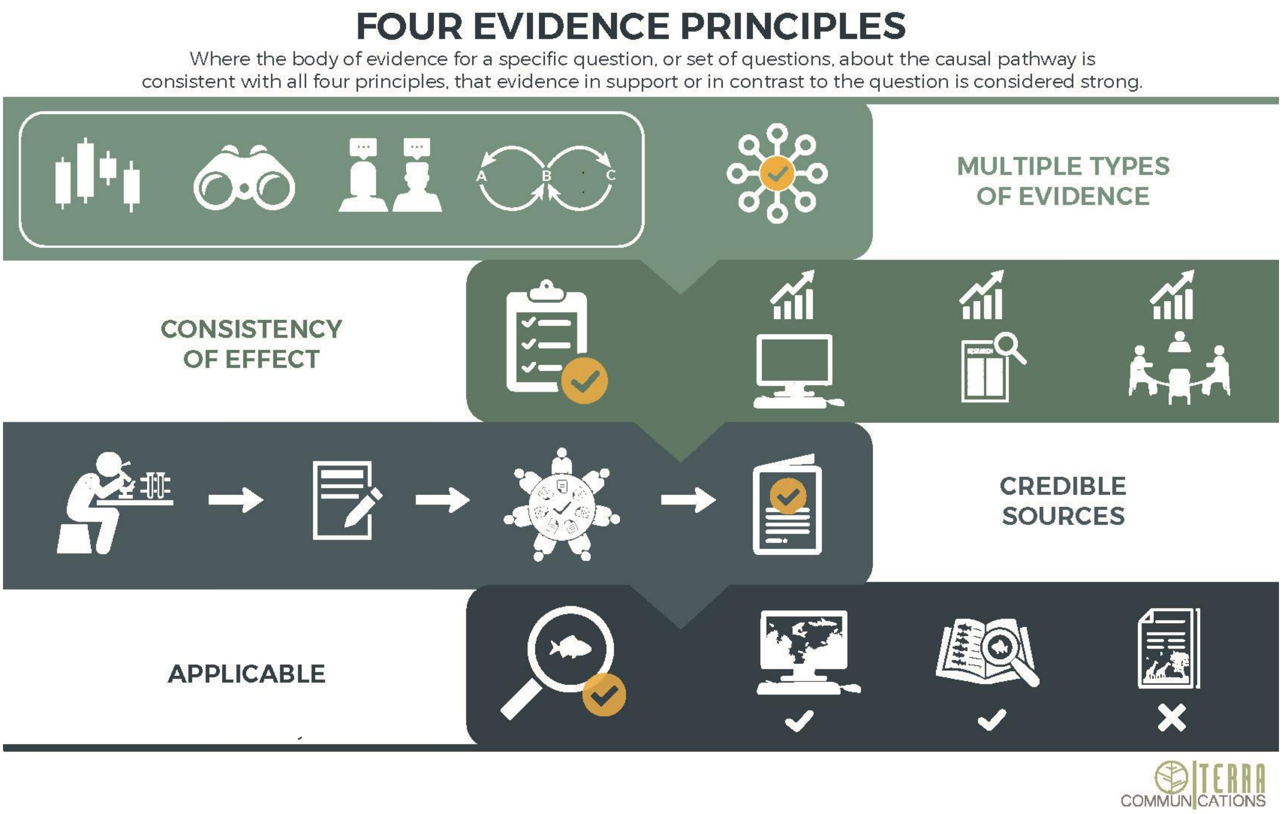

Figure 2 Schematic of the four principles that underpin an interdisciplinary approach to evidence