

Cross-discipline evidence principles for sustainability policy

Evidence-based approaches to sustainability challenges must draw on knowledge from the environment, development and health communities. To be practicable, this requires an approach to evidence that is broader and less hierarchical than the standards often applied within disciplines.

Edward T. Game, Heather Tallis, Lydia Olander, Steven M. Alexander, Jonah Busch, Nancy Cartwright, Elizabeth L. Kalies, Yuta J. Masuda, Anne-Christine Mupepele, Jiangxiao Qiu, Andrew Rooney, Erin Sills and William J. Sutherland

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, problems that are complex, wicked and transgress traditional disciplinary boundaries. Indicative of this focus, 12 of the 17 United Nations' Sustainable Development Goals directly reference linkages between human development challenges and environmental health, and thus, evidence-based approaches to the Sustainable Development Goals must draw on knowledge from the environment, development and health domains. In response, the environment, development and health communities are investing more in shared, cross-disciplinary approaches to evaluating the effectiveness of interventions. This effort requires a broader, less-hierarchical approach to evidence than those often applied within disciplines.

Different kinds of knowledge arise from research in disciplines that make fundamentally different philosophical and methodological assumptions¹ or from knowledge that is entirely outside the epistemological framework of most research². Because different types of knowledge are useful for different purposes³ and it is not possible to evaluate all knowledge using the same criteria, how knowledge, including evidence, is defined and interpreted has a major impact on how understanding complex problems and potential interventions is approached. Overcoming the cultural and philosophical barriers of working with very different forms of knowledge remains a general challenge^{2,4}. Here we address the narrower, but still difficult, problem of integrating different types of evidence underpinned by the assumption that it is possible to predict outcomes of an intervention.

A common interpretation of evidence among natural sciences and more positivist social science approaches is, the body of information relevant to judging whether a hypothesis is likely to be true or not⁵. Assessing the strength of evidence that implementing an intervention will result in a particular outcome (a causal hypothesis), is a critical step in evidence-based decision-making about whether, when and where to pursue an intervention (or which of many, possibly untested, interventions to pursue). Cross-disciplinary approaches to social–environmental challenges require causal associations and evidence across various domains to be considered⁶. For example, tackling the use of fire to clear tropical peat forests crosses the environment, health and development communities; 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.

Although there is no universal approach to assessing evidence, there is convergence within some disciplines, such as in clinical medicine⁷, toxicology and public health⁸. These discipline-specific approaches have grown out of calls since the 1990s for evidence-based practice and systematic reviews⁹. Similarly, a causal empiricist approach has become the dominant paradigm in development economics¹⁰. While calls for evidence-based practice have also encouraged the growth of systematic reviews in environmental management and conservation^{11,12}, broader consensus on an approach to evaluating and determining the strength and appropriateness of different types of evidence remains elusive.

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 cross-disciplinary understanding of

social–environmental systems embodied in the Sustainable Development Goals. Using a medical standard to assess evidence on environment outcomes, for example, would mean excluding much candidate evidence (for instance studies that are observational rather than experimental), thus reducing the ability to discriminate the relative strength of evidence supporting different interventions and likely missing key insights about whether and where an intervention will lead to the desired change¹³. Alternatively, assessing evidence in a manner that employs a different understanding of validity than medicine can be a barrier to collaboration across disciplines where health outcomes are concerned. Even within these disciplines that have a broadly comparable interpretation of evidence (for example, there is some objective truth out there about whether an intervention causes an outcome), commensurability is not easy to achieve. Nevertheless, this is the point. Increasingly we are faced with evaluating interventions that do not fall neatly into a single disciplinary paradigm. To do so, we require an approach to evidence assessment that can combine evidence from different disciplines.

Table 1 describes six different types of candidate evidence commonly available within the disciplines of health, development and the environment. There is no inherent rank order in terms of the strength of evidence these six types produce to support a hypothesis. Some forms of evidence are more appropriate for particular questions than others, for example, understanding whether an intervention will lead to a behaviour change may be better informed by qualitative evidence that focuses on context and perceptions around the behaviour, than by a quantitative study¹⁴. However, within the disciplines that work with each type of evidence, there is general agreement that some modes of information

Table 1 | Types of candidate evidence to assess causal associations

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 (or cases) under investigation, without characterizing 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. 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, indigenous 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 measurements that may or may not be part of study, for example, meteorological records.

The typology characterizes underlying sources and important differences between the types of candidate evidence commonly used, both explicitly and implicitly, in decisions about interventions in social-environmental 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. This set of evidence types reflects only forms of knowledge consistent with the view of evidence that is the focus of this manuscript.

collection or generation are more precise or reliable than others, and that some forms of analysis are less biased than others. These ‘methodological standards’ for each candidate type of evidence can be used to determine the quality of evidence provided¹⁵.

Evidence principles

We suggest that a reasonable and practicable approach to evidence assessment¹⁶, which recognizes and integrates different types of evidence, is possible. We propose that four characteristics of evidence represent foundational considerations when assessing a body of evidence for a specific question about causal associations (Fig. 1). Here, we briefly introduce each of these principles:

- The variety of evidence types that support an association between a phenomenon or intervention and observed outcomes
- The consistency of the effect found in the evidence about the causal association
- The credibility of the evidence sources being integrated or considered
- The applicability of the evidence to the question of interest

Multiple types of evidence. Each type of evidence described in Table 1 has its own strengths and weaknesses. Where the answer to a question of association 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 a reported relationship 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. Evidence assessment schemes commonly recognize that multiple, unrelated lines of evidence provide stronger overall evidence¹⁷, but other schemes do not explicitly 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 increases confidence in the answer to the question in view. Consistency implicitly places value on having a larger amount of evidence but also has multiple dimensions to it. Consistency can be considered in the direction (or sign) of an association, or for quantitative evidence, the size and the range or variance of an effect. Consistency of effect across studies is considered a central tenet of many evidence assessment schemes. This is particularly true of schemes looking to assess general claims of relationships between a treatment and outcome for medical interventions⁷ or exposure and hazard for public health questions¹⁸. Although 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 good explanations, often revealed through qualitative studies, for variation resulting from the basis for comparison. We intentionally use consistency rather than size or magnitude as a term because it is inclusive of evaluation by evidence types such as qualitative studies that do not generally involve magnitude estimations.

Credible sources. Where candidate evidence is available from sources widely seen as credible (that is, trusted and believed in), as judged by the prevailing standards for that type of candidate evidence, 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 bias on findings is minimized¹⁹. 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-review system, but one that defies easy characterization. Candidate evidence that comes from sources outside peer review can be credible, (for example, reports from UN organizations). However, where most candidate evidence comes from sources with the potential for perceived bias (for example, because of an organization’s agenda or funding), a thorough assessment of study designs 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 (for instance, similarity in the populations, interventions, and outcomes being considered), it is reasonable to assume that this evidence is relevant and therefore has greater potential for providing strong support. How applicable candidate evidence is to the question at hand is dependent on both the context for the evidence (for example, 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 implementation of the treatment, the measurement of the outcome and the basis for comparison or counterfactual (for example, was the

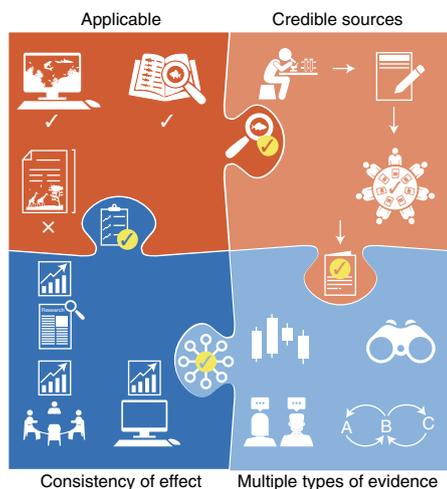


Fig. 1 | The four principles that underpin a cross-disciplinary approach to evidence. Where all four principles are met, the body of candidate evidence would provide strong support for a hypothesized causal association.

same outcome of interest measured or was a related outcome measured). The applicability principle is akin to the assessment of evidence directness found in other evidence grading schemes⁷, but with greater flexibility as to what constitutes fit with the question and context in focus.

The body of candidate evidence would provide strong support for a hypothesized causal association where (i) support comes from multiple types of evidence, (ii) there is consistency in the pattern of association, (iii) the evidence comes from credible sources and (iv) the evidence is highly applicable to the question of interest. Across previously published evidence schemes, a large number of factors or criteria have been proposed as relevant to an assessment of evidence strength (for example, whether a dose-response relationship can be defined, or whether the design of individual studies includes randomization). However, we (authors from a range of disciplines) consider the four characteristics described above as unequivocal indicators of the strength of evidence across disciplines.

In making decisions about interventions, it is important to know where there is strong support for a hypothesis, and if there is not strong support, in what ways it is not strong (that is, which of the four evidence principles is not satisfied). Where the body of evidence is inconsistent with these

principles, policymakers and practitioners should be aware of that lack of certainty in the cross-disciplinary knowledge base. That does not preclude greater certainty based on one particular type of evidence, as judged by the standards of the relevant discipline and by its relevance to the question at hand. We have focused on evidence in the context of evaluating interventions but there is need for evidence from a cross-disciplinary knowledge base throughout a policy process²⁰.

Assessing a body of evidence against each of the principles requires individual judgement and will be implemented variably by different people. This element of subjective judgment is a near universal feature of evidence assessment schemes. Our goal is to provide some foundational principles of 'strong evidence' that will facilitate understanding and resolution of these differences in judgment in the cross-disciplinary teams that are required to assess candidate evidence on interventions operating across the domains of development, environment and health. These principles intentionally move away from the hierarchical view of evidence types prevalent in evidence assessment schemes. Such a shift is necessary if we are to effectively confront social–environmental sustainability challenges with evidence. Engaging in the full complexity of social–environmental systems, however, challenges us to think even more broadly about what counts as knowledge and knowing. □

Edward T. Game^{1,2*}, Heather Tallis³, Lydia Olander⁴, Steven M. Alexander^{5,6,7}, Jonah Busch⁸, Nancy Cartwright^{9,10}, Elizabeth L. Kalies¹¹, Yuta J. Masuda¹², Anne-Christine Mupepele^{13,14}, Jiangxiao Qiu¹⁵, Andrew Rooney¹⁶, Erin Sills¹⁷ and William J. Sutherland¹⁸

¹The Nature Conservancy, South Brisbane, Queensland, Australia. ²University of Queensland, St Lucia, Queensland, Australia. ³The Nature Conservancy, Santa Cruz, CA, USA. ⁴Nicholas Institute for Environmental Policy Solutions, Duke University, Durham, NC, USA. ⁵National Socio-Environmental Synthesis Center, University of Maryland, Annapolis, MD, USA. ⁶Stockholm Resilience Centre, Stockholm University, Stockholm, Sweden. ⁷Environmental Change and Governance Group, Faculty of Environment, University of Waterloo, Waterloo, Canada. ⁸Center for Global Development, Washington DC, USA. ⁹Philosophy Department, Durham University, Durham

DH1 3HN, UK. ¹⁰University of California San Diego, La Jolla, CA, USA. ¹¹The Nature Conservancy, Durham, NC, USA. ¹²The Nature Conservancy, Seattle, WA, USA. ¹³Department of Biometry and Environmental System Analysis, University of Freiburg, Freiburg, Germany. ¹⁴Nature Conservation and Landscape Ecology, University of Freiburg, Tennenbacher Str. 4, 79106 Freiburg, Germany. ¹⁵School of Forest Resources and Conservation, Fort Lauderdale Research and Education Center, University of Florida, Florida, FL, USA. ¹⁶Division of the National Toxicology Program, National Institute of Environmental Health Sciences, Research Triangle Park, NC, USA. ¹⁷Department of Forestry and Environmental Resources, NC State University, Raleigh, NC, USA. ¹⁸Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge, UK.

*e-mail: egame@tnc.org

Published online: 14 September 2018

<https://doi.org/10.1038/s41893-018-0141-x>

References

1. Midgley, G., Nicholson, J. D. & Brennan, R. *Ind. Market. Manag.* **62**, 150–159 (2017).
2. Fazey, I. et al. *Energy Res. Social Sci.* **40**, 54–70 (2018).
3. Cornell, S. et al. *Environ. Sci. Pol.* **28**, 60–70 (2013).
4. Tengö, M. et al. *Curr. Opin. Environ. Sustain.* **26**, 17–25 (2017).
5. Cartwright, N. *Philos. Sci.* **73**, 981–990 (2006).
6. Munafò, M. R. & Davey Smith, G. *Nature* **553**, 399–401 (2018).
7. Guyatt, G. H. et al. *Brit. Med. J.* **336**, 924 (2008).
8. Morgan, R. L. et al. *Environ. Int.* **92**, 611–616 (2016).
9. Sackett, D. L. et al. *Brit. Med. J.* **312**, 71 (1996).
10. Panhans, M. T. & Singleton, J. D. *Hist. Polit. Econ.* **49**, 127–157 (2017).
11. Sutherland, W. J. et al. *Trends Ecol. Evol.* **19**, 305–308 (2004).
12. *Collaboration for Environmental Evidence, Guidelines for Systematic Review and Evidence Synthesis in Environmental Management Version 4.2* (Collaboration for Environmental Evidence, 2013).
13. Voß, J.-P. et al. *J. Environ. Pol. Plan.* **9**, 193–212 (2007).
14. Bennett, N. J. *Conserv. Biol.* **30**, 582–592 (2016).
15. Khagram, S. & Thomas, C. W. *Public Admin. Rev.* **70**, S100–S106 (2010).
16. Cartwright, N., Goldfinch, A. & Howick, J. J. *Child. Serv.* **4**, 6–14 (2010).
17. Norris, R. et al. *Freshw. Sci.* **31**, 5–21 (2011).
18. Rooney, A. A. et al. *Environ. Health Persp.* **122**, 711 (2014).
19. Montbeller, G. & Von Winterfeldt, D. *Risk Anal.* **35**, 1230–1251 (2015).
20. Clark, T. W. *The Policy Process: A Practical Guide for Natural Resources Professionals*. (Yale Univ. Press, New Haven, CT, 2002).

Acknowledgements

We thank I. Fazey for extensive input that thoroughly improved the manuscript and Z. Burivalova for input on the design of Fig. 1. This collaboration was supported by a grant from the David and Lucile Packard Foundation to H.T., E.T.G. and L.O. S.M.A. acknowledges support from the Social Sciences and Humanities Research Council of Canada and the National Socio-Environmental Synthesis Center through NSF grant no. DBI-1052875. W.J.S. is funded by Arcadia.