

DISCUSSION

Open Access

How to increase the potential policy impact of environmental science research

Gary S Bilotta^{1*†}, Alice M Milner^{2,3†} and Ian L Boyd^{2,4†}**Abstract**

This article highlights eight common issues that limit the policy impact of environmental science research. The article also discusses what environmental scientists can do to resolve these issues, including (1) optimising the directness of their study so that it examines similar processes/populations/environments/ecosystems to that of policy interest; (2) using the most powerful study design possible, to increase confidence in the identified causal mechanisms; (3) selecting a sufficient sample size, to reduce the chance of false positives/negatives and increase policy-makers' confidence in extrapolation of the findings; (4) minimizing the risk of bias through randomization of study units to treatment and control groups (reducing the risk of selection bias), blinding of study units and investigators (reducing the risk of performance and detection bias), following-up study units from enrolment to study completion (reducing the risk of attrition bias) and prospectively registering the study on a publically-available platform (reducing the risk of reporting and publication bias); (5) proving that statistical analyses meet test assumptions by reporting the results of statistical assumption checks, ideally publishing full datasets online in an open-access format; (6) publishing the research whether statistically significant or not, policy-makers are just as interested in the negative or insignificant results as they are in the positive results; (7) making the study easy to find and use, the title and abstract of an article are of high importance in determining whether articles are examined in detail or not and used to inform policy; (8) contributing towards systematic reviews on environmental topics, to provide policy-makers with comprehensive, reproducible and updateable syntheses of all the evidence on a given topic.

Keywords: Policy; Environmental science; Study design; Bias; Statistical power; Randomization; Statistical validity; Systematic review

Background

Evidence from environmental science is used to inform public policies but those policies sometimes deviate in significant ways from what the evidence may seemingly support. This can be a source of frustration for some environmental scientists. Frequently, deviation of policy from scientific evidence occurs because policy implementation is multidimensional and includes electoral, ethical, cultural, practical, legal and economic considerations [1]. Occasionally, deviation of policy from apparent scientific evidence occurs because of problems surrounding the quality or reporting of the scientific evidence itself. The following eight sections of this article identify common issues associated with evidence from environmental science

and discuss what environmental scientists can do to resolve these issues to increase the potential policy impact of their research.

Optimise the directness of the study

Policy-makers can make more use of evidence from studies that examine similar processes/populations/environments/ecosystems to that of policy interest, including consideration of the appropriateness of the temporal and spatial scales of observations. For example, one of the criticisms of the scientific evidence of the impacts of neonicotinoid insecticides on insect pollinators centres on the use of laboratory conditions to simulate exposures in the wild. In their article 'A restatement of the natural science evidence base concerning neonicotinoid insecticides and insect pollinators', Godfray et al. [2] state that 'the strengths of laboratory studies are that they allow carefully controlled experiments to be performed on individual insects subjected to

* Correspondence: g.s.bilotta@brighton.ac.uk

†Equal contributors

¹School of Environment and Technology, University of Brighton, Brighton BN2 4GJ, UK

Full list of author information is available at the end of the article

well-defined exposure. The weaknesses are that they are conducted under very artificial conditions (which may affect tolerance to external stress), any avoidance response by the insect is limited and hence the exposure dose and form is determined solely by the experimenter, and responses at the colony or population level are both difficult to study and to extrapolate to the field' [2]. The directness of a study is something that depends on the purpose for which the study is to be used, and this may only become fully apparent after a study is published. However, optimising directness of a study to policy questions can often be considered from the onset of study design (in balance with the degree of experimental control). Scientists can aid the decision by policy-makers, who are considering how similar a study is to the situation of policy interest, by reporting as much background information as possible on the study units and the conditions of the study.

Use the most powerful study design possible

Study design underlies how much confidence policy-makers will have in the findings. Non-randomised observational studies that lack control groups and simply report correlations between variables will typically attract less confidence than a randomised controlled study on the same topic. This is because policy-makers recognise that correlation between observations does not signify causation. Tyler Vigen has created a website called Spurious Correlations, which demonstrates this point in a number of amusing ways. Vigen trawls data sets and matches parameters until he comes up with a

correlation. In the example shown in Figure 1, Vigen presents the correlation (0.99) between 'US spending on science, space, and technology' and 'Suicides by hanging, strangulation and suffocation'.

The more data are trawled for patterns, the more likely it is that the patterns found will simply reflect chance associations. This might be innocuous as long as we are comparing clearly unrelated variables, such as those shown in Figure 1. But if environmental scientists find a chance correlation between two variables that just happen to have a plausible functional cause and effect relationship, then there is a higher risk of misinterpretation. Where there is a choice, this is why policy-makers will often place more confidence in evidence from scientific studies that have both control groups and treatment groups, and where the individual study units have been allocated to these different groups based on some random allocation process that is not possible to predict. Without a control group (i.e. study units that are dealt with in exactly the same way as study units in the experimental group except for the treatment applied), it is difficult to determine whether a given treatment really had an effect or whether, for example, there was a natural change over time in the outcome of interest that may be unconnected with the treatment. Without random allocation of study units to control and treatment groups, it is difficult to ensure that the groups are balanced at baseline with respect to known and unknown determinants of outcome, and therefore, it is difficult to ascertain whether variances in outcomes were caused by the treatment/intervention of interest. An example of an

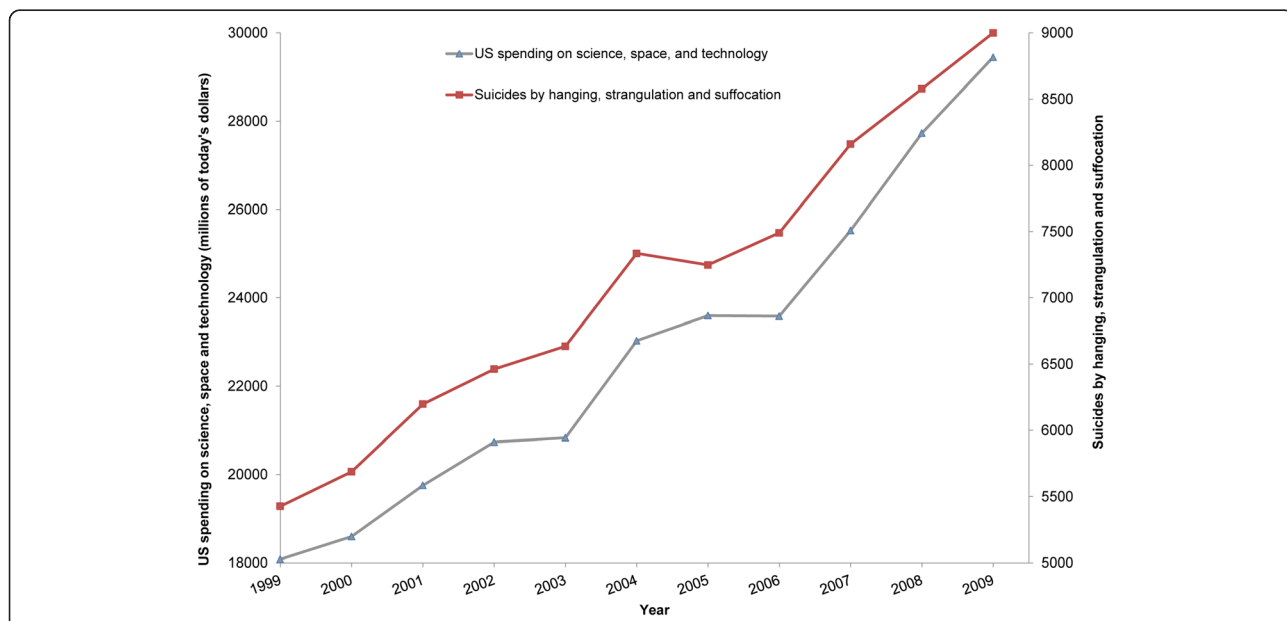


Figure 1 The spurious correlation. Correlation (0.99) between 'US spending on science, space, and technology' in *Millions of today's dollars* (US OMB) and 'Suicides by hanging, strangulation and suffocation' (US) (CDC). Reproduced from <http://www.tylervigen.com/>.

environmental study affected by the lack of random allocation of study units to control and treatment groups is provided by Peach et al. [3], who attempted to assess the effect of Countryside Stewardship Schemes (CSS) on populations of circl buntings (*Emberiza circlus*). This study surveyed the entire geographic range of the species, between 1992 and 1998, and compared changes in the abundance of circl buntings in tetrads (2 × 2 km squares) over time. There was some evidence, however, that the selection of sites to be managed under the CSS (though not carried out by the investigators) may not have been random and may have been related to the outcome of interest. The authors acknowledge that ‘the relatively high densities of circl buntings in 1992 on land that subsequently entered CSS agreements probably reflects a tendency for sites already supporting circl buntings to be more likely to apply for, and be offered, CSS status. Many of the CSS agreements...include Sites of Special Scientific Interest or County Wildlife Sites, and landowners of sites known to support circl buntings have been encouraged to apply for CSS status’ [3].

Select a sufficient sample size

The sample size and the sampling strategy (e.g. random, systematic, stratified) of a study are important determinants of how representative the study findings will be of the wider population (biotic or abiotic). The sample size of a study is also an important determinant of the validity of the statistical conclusions, and it is therefore of critical interest to policy-makers [4,5]. The adequacy of the sample size, or the ‘sample power’, to detect an effect or difference can often be estimated a priori using statistical power analyses or alternative Bayesian approaches [6]. The lower the power, the less likely it is that the study will detect an effect that exists and thus the more likely that it will falsely accept a null hypothesis. A statistical power of 0.8 means that for ten true hypotheses tested, two will be ruled out because their effects are not detected in the data. However, consideration and reporting of statistical power is rare in environmental science studies [4-6]. For example, a review of fisheries and aquatic science research papers [7] that did not reject some null hypothesis found 98% of the papers failed to report statistical power. To increase the potential policy impact of environmental studies, researchers should carry out a priori statistical power analyses (where this is possible) and should report the results of this in any publications arising from the work. If researchers are unable to achieve an acceptable statistical power through increasing their sample size, reducing their measurement errors and/or increasing their limits of acceptable change, then they should consider whether the study is worth conducting [8].

Minimise bias within the study

A bias is a systematic error resulting from poor study design or issues associated with conduct in the collection, analysis, interpretation and reporting of data. Biases can operate in either direction, causing an under- or an overestimation of effect, which if unaccounted for may ultimately affect the validity of the conclusions from a study [9]. Biases in scientific research can be cryptic [1], and it is usually impossible to know the extent to which they have affected the results of a particular study. However, Gluud [10] highlights, from investigating variations in the results of studies of the same intervention according to features of their study design, that there is empirical evidence that the following key aspects of study design help to minimize the risk of bias:

- *Randomization* minimizes the risk of *selection bias* (systematic differences between baseline characteristics of the groups that are to be compared).
- *Blinding* of study units and investigators minimizes the risk of *performance bias* (systematic differences between groups in the care that is provided or in exposure to factors other than the interventions of interest due to lack of blinding of investigators) and *detection bias* (systematic differences between groups in how outcomes are determined) due to participants’ or investigators’ expectations.
- *Follow-up* of study units from enrolment to study completion minimizes the risk of *attrition bias* (systematic differences between groups in withdrawals from a study/loss of samples).
- *Prospective study registration* (including a description of the number of study units, treatment protocols, duration of the study, primary outcomes to be measured and planned analyses) and *unselective reporting* of outcomes minimizes the risk of *reporting bias* (systematic differences between reported and unreported findings). Studies that are registered prospectively are also far less likely to be derailed by chance correlations. Various mechanisms for prospective scientific study registration already exist or are in development (e.g. the Open Science Framework).

Many of these aspects of study design are not technically difficult to implement and would not necessarily add significantly to the expense of a study, but they can significantly increase policy-makers’ confidence in the findings of a study.

Prove that statistical analyses meet test assumptions

Most statistical tests are based on a set of assumptions about the data that must be met prior to the application of the statistical analysis and testing of a hypothesis. For

example, virtually all parametric statistics have an assumption that the data come from a population that follow a certain distribution. Other assumptions include homoscedasticity (data from multiple groups have the same variance), linearity (data have a linear relationship) and independence (data are independent). Violating assumptions of a statistical test may make tests more or less likely to make type I or II errors (false positives and false negatives), which can lead to incorrect inferences about the cause-effect relationship and thus undermine meaningful research. Statistical procedures can and should be used to check that the statistical analyses meet the assumptions of the statistical test, and the results from these procedures need to be reported to verify that the statistical analyses are valid. Ideally, researchers should publish full datasets online with a digital object identifier (DOI) to persistently identify the dataset from which conclusions were drawn and enable subsequent researchers to use, interrogate, and test the data. There are multiple repositories for making data more widely available, such as Dryad - a curated general-purpose repository that makes the data underlying scientific publications discoverable, freely reusable and citable; Figshare - a repository that offers a means for sharing data and other research materials; and the Open Science Framework, which offers infrastructure for documenting, archiving and sharing data within collaborative teams and making research materials publicly available. Academic journals are also increasingly adopting policies for making data, protocols and analytical codes available. Publishing data enables easy verification of statistical conclusions and will facilitate inclusion of the data in meta-analyses by policy-makers.

Publish the research whether statistically-significant or not
Researchers, and sometimes the journals that publish research, are more likely to publish positive results (e.g. showing a statistically significant finding) than results that are negative (i.e. supporting the null hypothesis) or insignificant. This phenomenon is often referred to as publication bias. Song et al. [11] found that the principle reasons for non-publication of completed studies included lack of time or low priority (35%), unimportant results (20%) and journal rejection (10%); indicating that publishing bias primarily originates from researchers failing to write up and submit to journals when the results are negative or non-significant, and rather concentrated on 'wonderful results'. In reality, policy-makers are just as interested in the negative or insignificant results as they are in the positive results. In fact, as explained in an excellent animation in the Economist [12], the negative results can be much more trustworthy. Policy-makers need to know both positive and negative findings in order to make well-informed decisions, so scientists

should publish results whether they show a statistically-significant positive effect or not.

Make the study easy to find and use

Policy-makers do not have subscriptions to every environmental science journal that exists, neither do they have unlimited resources to search for every article related to a given topic. As a consequence, the title and abstract of an article are of high importance in determining whether articles are examined in detail or not and used to inform policy. Titles that are informative, concise and include keywords that identify the article's main concepts, variables and relationships between them increase the chance of policy-makers finding and using research. Structured abstracts have been found to contain more information, are more easily searched and help readers, including policy-makers, to find information more quickly compared to traditional abstracts [13]. There are publication guidance documents and checklists for standardised reporting of different types of studies (e.g. CONSORT for randomised controlled trials, STROBE for observational studies, PRISMA for systematic reviews and meta-analyses) and an updated list of reporting guidelines is maintained by the EQUATOR Network [14]. In addition, publishing research in an open access format or publishing pre-proof versions of the article online in accordance with publishers' rules increases the chance of research findings being accessed and used by policy-makers who do not have access to all academic journals.

Contribute to systematic reviews

There are alternative methods of influencing public policy other than conducting and reporting primary research: contributing to synthesising multiple sources of primary evidence is a good example of one of these alternative means. The scientific evidence-base on many environmental topics is large and continually growing [15]. Systematic reviews, such as those conducted by the Collaboration for Environmental Evidence (CEE; an open community of scientists and managers who, from their initial centres in Australia, Canada, South Africa, Sweden and the UK, prepare systematic reviews on environmental topics) can be extremely useful to policy-makers; providing a comprehensive, objective, reproducible and updateable synthesis of all the evidence on a given topic [16]. Policy-makers prefer systematic reviews to traditional narrative literature reviews as it is acknowledged that narrative literature reviews are more vulnerable to author bias which can occur when the review authors intentionally or unintentionally select or emphasise research according to their own opinions, prejudices or commercial interests. Furthermore, narrative literature reviews rarely consider, in a reproducible and meaningful manner, the methodological quality, degree of bias and therefore reliability

of the primary studies that are cited. These features of narrative literature reviews are more likely to lead to ill-informed environmental policies.

To date, the CEE has published more than 60 systematic reviews, with a further 30 in progress. These systematic reviews, which are all available from the CEE Library, cover a range of topics including pure environmental science questions such as ‘What is the evidence for glacial shrinkage across the Himalayas?’ [17], applied environmental management topics such as ‘Evaluating the biological effectiveness of fully and partially protected marine areas’ [18] and human - environment interaction questions such as ‘What is the evidence that scarcity and shocks in freshwater resources causes conflict instead of collaboration?’ [19]. Through contributing to further systematic reviews, it may be possible in the future for policy-makers to visit the CEE and other systematic review libraries for comprehensive syntheses on many topics relevant to their policy questions.

Conclusions

This article highlights eight common issues surrounding the quality or reporting of environmental science research that can limit its potential policy impact. The article also discusses what environmental scientists can do to resolve these issues to increase their potential impact on public policy. The recommendations made arise from the development of a best-practice approach to quality assessment of evidence from environmental science [20]. This quality assessment tool, known as the Environmental-GRADE tool, is adapted from a best-practice tool developed by the healthcare sector and used by the World Health Organization, the UK National Institute for Health and Care Excellence (NICE), and more than 20 health care bodies internationally [9].

The Environmental-GRADE tool describes four levels of evidence quality (high, moderate, low and very low). The highest quality rating is initially for randomised controlled trials, the low quality rating is generally for sound observational studies and the very low quality rating includes, but is not limited to, studies with critical problems and unsystematic observations (e.g. case studies). Assessors can, however, downgrade evidence to moderate, low or even very low quality evidence, depending on the presence of the three factors: (1) The risk of bias within the study - assessed using the Environmental-Risk of Bias Tool which was adapted from the Cochrane Collaboration’s Risk of Bias tool [21]. (2) The directness of the study - assessed using Environmental-GRADE tool criteria. (3) The precision of the effect estimates^a - assessed using Environmental-GRADE tool criteria. Observational studies can be upgraded to moderate or high quality if such studies yield large effects and there is no obvious bias explaining those effects; all plausible confounding factors

would reduce a demonstrated effect or suggest a spurious effect when results show no effect and/or if there is evidence of a dose-response gradient.

It is hoped that wider awareness of the quality assessment criteria used in Environmental-GRADE, and an understanding of the justification for these criteria, as highlighted in this article, will contribute to improved study design and reporting of environmental science research that will increase its potential policy impact. It will however remain the case, even when high quality scientific evidence is available, that public policy will be informed by a number of different evidence sources, including electoral, ethical, cultural, practical, legal and economic considerations. Nevertheless, policy-makers should publicly explain the reasons for policy decisions, particularly when the decision is not consistent with scientific advice, and in doing so should accurately represent the evidence.

Endnote

^aIn this case, imprecision refers to *random error*, meaning that multiple replications of the same study would produce different effect estimates.

Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

All authors contributed equally to the writing of this article. All authors read and approved the final manuscript.

Acknowledgements

This work arises from Natural Environment Research Council funding (Grant NE/L00836X/1 and NE/L008599/2) in collaboration with the United Kingdom’s Department for Environment, Food and Rural Affairs.

Author details

¹School of Environment and Technology, University of Brighton, Brighton BN2 4GJ, UK. ²Department for Environment, Food and Rural Affairs, London SW1P 3JR, UK. ³Department of Geography, Royal Holloway, University of London, Egham, Surrey TW20 0EX, UK. ⁴College Gate, University of St Andrews, St Andrews, Fife KY16 9AJ, UK.

Received: 19 October 2014 Accepted: 25 February 2015

Published online: 29 March 2015

References

- Boyd I. Research: a standard for policy-relevant science. *Nature*. 2013;501:159–60.
- Godfray HCJ, Blacquière T, Field LM, Hails RS, Petrokofsky G, Potts SG, et al. A restatement of the natural science evidence base concerning neonicotinoid insecticides and insect pollinators. *Proc R Soc B Sci*. 2014;281(1786):20140558.
- Peach WJ, Lovett LJ, Wotton SR, Jeffs C. Countryside stewardship delivers ciril buntings (*Emberiza cirilus*) in Devon, UK. *Biol Conserv*. 2001;101:361–73.
- Peterman RM. The importance of reporting statistical power: the forest decline and acidic deposition example. *Ecology*. 1990;71:2024–7.
- Reynolds JH, Thompson WL, Russell B. Planning for success: identifying effective and efficient survey designs for monitoring. *Biol Conserv*. 2011;144(5):1278–84.
- Legg CJ, Nagy L. Why most conservation monitoring is, but need not be, a waste of time. *J Environ Manag*. 2006;78(2):194–9.
- Peterman RM. Statistical power analysis can improve fisheries research and management. *Can J Fish Aquat Sci*. 1990;47(1):2–15.

8. Manly BF. The design and analysis of research studies. Cambridge: Cambridge University Press; 1992.
9. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *Br Med J*. 2011;343:d5928.
10. Gluud LL. Bias in clinical intervention research. *Am J Epidemiol*. 2006;163(6):493–501.
11. Song F, Parekh S, Hooper L, Loke YK, Ryder J, Sutton AJ, et al. Dissemination and publication of research findings: an updated review of related biases. *Health Technol Assess*. 2010;14(8):iii, ix–xi.
12. Economist. Unlikely results: why most published scientific research is probably false. 2014. <http://www.economist.com/node/21587349>. Accessed 17th August.
13. Hartley J, Sydes M, Blurton A. Obtaining information accurately and quickly: are structured abstracts more efficient? *J Inf Sci*. 1996;22(5):349–56.
14. Simeria I, Moher D, Hirst A, Hoey J, Schulz KF, Altman DG. Transparent and accurate reporting increases reliability, utility, and impact of your research: reporting guidelines and the EQUATOR Network. *BMC Med*. 2010;8(1):24.
15. Larsen PO, von Ins M. The rate of growth in scientific publication and the decline in coverage provided by Science Citation Index. *Scientometrics*. 2010;84(3):575–603.
16. Bilotta GS, Milner AM, Boyd IL. On the use of systematic reviews to inform environmental policies. *Environ Sci Pol*. 2014;42:67–77.
17. Miller J, Rees G, Warnaars T, Young G, Shrestha A, Collins D. What is the evidence about glacier melt across the Himalayas? *CEE Review* 2012, 10–008. Collaboration for Environmental Evidence: www.environmentalevidence.org/SR10008.html.
18. Sciberras M, Jenkins S, Kaiser M, Hawkins S, Pullin A. Evaluating the biological effectiveness of fully and partially protected marine areas. *Environ. Evid*. 2013; 2 (4). In: <http://www.environmentalevidencejournal.org/content/2/1/4>.
19. Johnson V, Fitzpatrick I, Floyd R, Simms A. What is the evidence that scarcity and shocks in freshwater resources cause conflict instead of promoting collaboration? *CEE Review* 2011, 10–010. Collaboration for Environmental Evidence. In: www.environmentalevidence.org/SR10010.html.
20. Bilotta GS, Milner AM, Boyd IL. Quality assessment tools for evidence from environmental science. *Environ Evid*. 2014, 3: (14). <http://www.environmentalevidencejournal.org/content/3/1/14>.
21. Turner L, Boutron I, Hróbjartsson A, Altman DG, Moher D. The evolution of assessing bias in Cochrane systematic reviews of interventions: celebrating methodological contributions of the Cochrane Collaboration. *Syst Rev*. 2013;2(1):79.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► springeropen.com
