



TRANSFORMING CONSERVATION

A Practical Guide to Evidence
and Decision Making

EDITED BY WILLIAM J. SUTHERLAND



<https://www.openbookpublishers.com>



©2022 William J. Sutherland. Copyright of individual chapters is maintained by the chapter's authors.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). This license allows you to share, copy, distribute and transmit the text; to adapt the text for non-commercial purposes of the text providing attribution is made to the authors (but not in any way that suggests that they endorse you or your use of the work). Attribution should include the following information:

William J. Sutherland, *Transforming Conservation: A Practical Guide to Evidence and Decision Making*. Cambridge, UK: Open Book Publishers, 2022, <https://doi.org/10.11647/OBP.0321>

Further details about Creative Commons licenses are available at <https://creativecommons.org/licenses>

All external links were active at the time of publication unless otherwise stated and have been archived via the Internet Archive Wayback Machine at <https://archive.org/web>

Updated digital material and resources associated with this volume are available at <https://doi.org/10.11647/OBP.0321#resources>

Every effort has been made to identify and contact copyright holders and any omission or error will be corrected if notification is made to the publisher.

ISBN Paperback: 978-1-80064-856-2

ISBN Hardback: 978-1-80064-857-9

ISBN Digital (PDF): 978-1-80064-858-6

ISBN Digital ebook (EPUB): 978-1-80064-859-3

ISBN Digital ebook (AZW3): 978-1-80064-860-9

ISBN XML: 978-1-80064-861-6

ISBN HTML: 978-1-80064-862-3

DOI: 10.11647/OBP.0321

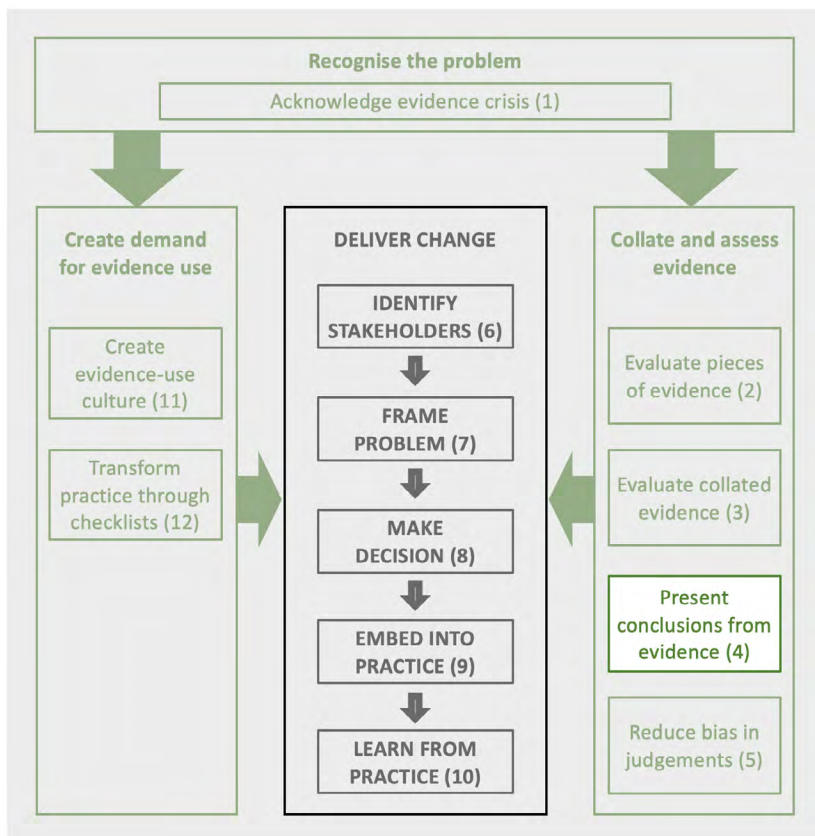
Cover image: Tiger Mountain / Issaquah, WA by Dave Hoefler.

Design by Katy Saunders

4. Presenting Conclusions from Assessed Evidence

William J. Sutherland^{1,2}, Tatsuya Amano^{3,4}, Nicolas Boenisch⁵, Samantha H. Cheng^{6,7}, Alec P. Christie^{1,2,8}, H. Charles J. Godfray⁹, Matthew J. Grainger¹⁰, Angela R. McLean⁹, William H. Morgan¹, Nancy Ockendon¹¹, Nicola P. Randall¹², Ullrika Sahlin¹³, Thomas B. White¹

Applying evidence builds on the conclusions of the assessment of the evidence. The aim of the chapter is to describe a range of ways of summarising and visualising different types of evidence so that it can be used in various decision-making processes. Evidence can also be presented as part of evidence capture sheets, argument maps, mind maps, theories of change, Bayesian networks or evidence restatements.



¹ Conservation Science Group, Department of Zoology, University of Cambridge, The David Attenborough Building, Pembroke Street, Cambridge, UK.

² Biosecurity Research Initiative at St Catharine's College (BioRISC), St Catharine's College, Cambridge, UK.

³ School of Biological Sciences, University of Queensland, Brisbane, 4072 Queensland, Australia

⁴ Centre for Biodiversity and Conservation Science, The University of Queensland, Brisbane, 4072 Queensland, Australia

⁵ Foundations of Success Europe, 7211 AA Eefde, The Netherlands

⁶ Center for Biodiversity and Conservation, American Museum of Natural History, New York, NY 10026, USA

⁷ Global Science, World Wildlife Fund, Washington, D.C., USA

⁸ Downing College, Regent Street, Cambridge, UK

⁹ Oxford Martin School, University of Oxford, Oxford, UK

¹⁰ Norwegian Institute for Nature Research (NINA), PO Box 5685 Torgarden, 7485 Trondheim, Norway

¹¹ Endangered Landscapes Programme, Cambridge Conservation Initiative, David Attenborough Building, Pembroke Street, Cambridge, UK

¹² Centre for Evidence-based Agriculture, Harper Adams University, Newport, Shropshire, UK

¹³ Centre for Environmental and Climate Science, Lund University, Lund, Sweden

Contents

4.1 Principles for Presenting Evidence

4.2 Describing Evidence Searches

4.3 Presenting Different Types of Evidence

4.4 Presenting Evidence Quality

4.5 Balancing Evidence of Varying Strength

4.6 Visualising the Balance of Evidence

4.7 Synthesising Multiple Evidence Sources

References

What does the evidence say? This is a common question and one that is critical to the process of decision-making (Chapter 8) and applying evidence, whether writing plans or deciding upon policies (Chapter 9). The aim of this chapter is to present a range of means of presenting evidence that gives a fair summary whilst making clear any limitations or biases.

This chapter considers the general principles, reviews the various elements that are involved in finding and presenting evidence, describes some of the main types of evidence and how these might be presented, and describes some approaches for presenting multiple pieces of evidence.

4.1 Principles for Presenting Evidence

In applying evidence, it is important to understand the conclusions drawn and the likelihood that those conclusions are accurate and relevant. The core general principles for presenting evidence (see Box 4.1) are designed to reduce the risk that the evidence is misapplied due to the conclusions being presented incorrectly; the limitations not being made clear; or because the conditions are very different.

In practice, evidence is presented in a range of formats that differ in the detail provided. In some cases, such as reports and assessments, the key evidence is likely to be well documented with details of the source and assessment of quality. However, in practice, much evidence is provided by experts. In this case, the responsibility is on the expert to ensure that if caveats are needed then they are provided (for example, ‘although the evidence is weak it seems...’, ‘the evidence is conflicting but my interpretation is...’, ‘the only studies are on lizards but it seems likely that snakes would respond the same way’). Thus if a responsible evidence-based expert provides no caveat then it should be reasonable to assume the expert holds no such concerns. In practice, it is worth checking if the expert does this and also checking the details for key assumptions.

Box 4.1 General principles for presenting evidence

1. The means by which the evidence was obtained is clear.
2. The type and source of evidence is clear.
3. The manner in which it is communicated is tailored to the intended audience.
4. The presentation and interpretation of the evidence is, as far as possible, neutral and balanced.
5. Any biases or limitations are made clear.
6. Conflicting evidence is presented.
7. The strength of the evidence behind recommendations is transparent.
8. If an atypical study is presented then this is made clear ('communities vary in their enthusiasm for this project but some are very enthusiastic, for example...').

4.2 Describing Evidence Searches

Understanding how the evidence is presented is key for interpreting the results, as it indicates how complete the evidence is, and the likelihood of any bias.

4.2.1 Describing the types and results of searching

Here, we describe how different types of evidence could be obtained. Table 4.1 outlines possible content for describing a range of search methods.

Table 4.1 Suggested possible content, with examples, for presenting different means of searching for evidence.

Source	Possible content	Examples
Search of actions from Conservation Evidence	Describe search, give date of literature covered (given in the synopsis). Reflect on the actions and the quality, quantity, and relevance of evidence in relation to the topic of interest.	Using the search term <i>roads</i> and refining using <i>reptiles</i> gave 10 actions with at least one study. The action on crossing structures has 16 papers (the eight studies looking at survival include one randomised, controlled, before-and-after study and one review) and that on tunnels has 15 (the three studies examining survival comprise a site comparison studies, a before-and-after study and a replicated test). The reptile synopsis covers literature up to 2018.

Source	Possible content	Examples
Search for reviews from CEEDER	Describe the search giving the date of retrieved reviews. Reflect on the actions and the quality, quantity and relevance of evidence.	Searching for <i>glyphosate</i> on 27.05.2022 gave two reviews. A 2020 systematic review of the literature in English on the effect of applying glyphosate on crop yield in Europe found 67 experimental comparisons. A 2018 narrative review discusses the impact of glyphosate on microorganisms, plants and animals.
Search for studies from Conservation Evidence	Describe the search and types of paper and relevance to the issue of interest.	A 27.05.2022 search for <i>meles</i> gave the following studies relating to the European Badger <i>Meles meles</i> : 9 roads and infrastructure, 2 reintroductions, 3 reducing impact as predators, 2 reducing the impact of TB, 5 reducing the impact on crops.
Systematic literature searches	Describe what was searched, how and when.	A systematic search carried out as part of a paper, broad review, or plan but not published as a systematic review.
Informal search	Describe search and outputs.	A quick search of Google for <i>status of smooth snake in Finland</i> provided a 1981 provisional atlas and a map compiled by IUCN in 2017.

4.2.2 Constraints on generality statements

Evidence collations often provide general conclusions, for example, to cover wide geographic areas or groups of species. However, these are based on the evidence available, which often has taxonomic, habitat and geographic biases (Christie et al., 2020a, 2020b). For example, an action may have been assessed as being beneficial for birds, but all of the studies testing the action may have been carried out in one habitat type for a small number of species. Therefore, before making any decisions it is vital to consider the evidence in more detail to assess its relevance for your study species or system and to state the relevance and any constraints in any outcome, for example, *there are numerous good studies showing this resulted in higher breeding success, but all the studies are on Salmonids.*

4.3 Presenting Different Types of Evidence

The aim of this section is to consider how different types of evidence, including describing costs, may be summarised.

4.3.1 Conservation Evidence summaries

Conservation Evidence collates and summarises individual studies on a topic in a consistent manner to ease comparisons and ensure all the key components are included. For example:

A search of Conservation Evidence for 'tree harvesting' found seven individual studies and seven relevant actions.

Conservation Evidence uses a standard format for describing studies that test interventions, which may be of wider use. The following colour-coded format is here illustrated with an example.

A [TYPE OF STUDY] in [YEARS X-Y] in [HOW MANY SITES] in/of [HABITAT] in [REGION and COUNTRY] [REFERENCE] found that [INTERVENTION] [SUMMARY OF ALL KEY RESULTS] for [SPECIES/HABITAT TYPE]. [DETAILS OF KEY RESULTS, INCLUDING DATA]. In addition, [EXTRA RESULTS, IMPLEMENTATION OPTIONS, CONFLICTING RESULTS]. The [DETAILS OF EXPERIMENTAL DESIGN, INTERVENTION METHODS and KEY DETAILS OF SITE CONTEXT]. Data was collected in [DETAILS OF SAMPLING METHODS].

For the type of study there is a specific set of terms that are used (Table 4.2). For the results, for the sake of brevity, only key results relevant to the effects of the intervention are included. For Site context only those nuances that are essential to the interpretation of the results are included.

As an example:

A replicated, randomized, controlled, before-and-after study in 1993–1999 of five harvested hardwood forests in Virginia, USA (1) found that harvesting trees in groups did not result in higher salamander abundances than clearcutting. Abundance was similar between treatments (group cut: 3; clearcut: 1/30 m²). Abundance was significantly lower compared to unharvested plots (6/30 m²). Species composition differed before and three years after harvest. There were five sites with 2 ha plots with each treatment: group harvesting (2–3 small area group harvests with selective harvesting between), clearcutting and an unharvested control. Salamanders were monitored on 9–15 transects (2 x 15 m)/plot at night in April–October. One or two years of pre-harvest and 1–4 years of post-harvest data were collected.

(1) Knapp S.M., Haas C.A., Harpole D.N. and Kirkpatrick, R.L. 2003. Initial effects of clearcutting and alternative silvicultural practices on terrestrial salamander abundance. *Conservation Biology* 17: 752–62.

Table 4.2 The terms used to describe study designs in Conservation Evidence summaries.

Term	Meaning
Replicated	The intervention was repeated on more than one individual or site. In conservation and ecology, the number of replicates is much smaller than it would be for medical trials (when thousands of individuals are often tested). If the replicates are sites, pragmatism dictates that between five and ten replicates is a reasonable amount of replication, although more would be preferable. The number of replicates is stated when given. Replicates should reflect the number of times an intervention has been independently carried out, from the perspective of the study subject. For example, 10 plots within a mown field might be independent replicates from the perspective of plants with limited dispersal, but not independent replicates for larger motile animals such as birds. In the case of translocations/release of captive bred animals, replicates should be sites, not individuals.
Randomized	The intervention was allocated randomly to individuals or sites. This means that the initial condition of those given the intervention is less likely to bias the outcome.
Paired sites	Sites are considered in pairs, within which one was treated with the intervention and the other was not. Pairs, or blocks, of sites are selected with similar environmental conditions, such as soil type or surrounding landscape. This approach aims to reduce environmental variation and make it easier to detect a true effect of the intervention.
Controlled*	Individuals or sites treated with the intervention are compared with control individuals or sites not treated with the intervention. The treatment is usually allocated by the investigators (randomly or not), such that the treatment or control groups/sites could have received the treatment.
Before-and-after	Monitoring of effects was carried out before and after the intervention was imposed.
Site comparison*	A study that considers the effects of interventions by comparing sites that have historically had different interventions (e.g. intervention vs no intervention) or levels of intervention. Unlike controlled studies, it is not clear how the interventions were allocated to sites (i.e. the investigators did not allocate the treatment to some of the sites).
Review	A conventional review of literature. Generally, these have not used an agreed search protocol or quantitative assessments of the evidence.
Systematic review	A systematic review follows structured, predefined methods to comprehensively collate and synthesise existing evidence. It must weigh or evaluate studies, in some way, according to the strength of evidence they offer (e.g. sample size and rigour of design).
Study	If none of the above applies, for example, a study measuring change over time in only one site and only after an intervention or a study measuring the use of nest boxes at one site.

* Note that 'controlled' is mutually exclusive from 'site comparison'. A comparison cannot be both controlled and a site comparison. However, one study might contain both controlled and site comparison aspects e.g. study of fertilised grassland, compared to unfertilised plots (controlled) and natural, target grassland (site comparison).

4.3.2 Systematic maps

Systematic maps (Section 3.2.1) are overviews of the distribution and nature of the evidence found in literature searches. Standardised descriptive data from individual studies are usually extracted and presented within searchable databases or spreadsheets, which accompany written systematic reports and the descriptive statistics. As critical appraisal of included studies does not usually take place in systematic maps, to avoid vote-counting, systematic maps authors do not usually summarise the outcomes of included studies.

Their description includes statistics about the studies included, such as the number of papers on different sub-topics or from different regions, a summary of the quantity and the quality of the evidence identifying areas of strong evidence and knowledge gaps. It also includes a description of the process used (including whether Collaboration for Environmental Evidence guidelines and ROSES reporting standards are adopted), date carried out, and sources and languages searched. For example:

Adams et al. (2021) mapped out the literature on the effects of artificial light on bird movement and distribution and found 490 relevant papers. The most frequent subjects were transportation (126 studies) and urban/suburban/rural developments (123) but few were from mineral mining or waste management sectors. Many studies are concerned with reducing mortality (169 studies) or deterring birds (88). The review followed a published protocol and adopted the Collaboration for Environmental Evidence guidelines and ROSES reporting standards. The search, on August 21 2019, was for literature in English in the Web of Science Core Collection, the Web of Science Zoological Record, Google Scholar and nine databases.

Or for a brief version:

One good quality, recent systematic map of the English literature (Adams et al., 2021) found 88 papers on using light to deter birds.

4.3.3 Systematic reviews

Systematic reviews use repeatable analytical methods to extract secondary data and analyse it and can also be carried out for qualitative studies. Their description includes the objective, number of studies, statistics and the main results. Description of any meta-analysis should describe the summary or averaged effect size (e.g. the mean difference between having the focal intervention and not having it), the uncertainty around this (e.g. 95% confidence limits) and the heterogeneity (variability in reported effect sizes) between studies included in the analysis and the main results of any meta-regression used to explore the drivers of the between-study heterogeneity. A description of the process used (whether Collaboration for Environmental Evidence guidelines and ROSES reporting standards are adopted), date carried out and sources and languages searched, should also be described.

For example, for the systematic review given in Chapter 3:

An English language systematic review, following Collaboration for Environmental Evidence guidelines, produced 78 primary studies that assess whether tree retention at harvest helps

to mitigate negative impacts on biodiversity, (Fedrowitz et al., 2014). Meta-analysis showed that overall species richness (15 studies) and abundance (50 studies) were significantly (hedges $d < 0.001$) higher with tree retention at harvest compared to clear cutting. However, there were significant differences among taxonomic groups in richness-response to retention cuts compared with un-harvested forests.

Or more briefly:

A good quality meta-analysis of 78 studies showed tree harvest resulted in more biodiversity than clear cutting, but the responses varied between groups.

4.3.4 Dynamic syntheses

Metadataset (see Section 3.7) provides a means for running meta-analysis for the subject of concern. For example for invasive species the user can select the species, countries included action and outcome. It generates a Forest plot with effect sizes and confidence intervals, a paragraph summarising the main conclusions and a funnel plot showing the precision of studies against their effect size results as one check for publication bias. Their description includes the comparison selected and geography included, the response ratio, statistical significance, the number of studies and papers and any significant heterogeneity.

For example, the example shown in Figure 3.4, of analysing the effectiveness of different action on invasive Japanese knotweed could be summarised as follows.

*The effect of chemical and mechanical control on the abundance of invasive Japanese knotweed (*Fallopia japonica*) was selected for analysis. Chemical control reduced abundance by 63% lower compared to the untreated control (response ratio = 0.37; significance $P = 0.0084$; 193 data points from 13 studies in 12 publications). For physical disturbance, the abundance was 16% lower than the control but this was not statistically significant (response ratio = 0.84; $P = 0.579$; 27 data points from 8 studies in 8 publications). For both analyses, there was significant heterogeneity between data points ($P < 0.0001$).*

Or for a brief version

Analysis using metadataset showed a significant 63% reduction in Japanese knotweed abundance when treated with herbicide but a non-significant 16% reduction with physical disturbance.

4.3.5 Conservation Evidence actions

State date, name of action, assessment category and describe the nature of the evidence in relation to the topic of interest. One key principle is to be careful of 'vote counting' — presenting the number of studies showing different results when these might differ in quality.

The action 'Add mixed vegetation to peatland surface' (2018 synopsis) has 18 studies and is classified as Beneficial. Seventeen replicated studies looked at the impact on Sphagnum moss cover (five were also randomised, paired, controlled, before-and-after): in all studies Sphagnum

moss was present (cover ranged from <1 to 73%), after 1–6 growing seasons, in at least some plots. Six of the studies were controlled and found that Sphagnum cover was higher in plots sown with vegetation including Sphagnum than in unsown plots. Five studies reported that Sphagnum cover was very low (<1%) unless plots were mulched after spreading fragments.

Or for a brief version:

The 17 replicated studies on Conservation Evidence looking at adding mixed vegetation showed in all cases there was some Sphagnum in at least some plots but in five cases at low density.

4.3.6 Costs of actions

Most conclusions will wish to give some indication of cost, yet, as described in Chapter 2, the reporting of costs is usually unsatisfactory. Either costs are not given or are given without stating what is included. One solution is the standardised collation and reporting of costs to help think through and report the direct costs of conservation actions. Iacona et al. (2018) developed a framework to help record the direct accounting costs of interventions. These steps are outlined below. Major accounting costs include the costs of consumables, capital expenditure, labour costs and overheads.

- Step 1: State the objective and outcome of the intervention.
- Step 2: Provide context, and the method used for the intervention.
- Step 3: When, where and at what scale was the intervention implemented?
- Step 4: What categories of cost are included? (see Table 2.4) What are these costs and how might they vary with context?
- Step 5: State the currency and date for the reported costs.

White et al. (2022b) created a further 6-step framework for thinking through the wider economic costs and benefits of different conservation actions that can be used to assess cost-effectiveness:

- Step 1: Define the intervention/programme.
- Step 2: Outline the costing perspective and reporting level.
- Step 3: Define the alternative scenario.
- Step 3: Define the types of cost and benefits included or excluded (see Table 2.4).
- Step 5: Identify values for the included cost categories.
- Step 6: Record cost metadata (e.g. date, currency).

In a similar style to the layout of the CE effectiveness summary paragraphs (Section 4.3.2) the following suggestions give suggested standard frameworks for reporting different elements of costs in line with cost reporting frameworks (e.g. Iacona et al. 2018). This is a flexible format covering a range of levels of detail.

For marginal intervention costs for the conservation organisation (probably the majority of the reported costs):

First sentence on total costs and what is included:

Total cost was reported as XX in [currency;] in [year] (XX USD) at [information on the scale e.g. XXkm, per hectare; breakdown not provided]. Costs are given over a [XX] time horizon and a discount rate of [XXX] applied. They included [types of cost] and excluded [types of cost] /

Possible additional sentence on the financial benefits of action.

Financial benefit was reported as: XX in [currency; [XX equivalent USD]; breakdown not provided), resulting in a net cost of XX. The benefits only included [types of benefit].

Additional sentence on the breakdown of the specific costs if provided.

Cost breakdown: XX [XX USD] for XX materials, XX [XX USD] for XX hours/days of labour, XX [XX USD] for overheads, XX [XX USD] for capital expenditure... [continue for all types of cost reported] (For further details see original paper).

For project level or organization costs, where the costs of interventions are given as part of a broader package:

Total cost of the action was reported as part of a [describe type and scale of conservation project]. The cost of this project was reported as XX in [currency; no breakdown provided] including [specific activities]. Costs are given over a [XX] time horizon and a discount rate of [XXX] applied. They included [types of cost] and excluded [types of cost]

Note: *Where costs or benefits are incurred by actors other than the conservationist/researcher, this should be noted. E.g., ‘the total included Opportunity costs to local communities, payments made by volunteers to travel to the nature reserve.’*

As examples:

Total costs reported as £1,000 GBP in 2022 (\$1,200 USD) for 2km of fencing. Costs are given over a 5 year time horizon with no discount rate applied. Cost breakdown: £100/2km for wire fencing and £900 for 5 day of labour for 4 staff. Financial benefit was reported as £200 (\$240 USD), resulting in a net cost of £800 (\$960 USD). Benefits included the reduced loss of livestock to local stakeholders.

Total cost reported as €10,000 EUR in 2008 (\$12,000 USD) for the translocation of 10 animals (€1000 / animal). Cost breakdown: €2000 [\$2400 USD] for tracking devices, supplementary food, and veterinary supplies, €6000 [\$7200 USD] for one staff employed by the project for two months and €2000 [\$2400 USD] for overheads (project management). Financial benefits reported as €10,000 (\$12,000 USD), resulting in a net cost of €0 (\$0 USD). Benefits included the estimated tourism value of each reintroduced eagle, although benefits are accrued by the wider sector.

As an example of a brief version:

In 2008, this project cost an extra US\$550 comprising \$250 for fuel and \$300 for chainsaw and safety equipment.

4.4 Presenting Evidence Quality

4.4.1 Information, source, relevance (ISR) scores

Figure 4.1 shows a range of information, source, reliability (ISR) cuboids as introduced in Chapter 2. For a particular assumption in a particular context, the ISR score is an assessment of whether a particular piece of evidence, whether a systematic review or observation, is relevant, and was generated in a reliable manner and from a source considered reliable. Such assessments will often not be public, especially if it would be embarrassing for the source to discover how their reliability is assessed. In other cases, it may be possible to compare different examples of the same type of evidence using predefined criteria or tools available for specific evidence types. One such tool is the CEESAT checklist, <https://environmentalevidence.org/ceeder/about-ceesat/>, which enables users to appraise and compare the reliability of the methods and/or reporting of evidence reviews and evidence overviews (such as systematic maps) by assessing the rigour of the methods used in the review, the transparency of the reporting, and the limitations of the review or overview.

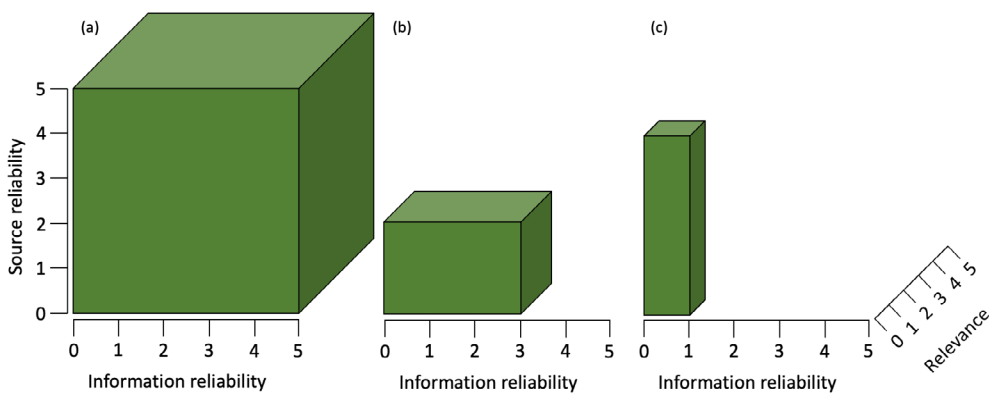


Figure 4.1 Cuboids of different strengths of evidence can range in weight from 0 (no weight) to 125 (5^3). In practice they will usually be cuboids of different dimensions. (a) strongly supported on all three axes; ISR = 5.5.5, (b) limited information and source reliability but highly relevant; ISR = 3,2,2 (c) high source but little information reliability or relevance; ISR = 1,1,5. (Source: authors)

Chapter 2 outlines some factors that influence the scores given for each of the three ISR elements. Our suggestion is that ISR becomes a standard way of assessing the weight of different pieces of evidence including different types of knowledge with the ISR score at the end of the claim: *During lockdown, when the reserve was closed, cranes were said to have nested within 3m of the trackway in part of a reserve where they have never nested before* ($I = 3, S = 4, R = 5$).

The weight of a piece of evidence can then be presented by multiplying the three ISR axes up to a maximum weight of 125 (5^3). Of course, if the value of one axis is zero, meaning that an element of the evidence either cannot be trusted or is judged to be irrelevant, then the weight is zero. Table 4.3 suggests how these weights can be converted into descriptions of the strength of evidence for a single piece of evidence.

Table 4.3 Conversion of weights of single pieces of evidence (from multiplying three axes) into descriptions of evidence strengths.

Weight	Description of evidence strength
0–1	Unconvincing piece of evidence
2–8	Weak piece of evidence
9–27	Fair piece of evidence
28–64	Reasonable piece of evidence
65–125	Strong piece of evidence

4.5 Balancing Evidence of Varying Strength

Whilst in some cases, evidence will either clearly refute or support an assumption (as in the above examples), in many situations there is a range of relevant evidence that varies in its strength and direction in relation to a particular assumption. This is often the case when assessing the effectiveness, costs, acceptability, or feasibility of actions, where evidence can, for example, show an action to be beneficial to different extents, have no effect, or have negative impacts. Where this is the case, one approach, modified from Salafsky et al. (2019), is to imagine placing the various pieces of evidence along a balance according to the direction and magnitude of the effect they report, from large negative effects, through to large positive effects (in this case, effectiveness; Christie et al., 2022) (Figure 4.2).

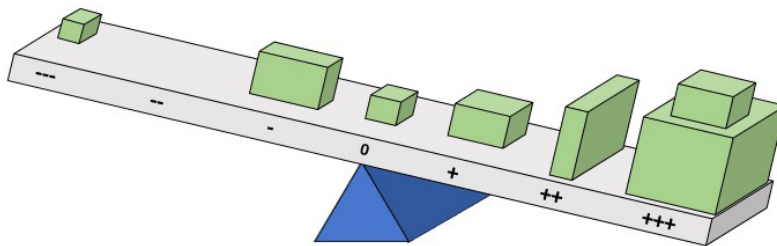


Figure 4.2 A means of visualising the balance of evidence behind an assumption. The pieces of evidence are placed on the balance according to the extent to which it supports or rejects the assumption. The greater the tilt the greater the confidence in accepting (or rejecting) an assumption. (Source: authors)

How can this visualisation be converted into a conclusion? Table 4.4 provides a way of assessing the evidence for an assumption about an action. The cumulative evidence score is the sum of the weights of each block of evidence. The weight of each block can be calculated by multiplying the ISR (information, source, relevance) scores (Section 4.2.2). Adding the weights of all evidence blocks together gives the cumulative weight of evidence. The mean effectiveness category comes from assessing the effectiveness associated with each piece of evidence and then taken

as the average (the weighted median — place the evidence in order of effectiveness, add up the ISR values until you reach the figure that is half the value for the cumulative evidence, take the effectiveness linked to piece of evidence).

Table 4.4 Converting the combined evidence into statements of the strength of evidence.
The maximum score for a piece of evidence is 125 (5 x 5 x 5) so strong evidence requires, say, one really strong study conducted in a manner that minimises bias and overwhelming evidence requires over two high-quality studies.

Cumulative evidence score	Evidence category
>250	Overwhelming evidence
101–250	Strong evidence
51–100	Moderate evidence
11–50	Weak evidence
1–10	Negligible evidence

4.5.1 Statements of effectiveness

The cumulative evidence and mean effectiveness can be combined to give statements of the effectiveness for actions — the Strategic Evidence Assessment (SEA) model (Table 4.5). These statements can then be used in other processes, such as writing management plans or funding applications.

4.5.2 Words of estimative probability

When President Kennedy was deciding whether to proceed with the proposal for the disastrous Bay of Pigs invasion of Cuba, the US Joint Chiefs of Staff stated that the proposal had a ‘fair chance’ of success, which Kennedy interpreted as likely to work. The Chiefs actually meant that they judged the chances of success as ‘3 to 1 against’ (Wyden 1980). As a result of such lessons, the accepted least ambiguous approach is to attempt estimating probabilities. Studies have shown, at least in medicine, that giving natural frequencies ‘Three out of every 10 patients have a side effect from this drug’ leads to fewer problems of interpretation than probabilities, such as ‘You have a 30% chance of a side effect from this drug’ (Gigerenzer and Edwards, 2003).

Some dislike judging probabilities (how certain we are about something) or estimating frequencies (how often something happens), partly because it seems to give unjustified accuracy. An alternative approach is to express likelihoods in the form of standardised terms. Table 4.6 gives the language for communicating probabilities recommended by the Intergovernmental Panel on Climate Change (IPCC, 2005). Table 4.7 gives the opposite: examples of some ‘weasel’ words that are likely to be ambiguous and so should be avoided.

Table 4.5 The Strategic Evidence Assessment model. Converting the combined evidence into statements of effectiveness combining the scores from Table 4.4 with effectiveness.

Cumulative evidence score	Median evidence category	Considerable harms	Moderate harms	Minor harms	No effect	Little benefit	Moderate benefit	Considerable benefit
>250	Overwhelming evidence	Overwhelming evidence of considerable harms	Overwhelming evidence of moderate harms	Overwhelming evidence of minor harms	Overwhelming evidence of no effect	Overwhelming evidence of little benefit	Overwhelming evidence of moderate benefit	Overwhelming evidence of considerable benefit
101–250	Strong evidence	Strong evidence of considerable harms	Strong evidence of moderate harms	Strong evidence of minor harms	Strong evidence of no effect	Strong evidence of little benefit	Strong evidence of moderate benefit	Strong evidence of considerable benefit
51–100	Moderate evidence	Moderate evidence of considerable harms	Moderate evidence of moderate harms	Moderate evidence of minor harms	Moderate evidence of no effect	Moderate evidence of little benefit	Moderate evidence of moderate benefit	Moderate evidence of considerable benefit
11–50	Weak evidence	Weak evidence of considerable harms	Weak evidence of moderate harms	Weak evidence of minor harms	Weak evidence of no effect	Weak evidence of little benefit	Weak evidence of moderate benefit	Weak evidence of considerable benefit
1–10	Negligible evidence	Negligible evidence of considerable harms	Negligible evidence of moderate harms	Negligible evidence of minor harms	Negligible evidence of no effect	Negligible evidence of little benefit	Negligible evidence of moderate benefit	Negligible evidence of considerable benefit

Table 4.6 Terms suggested by the IPCC (2005) for referring to probabilities.

Probability	Recommended term
>99%	Virtually certain
90–99%	Very likely
66–90%	Likely
33–66%	About as likely as not
10–33%	Unlikely
1–10%	Very unlikely
<1%	Exceptionally unlikely

Table 4.7 Examples of ‘weasel’ terms likely to be ambiguous.

A chance	Cannot dismiss	Could	Maybe	Perhaps	Somewhat
Believe	Cannot rule out	Estimate that (or not)	Might	Possibly	Surely
Cannot discount	Conceivable	May	Minor	Scant	Suggest

4.6 Visualising the Balance of Evidence

In some cases it may be particularly informative to visualise the collated evidence base, ideally communicating both the weight of the evidence pieces and the results they found. Here we present two different options for visualising a collated evidence base that are best suited to slightly different types of results.

4.6.1 Ziggurat plots

Ziggurat plots (Figure 4.3) are best suited to results where the outcomes can be expressed categorically (e.g. varying discrete levels of effectiveness). The width of each bar represents its ISR score (out of 125) with studies piled in order of ISR score to give a ziggurat shape. This has the merit of showing the distribution of the evidence.

In some cases, the median may be a poor representation of the distribution, for example, because the results are strongly bimodal. In this case add a comment e.g. ‘median effect size mixed, but results strongly bimodal’.

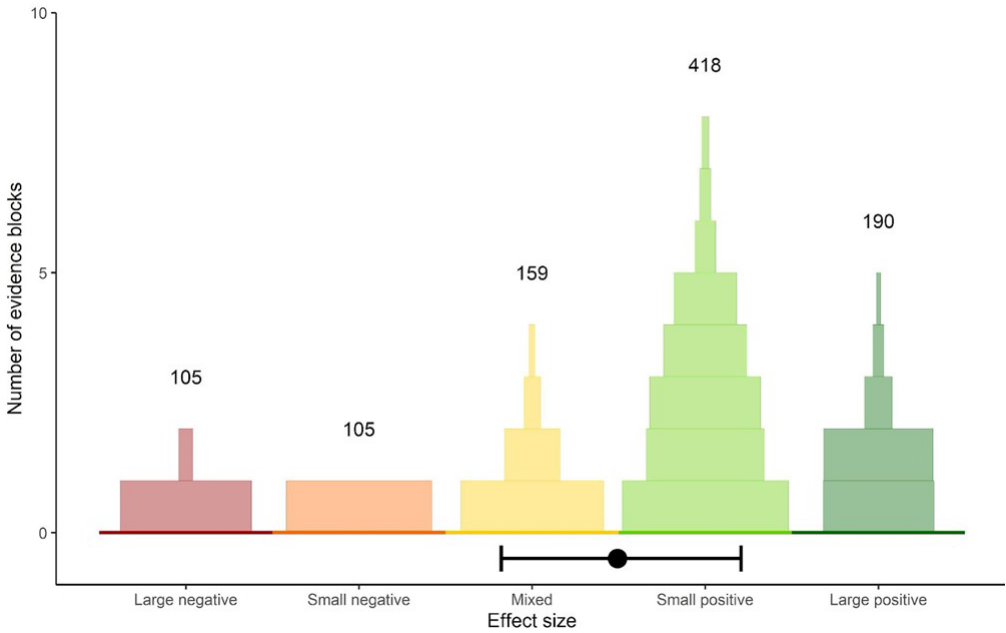


Figure 4.3 A ziggurat plot in which each study is a horizontal bar whose width is the information reliability, source reliability, and relevance (ISR) score (up to 125). The studies are collated for different categories of effectiveness. The number above each pile of evidence blocks shows the total evidence score for that pile. In order to derive an average effect size, we defined effect size categories as consecutive integers (large negative = -2; small negative = -1; mixed = 0; small positive = 1; large positive = 2) and calculated the weighted mean (filled black point) and 95% confidence intervals. The confidence interval was calculated by bootstrapping the weighted mean. (Source: authors). The following R code can be used to carry out this process, which we have adapted from code produced by Stackoverflow users Tony D and Ben <https://stackoverflow.com/questions/46231261/bootstrap-weighted-mean-in-r>. Text in quote marks must be altered by the user:

```
library(boot)
df <- data.frame(x= "degree of support", w= "ISR score")
wm <- function(d,i){
  return(weighted.mean(d[i, 1], d[i, 2]))
}
bootwm <- boot(df, wm, R=10000)
boot.ci(boot.out = bootwm)
```

Figure 4.4 shows a range of possible outcomes. In some cases, the evidence may be clear and convincingly show either a positive effect or no effect. In others the results may be bimodal, for example, if combining studies of reptiles but lizards and snakes show very different responses. In other cases, the responses may be unclear – perhaps as a range of different unknown variables are important.

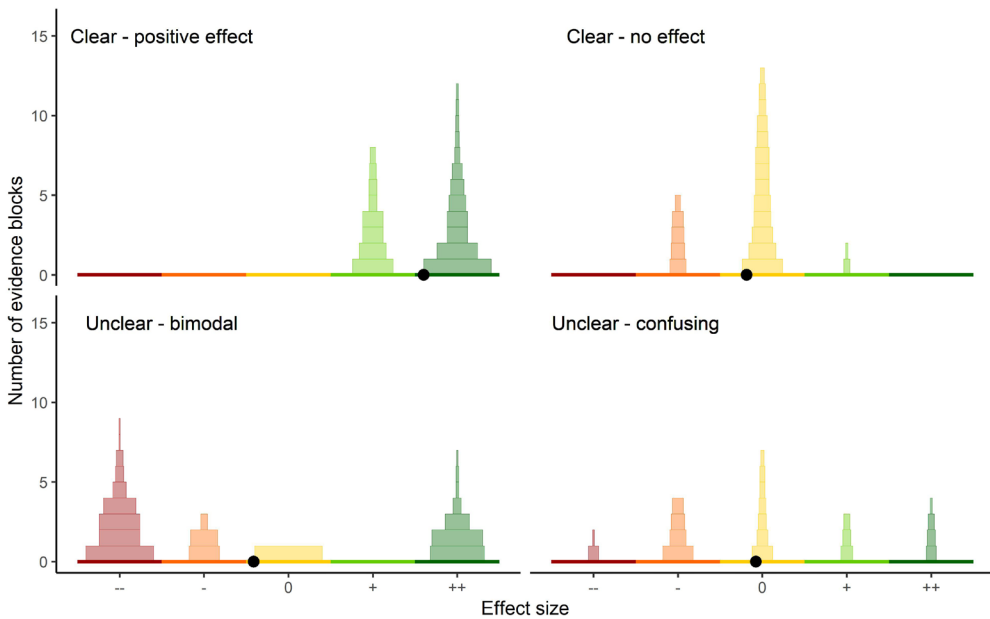


Figure 4.4 A range of possible outcomes of ziggurat plots. Those on the top show clear results, with either a moderately large effect size (left) or evidence for negligible effect (right). Those underneath show unclear effects as either bimodal (left) or scattered results (right). The filled black point shows the average effect size, calculated as the weighted mean. (Source: authors)

4.6.2 Weighted histogram plots

For continuous results (an effect size, or counts of a particular species, for example) however, a ziggurat plot may hide useful information, and the collated evidence base can be better visualised with a weighted histogram plot (Figure 4.5). Sources of evidence are ordered by the magnitude of the result they report, and the height of the bar is the IRS score (out of 125). Evidence pieces that report the same effect sizes are stacked on top of each other.

4.7 Synthesising Multiple Evidence Sources

Combining knowledge sources is key. The evidence that decision makers draw on will usually be a combination of scientific knowledge, local knowledge and experience. For example, when managing a reserve, much of the decision making will be based on the experience and local knowledge of the reserve manager and team ('I think this area is too close to the forest for the species to breed', for example), informed by knowledge from the local community, local naturalists, fishers, etc and scientific knowledge that has generality.

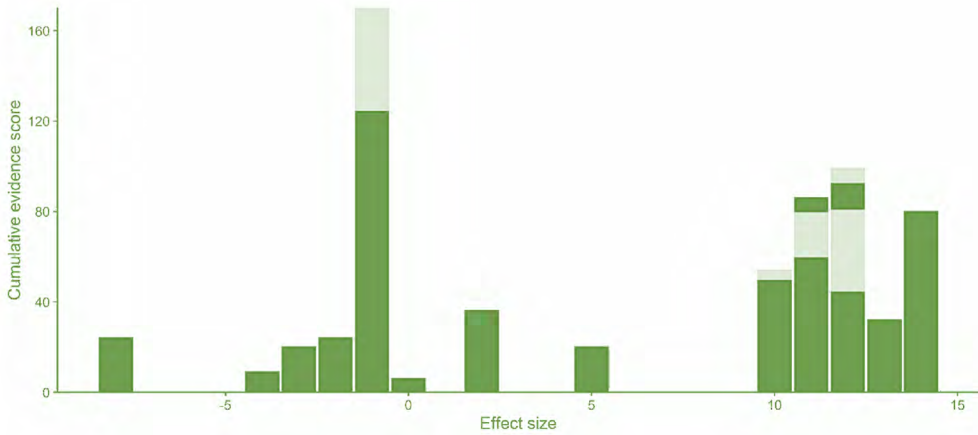


Figure 4.5 A weighted histogram plot in which each piece of evidence is represented by a vertical bar, whose height is the information reliability, source reliability, and relevance (ISR) score. Evidence pieces are arranged in order of the magnitude of the effect size, and where they report the same effect size, they are stacked on top of each other, with alternating colours for each evidence piece in the stack. (Source: authors)

4.7.1 Evidence review summary statements

The result of the collation or analysis can also be presented in a summary. The aim is to describe the results so that they provide a summary of the amount and nature of the evidence (e.g. quality/reliability, coverage: species, habitats, locations), what it tells us (e.g. effects on key outcome metrics), and any limitations. The different elements of this chapter can be combined, as shown in Table 4.8, to provide a summary of the evidence.

Table 4.8 The different main elements of summarising evidence described in this chapter with an illustrative sentence.

Element	Example
Type of search	Searching for ‘earthworm’ on CEEDER to look for the impact of invasive earthworms
Type of evidence	produced one good quality meta-analysis from 2018 based on 430 observations from 30 independent studies,
Constraints on generality statement	with many studies looking at impacts on Collembola and oribatid mites but few on mycorrhizal fungi,
Study conclusions	showing compelling evidence for negative effects on below-ground biodiversity of recipient ecosystems,
Information on costs	but provided no figures on costs.

Such summaries can highlight knowledge clusters and thus gaps in the evidence, which are important to take into consideration. However, it is important that summaries are not used to

vote count, i.e. draw conclusions based on the number of studies showing positive vs negative results, which is usually a misleading method of synthesis (Stewart and Ward, 2019). Studies are generally not directly comparable or of equal value, and factors such as study size, study design, reported metrics, and relevance of the study to your situation need to be taken into consideration, rather than simply counting the number of studies that support a particular interpretation. Chapter 2 provides suggestions for describing single studies.

In the UK three replicated randomised controlled experiments showed that adding a collar and bell to domesticated cats reduced the predation rate of small mammals by about a half; one of these also showed that effects were similar for bells and a collar-mounted sonic device. Experiments outside the UK showed that a pounce protector (a neoprene flap that hangs from the collar) stopped 45% of cats from catching mammals altogether. A study in the USA showed cats provided with brightly patterned collars brought home fewer mammals than did cats with no collars in autumn, but not in spring. In all cases the cats had a history of bringing back whole prey and the catches were recorded by their owners.

Or for a brief version:

Domesticated cats took roughly half as many small mammals if fitted with collar and bell, collar and sonar device or a neoprene flap hanging from the collar. Brightly coloured collars were less effective.

4.7.2 Evidence capture sheets

Perhaps the most straightforward method for bringing together multiple sources of evidence is to use an evidence capture sheet. Table 4.9 presents an example of such a sheet, with each row constituting a different source of evidence, and each column providing vital information about the validity of the source and the result that it presents. While evidence capture sheets comprehensively communicate the key information of each source of evidence, they stop short of providing an overall conclusion of what the evidence base says.

4.7.3 Argument maps

Argument maps, also called argument diagrams, have three main roles: easing the creation of a logical argument, presenting the basis of an argument so it is easy to follow and reframing an existing argument so making the assumptions, logic and any gaps transparent. The heart of the argument map is to present the case for and against a position in a logical manner. Their merit is that, if done well, they organise and clarify the evidence and reasoning for a contention. These can be done collectively as a way of understanding the basis of any disagreement. They have the considerable advantage that they can easily be converted into text for a decision maker to make the final decision. If using a software package this conversion can be done automatically.

Table 4.9 Example of tabular presentation of evidence for a proposed project that plans to introduce natural grazing with ponies to the montado habitat in Iberia to increase biodiversity. I = information reliability, S = source reliability, R = relevance all on 0–5 scales. Note the evidence strength is for the particular problem rather than an overall assessment.

Evidence source ^a	Type of evidence ^b	Direction and strength of results ^c	I ^d	S ^e	R ^f	Description of evidence strength*
Experience of colleague	Anecdotal	Colleague has been told that there used to be higher densities of livestock and also more flowers in the area 50 years ago	1	2	5	Weak
Scientific study (Leal et al. 2019)	Experimental	A replicated controlled study found that moderate levels of grazing benefited most species of bird.	3	5	5	Strong
Grazelife report www.grazelife.com	Observational? (not specified)	Herbivory leads to more diverse vegetation and benefits many species	3	3	4	Moderate
https://www.conservationevidence.com/actions/1628	Synthesis	Three of five studies across Europe found that increasing grazing intensity increased plant diversity in shrublands	2	5	3	Moderate

a. E.g. peer-reviewed article, expert opinion, grey literature report, personal experience.

b. E.g. synthesis, experimental, observational, anecdotal, theoretical/modelling.

c. Was the result strongly positive, weakly positive, mixed, or no effect?

d. How convincing is the evidence? Depends on the type of evidence, for example, may depend on experimental design and sample size. (0–5)

e. Is the source biased or independent? (0–5)

f. Does it apply to your problem? This will depend on habitat, geography or socio-economic similarity. (0–5)

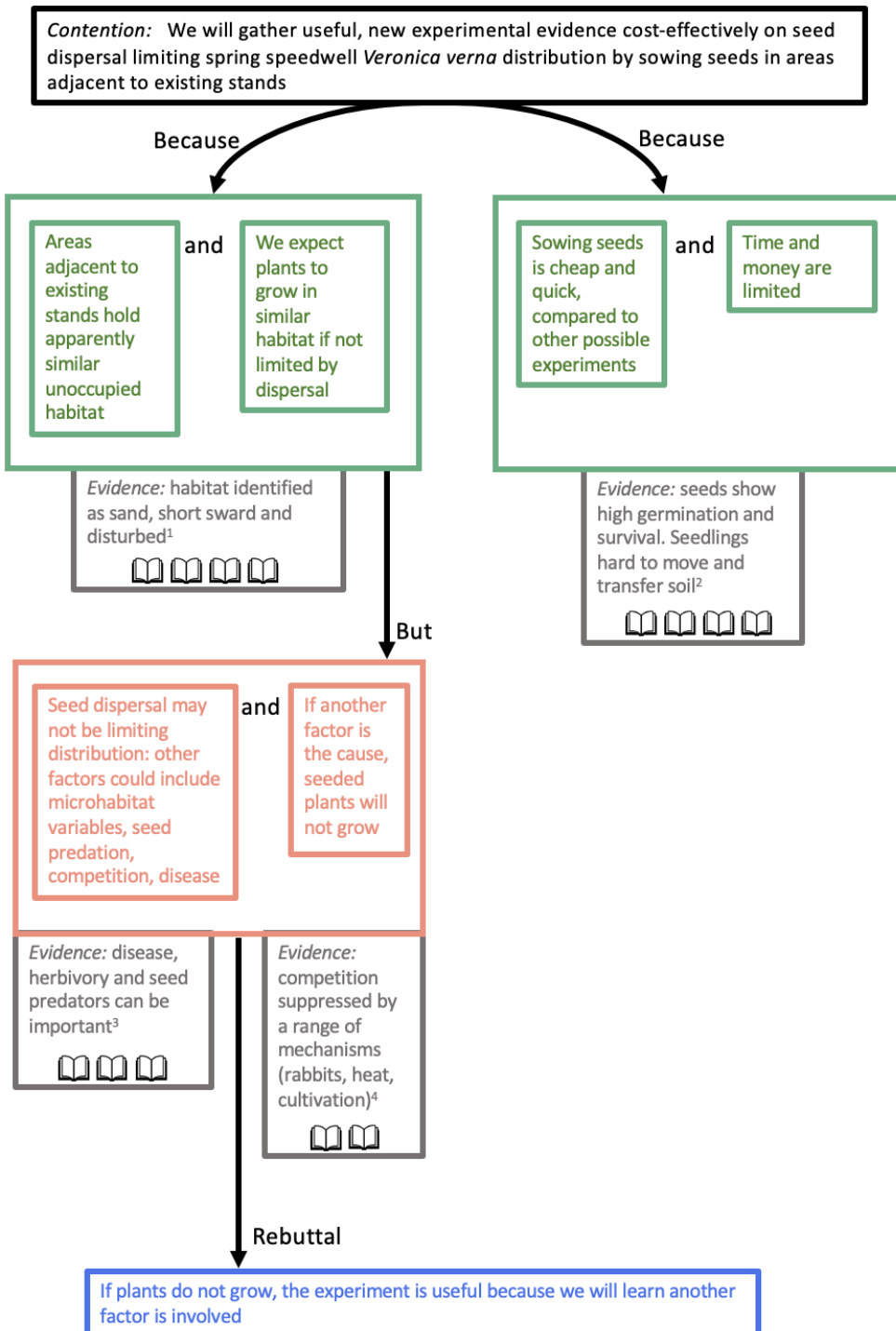


Figure 4.6 An example of an argument map to decide whether or not to introduce a plant to new locations. The main elements are: main contention, premise, counterargument, rebuttal and evidence. 📖📖📖📖📖 = overwhelming evidence; 📖📖📖📖 = strong evidence; 📖📖📖 = moderate evidence; 📖📖 = weak evidence; 📖 = negligible evidence. The evidence is elaborated in Box 4.2. (Source: authors)

Box 4.2 Evidence sources for plant reintroduction argument map

1. Heathland sites with *Veronica verna* all have low skeletal vegetation and open patches (Watt 1971). Experiments show it benefits from disturbance (Nabe-Nielsen et al. 2021).
2. In an experiment 22% of 700 *V. verna* seeds germinated in the autumn after sowing (Norman Sills pers. comm.). Practitioner experience is that seedlings are harder to transplant and result in transferring soil (Tim Parkhurst pers. comm.).
3. Extensive ecological literature shows there can be a serious impact of disease, herbivory and seed predators on annual plants. Herbivory by unknown invertebrates and slugs occurred on closely related species (Boutin and Harper, 1991). Five of 11 seedlings in an outdoor box were grazed (by slugs?), but four regrew (Norman Sills pers. comm.). Potentially could limit distribution but entirely speculative.
4. All existing sites with little competition. Rabbit grazing is traditionally important (Back from the Brink, 2021). At Weeting Heath arable weed reserve is cultivated. Hot dry conditions seem likely to limit competitors (*V. verna* grows over winter and flowers in spring) but no evidence.

They start with a statement, the main contention, that could be correct or not. The arguments for it being right, the premises, are then listed followed by the arguments for it being wrong, the counterarguments. Co-premises are two statements that both have to be correct. Lines are only drawn between boxes to specify that something is a reason to believe or a reason not to believe something else. The process forces the logic as to whether you should believe the claim and why or why not.

There are three basic rules (<https://www.reasoninglab.com/wp-content/uploads/2013/10/Argument-Maps-the-Rules.pdf>):

1. The **Rabbit Rule** (you cannot pull a rabbit out of the hat by magic): every significant word, phrase or concept appearing in the contention must also appear in one of the premises; i.e. every term appearing above the line has to appear below the line. In the example here the terms are sow, spring speedwell, seeds, adjacent areas, distribution, and dispersal.
2. The **Holding Hands Rule**: every term appearing below the line has to appear above; if something appears in a premise but not in the contention, it must appear in another premise.
3. **The Golden Rule**: every simple argument has at least two co-premises.

Figure 4.6 shows an argument map created by members of the Breckland Flora Group (Jo Jones, David Dives, Julia Masson, Tim Parkhurst, Norman Sills, William Sutherland, and James Symonds) guided by Mark Burgman. They are interested in planning a reintroduction programme for spring speedwell *Veronica verna*. Creating this map helped think through the logic for carrying out an experiment to assess whether the species is restricted in range due to

lack of dispersal or whether there it is constrained by very specific, but not fully understood, habitat requirements.

Argument maps can easily be sketched out, but there are numerous argument-mapping tools available. Argument mapping has been adopted by the New Zealand government for biosecurity analyses.

4.7.4 Mind maps

Mind maps, as popularised by Buzan (1974), express the relationships between different aspects of a problem. They are also sometimes called spider diagrams or brainstorms. Although often considered as a means of note taking, they can also be used for generating new ideas. Of relevance here they can be used to summarise information around an issue. They are a less structured process than a theory of change or argument map and so are useful for pulling together information to provide a clear account of the components of the issue. One advantage of summarising with such a visual model is that it makes the logic explicit and thus invites further comments. They can be drawn by hand or created using various packages available. These then can be created by a person or team putting together the strands of evidence.

These are created by starting with a central idea or topic, placed in the middle of the map. Branches are then added to this central topic. Each branch is then explored. If the evidence is extracted it can be assessed and added. Another option is to draw a preliminary map based on expert opinion and then adjust as evidence is added.

Figure 4.7 shows an example of a mind map created after a group discussion with experts (Malcolm Ausden, Jennifer Gill, Rhys Green, Jennifer Smart, William Sutherland, and Des Thompson). Box 4.3 lists some supporting evidence. This is intended to be illustrative of the process rather than a comprehensive review of the literature and concepts.

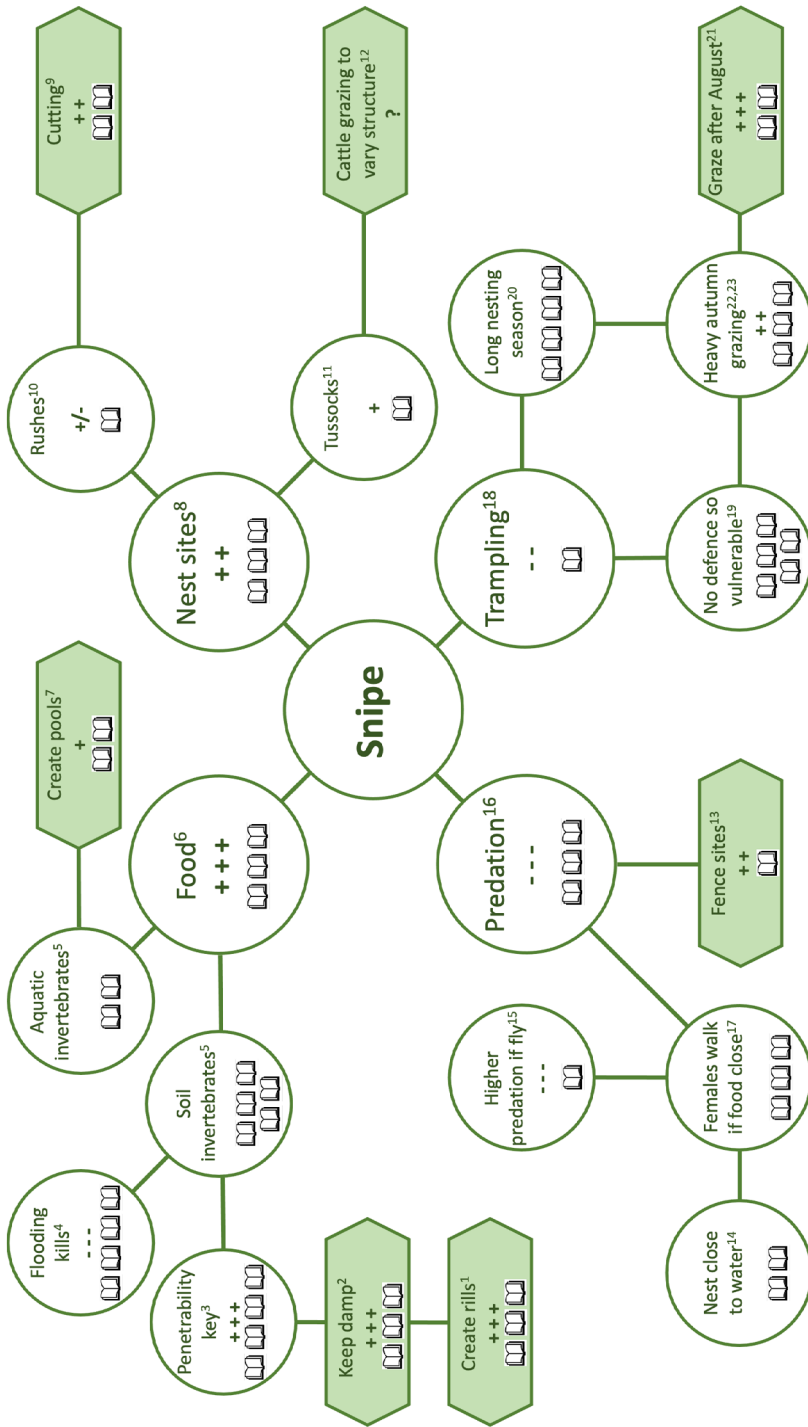


Figure 4.7 A mind map of possible means of managing lowland grassland habitat to improve the common snipe *Gallinago gallinago* population.
 [Icon: 100% filled] = overwhelming evidence; [Icon: 75% filled] = strong evidence; [Icon: 50% filled] = moderate evidence; [Icon: 25% filled] = weak evidence; [Icon: 0% filled] = negligible evidence; ? = no evidence; --- = considerable harm; -- = moderate harm; +/- = mixed effect; + = little benefit; ++ = moderate benefit; +++ = considerable benefit. See Box 4.3 for the evidence sources. (Source: authors)

Box 4.3 Evidence sources for mind map on snipe management

1. Creating rills (shallow linear features) initially beneficial in Otmoor enabling snipe to breed even where soil type doesn't allow them to probe but numbers subsequently crashed (Malcolm Ausden, Rhys Green, pers. comm.)
2. Penetration resistance of the soil related to depth of water table (Green, 1988)
3. Breeding season shorter if soil becomes impenetrable (Green, 1988)
4. Earthworms killed if flooded (Ausden et al., 2001)
5. Mainly feed on earthworms and Diptera larvae from soil but some aquatic invertebrates especially Chironomid larva (Green, 1988; Cramp and Simmons, 1983)
6. Largely restricted to areas where food is available; snipe were more likely to persist in fields where the soil conditions were wet and soft (Smart et al., 2008)
7. Pools used when cannot probe ground or there is low soil invertebrate biomass nearer to the nest (Green et al., 1990) Pools used when cannot probe ground so last resort (Rhys Green, pers. comm.)
8. Restricted to sites with some long grass (widespread experience; Cramp and Simmons, 1983)
9. Rush cutting benefits both lapwing and redshank (Robson and Allcorn, 2006); in Cumbria, England, number nesting increased from 1 to 11 following several interventions: rush cutting, higher water levels, more intensive grazing regime, and scrape creation (Holton and Allcorn, 2006). Unknown if it also benefits snipe.
10. Rushes provide some cover but deleterious if too abundant; upland snipe prefer rushes (Hoodless et al., 2007)
11. Some say prefer long, tussocky grass for nesting (Mason and MacDonald, 1976; Cramp and Simmons, 1983) but Rhys Green (pers comm) has found about 300 snipe nests but never one in a tussock and no relation between vegetation height and snipe density in wet grassland studied (Green 1986)
12. Widespread experience shows suitable habitat with muddy poached areas and suitable height for nesting best created by cattle grazing
13. Predator-exclusion fencing led to increased productivity of lapwing (Malpas et al., 2013)
14. In Otmoor 90% nest within 1 m of rill (Rhys Green, pers. comm) but no relation in upland site (Hoodless et al., 2007)

15. One of 10 radio-tracked snipe taken by kestrel; at least another eight snipe remains found around that kestrel's nest; likely to be more vulnerable if fly (Green et al., 1990)
16. 60% loss of eggs to predators (Green, 1988); some adults predated in the breeding season (Green et al. 1990)
17. Females walk to feed unless sites are within 70 m, are too dry, or have insufficient food (Green et al., 1990)
18. Nest loss is correlated with cattle density with 19% lost to trampling in one study (Green, 1988); in USA experiment, trampling is greater where cattle are present but overall success is unaffected (Popotnik and Giuliano, 2000)
19. Do not defend nest (widespread field experience; Cramp and Simmons, 1983) so particularly vulnerable to trampling
20. Long nesting season (Mason and MacDonald, 1976; Cramp and Simmons, 1983)
21. As can still be nesting in August if earlier clutches fail and fields remain suitably wet, it makes sense to delay grazing until late summer, but this can mean the cut hay is of low quality and therefore difficult to dispose of, and might affect sward quality for Snipe and other species in the longer-term (Malcolm Ausden, pers. comm.).
22. Grazing needed to maintain a suitable sward but can trample nests, so follows that heavier grazing in autumn can be used to help create a suitable sward for the following spring. Although with recent warmer winters autumn grazing may be insufficient to maintain short sward for lapwing creating, which may create demand for early spring grazing (Green, pers. comm.)

In practice most decisions appear to comprise an exploration of components of the issue, rather like an informal mental version of a mind map, followed by a pronouncement of the conclusion. Our experience is that it is often unclear what is justifying the decision. A sketched version of the discussion provides a reminder of the issues and helps force consideration of the components.

4.7.5 Evidence underlying theories of change

A theory of change (CMP, 2020) is a series of causally linked assumptions about how a team thinks its actions lead to intermediate results and to target outcomes (see Chapter 7 for further details). Displaying a theory of change in a strategy pathway diagram (see an example in Figure 4.8 below, repeated from 7.3) facilitates the identification of analytical questions and assumptions that can be tested with evidence.

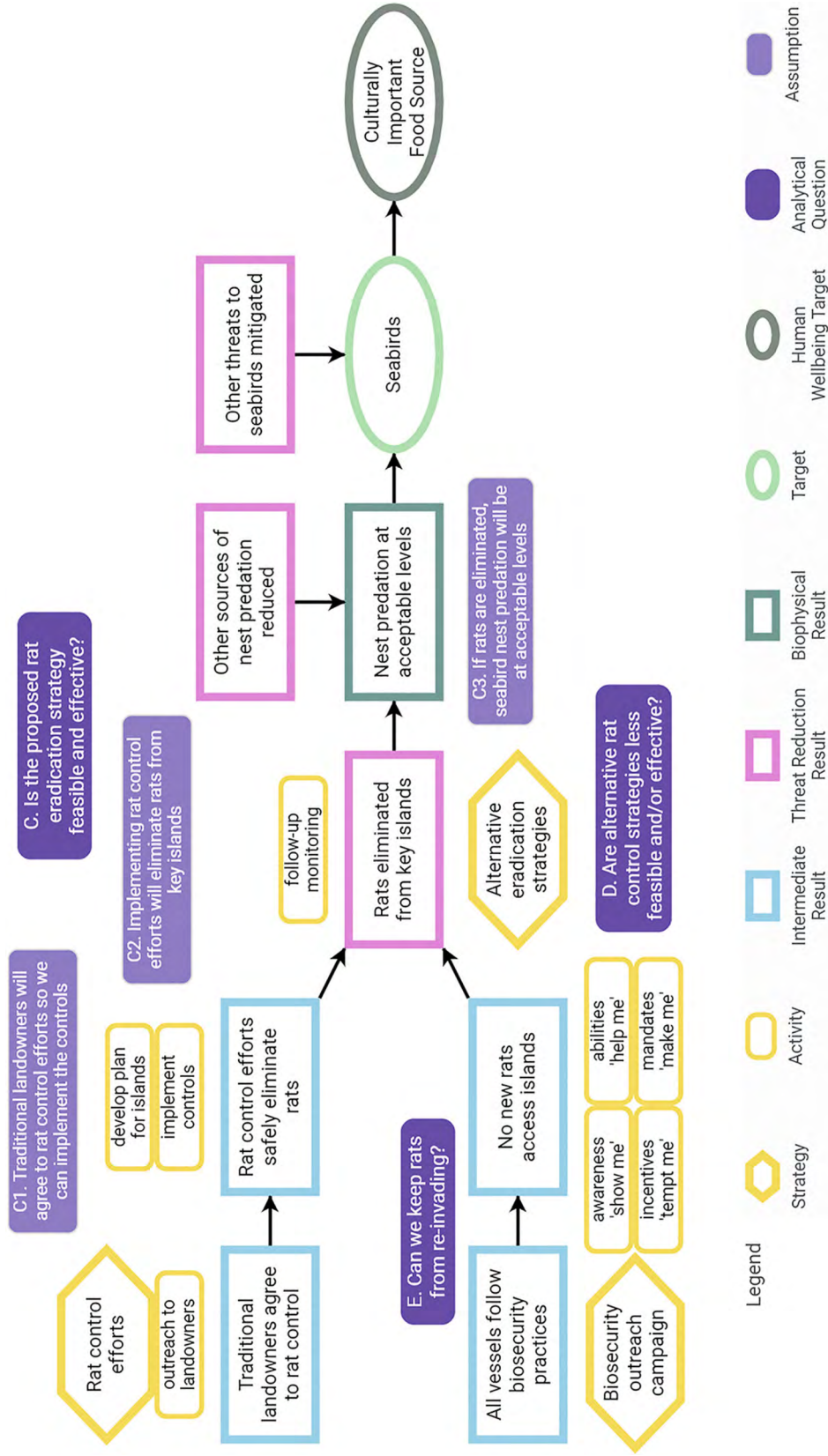


Figure 4.8 A theory of change pathway showing analytical questions and assumptions. This example is derived from several different real-world rat eradication projects. Note, assumptions for questions D and E are not shown. (Source: Salafsky et al., 2022, CC-BY-4.0)

Table 4.10 Summary of current evidence for analytical questions relating to the theory of change shown in Figure 4.7. (Source: Salafsky et al., 2022)

Analytical question	Who, date	Summary of current evidence for question	Current confidence in evidence	Implications	Additional information needs
C. Is the proposed rat eradication strategy feasible and effective?	Team, Jun 2021	While it is clear from other sites that rat eradication would be technically feasible in the short-term, it is not yet clear whether this will be socially acceptable given the potential lack of durability of the solution.	Need more info	Explore with the community whether eradication is acceptable.	Compare rat population levels in treated and untreated areas. Conduct key informant interviews with tribal elders.
D. Are alternative rat control strategies less feasible and/or effective?	Team Nov, 2021	Our experience is that other means of controlling rats such as trapping or introducing natural predators would not be technically feasible in our islands.	Confident, but...	Continue to monitor the emergence of new strategies that may be more effective.	Stay abreast of new literature/examples of rat eradication strategies.
E. Can we keep rats from re-invading?	Team, Dec 2021	It is not clear whether it is feasible to keep rats from re-invading; experience from other sites is that this requires a major investment of resources and that re-invasion may actually be facilitated by the previous invasive species.	Not confident	Unless we can figure out how to solve this problem, eradication will only be a temporary fix at best.	Stay abreast of new literature/examples of biosecurity strategies.

Once evidence is collected, theories of change can provide the backbone to assess the evidence and contextualise it for a respective conservation action. Often, evidence capture sheets (Table 4.10) are used to relate available evidence to the theory of change and form the basis for drawing overall conclusions about key questions and assumptions, and the effectiveness of the strategy pathway.

4.7.6 Bayesian networks

Bayesian networks (BNs), as outlined in Box 4.4, are powerful statistical models that represent suspected, theorised or known causal relationships between variables supported by data and knowledge (i.e. evidence) (Jensen, 2001; Stewart et al., 2014). They comprise a network of nodes (i.e. variables or factors) connected by directed arcs (i.e. edges, links or arrows) contextualising the underlying probabilistic relationships. BNs can be built by combining a range of data sources including empirical evidence and expert elicitation. These models are particularly useful because they explicitly and mathematically incorporate uncertainty in a transparent way (Landuyt et al., 2013). Due to the graphical nature of BNs, complex information can be

Box 4.4. A simple example of a Bayesian network

Imagine a scenario where we arrive home from work and notice the grass in our front garden is wet. There are only two possible causes of the wet grass; it has rained, or our sprinkler (automatic watering system) has been on. How can we decide what is the cause of the wet lawn?

First, we build a causal network representing the relationships between the three nodes (Figure 4.9).

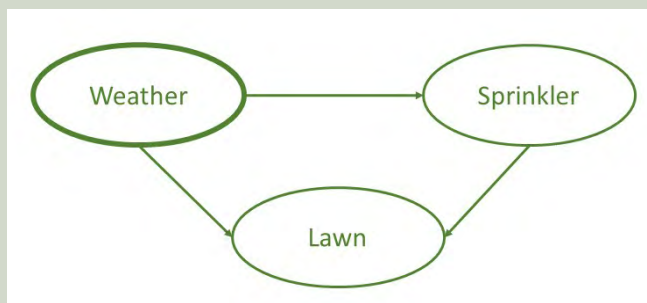


Figure 4.9 A causal diagram of the relationship between the three nodes in the system. The state of the Sprinkler node and the Lawn node are conditionally dependent on the state of the Weather node. (Source: authors)

Each node has a discrete number of states it can be in. The weather node can be in one of three states: Sunny, Cloudy or Rainy. The Sprinkler node can be in one of two states: On or Off and the Lawn node can also be in one of two states: Wet or Dry.

We are not sure what the weather was but we can use the weather forecast which tells us the probability that it was sunny is 40%, that it was cloudy is 30% and that it was raining 30%. We can use expert opinion or data to build the underlying conditional probabilities tables which encode the relationships between the nodes (Table 4.11). The probability that the Sprinkler is On when the state of the Weather node is Sunny is 80%. The probability that the Sprinkler is On when the state of the Weather node is Rainy is 10%. Note that each column in the conditional probability table sums to 100%.

Table 4.11 A conditional probability table for the Sprinkler node given the three states of the Weather node.

Weather	Sunny	Cloudy	Rainy
Sprinkler On	80	30	10
Sprinkler Off	20	70	90

If we have no evidence about the state of any of the nodes, we can see that the model displays the probability of each state for each node (Figure 4.10). When we have no evidence about the system, we can still gain useful information from the conditional probabilities. In this case, the state of the Lawn node is more likely to be Wet.

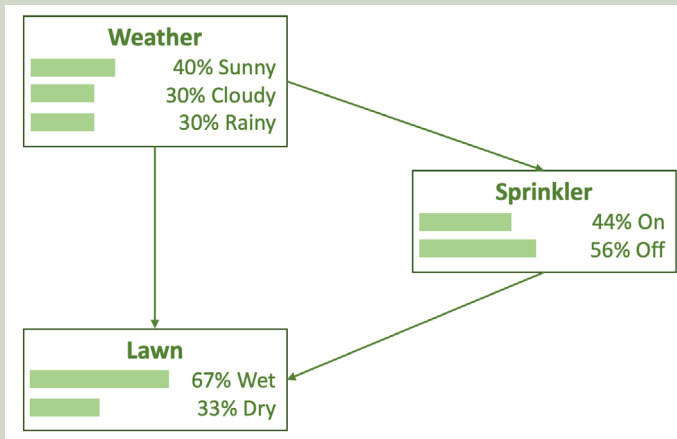


Figure 4.10 A Bayesian network of the lawn and sprinkler system shown in Figure 4.9 including the expected weather conditions, with the resulting consequences for sprinkler use. The probability of rain and the probability of the sprinkler being on gives the overall probability that the grass will be wet. (Source: authors)

The true power of Bayesian networks is that they can not only calculate complex combinations of conditional probabilities to deduce the probable state of a node (something that humans struggle to do) but that we can use evidence from any node in the network to determine the probable state of any other node (called abductive learning).

As we know the state of the Lawn node (it was wet when we arrived home) we can set the node (known as 'setting evidence') to 100% and that will give us the probability of each state in the Weather and Sprinkler nodes (Figure 4.11). If the Lawn is 100% Wet then with no knowledge of the state of the Sprinkler we know that the probability that it was sunny is higher than the probability that it was rainy or cloudy. Therefore the probability that the Sprinkler is the cause of our wet lawn is 59%.

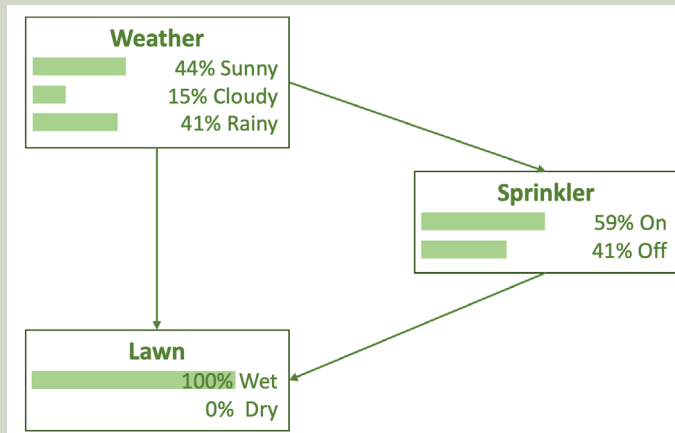


Figure 4.11 The Bayesian network shown in Figure 4.10 but with the lawn set to wet (as observed). From setting the lawn node can then determine the state of the weather and sprinkler nodes. (Source: authors)

represented in an intuitive manner that is easily interpreted by non-technical stakeholders (Spiegelhalter et al. 2004).

Developments in computing technology has allowed the rapid development of BNs over the last 10 to 15 years. BNs can now include both categorical and continuous variables and can be dynamic (across space and/or time), which has widened the scope for the application in applied conservation. For example, Stephenson et al. (2018) developed a dynamic BN to investigate the drivers of changes in pot-fishing effort distribution along the Northumbrian coastline of Northeast England across seasons (Figure 4.12). Empirical data on fishing vessel characteristics, habitat type and catch statistics were combined with qualitative data from interviews with fishers. This allowed for the identification of the key drivers of the change in fishing activity between 2004 and 2014; technological changes have increased the ability of fishers to fish in poor weather conditions. Stephenson et al. (2018) could then use their BN to develop management scenarios and test the possible outcomes of changing, for example, regulations on the number of fishing days or the exclusion of fishing from certain areas along the coast.

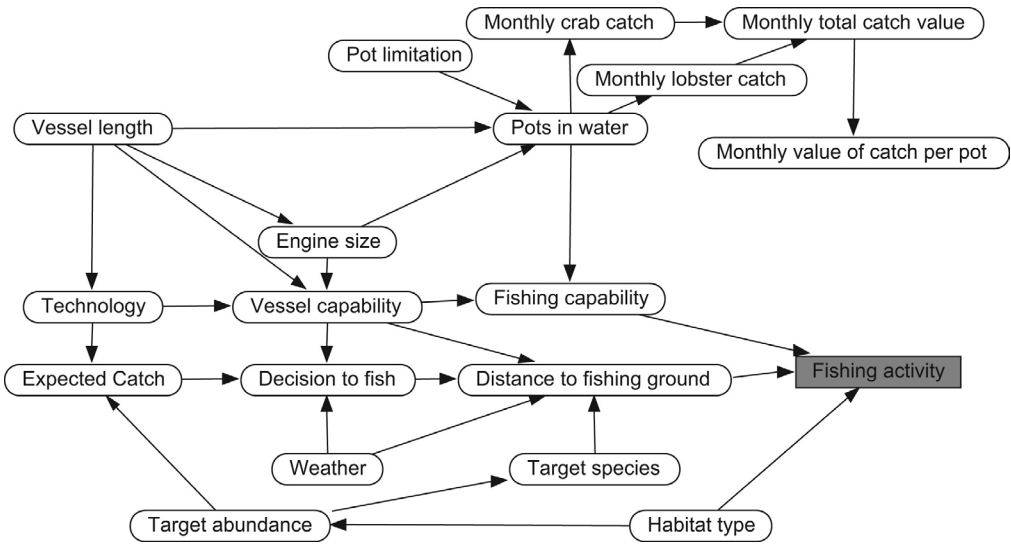


Figure 4.12 Inference diagram of the factors influencing pot-fishing activity along the Northumbrian coast, UK. (Source: Reprinted from Stephenson et al., 2018, with permission from Elsevier)

BNs can be used in evidence synthesis to contextualise evidence into a decision. A well conducted systematic review and meta-analysis can often only provide evidence for a small part of the full scope of a decision problem. For example, Carrick et al. (2019) asked the question ‘is planting trees the solution to reducing flood risks?’. Using a systematic review and meta-analysis they could only answer a lesser question about the effect of trees on stream flow. Flood risk is a wider issue and consists of many other factors that can influence the outcome. Using a BN to place the results of the systematic review into the full decision context, Carrick et al. (2019) were able to show the sensitivity of Flood Risk node to uncertainty in the other nodes in the network (Figure 4.13). This, in effect, allows one to quantify which uncertainties should be reduced first to get a better understanding of flood risk. These included, for example, infrastructure density (where roads and houses are placed on the landscape) and the capacity of reservoirs to slow down water flow into river catchments.

Bayesian networks can encompass decision nodes and utility nodes to make them explicit decision support tools (sometimes called Bayesian Decision Networks – see Nyberg et al. 2006). A decision node represents two or more choices that influence the values of outcome nodes. For example, the decision to mow or not to mow a grassland or to irrigate or not irrigate a crop could be included in the network. Utility nodes which show the value of an action in units of importance to the decision maker (this could be financial cost or expected change in the number of species) are often linked to these decision nodes to allow trade-offs between costs and benefits to be explored.

Several commercial packages (with limited free options) exist such as HUGIN, NETICA and GENIE and open source R packages such as bnlearn (Scutari, 2010) and caret (Kuhn, 2022) that can be used to develop BNs. Hybrid networks (combining continuous and categorical data) can

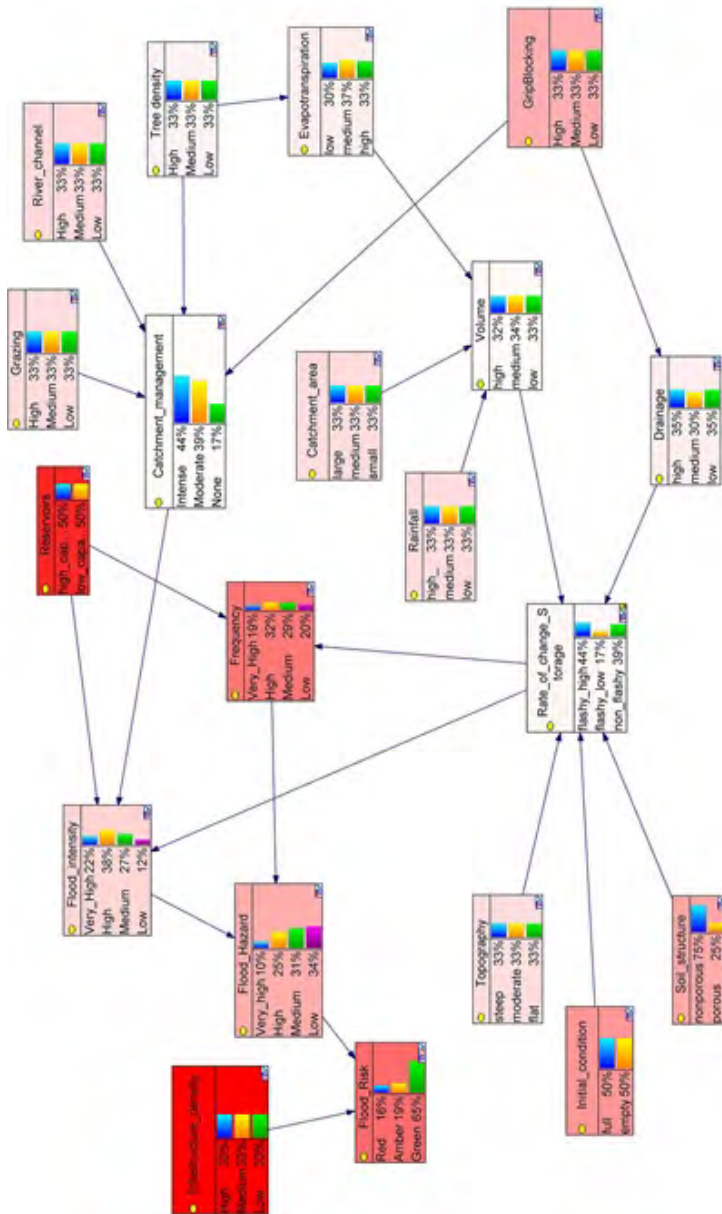


Figure 4.13 Bayesian networks are used to contextualise evidence for decision making. Here, 'Flood Risk' is dependent upon physical factors (e.g. soil structure, rainfall etc.) combined with human "interventions" such as storage dams ('Reservoirs') and changing upland drainage regimes ('Grip Blocking'). An added benefit of using Bayesian networks is that the sensitivity of the focal node ('Flood Risk') to uncertainty in the other nodes in the network (darker shades of red indicate greater importance for reducing uncertainty). Reducing uncertainty in, for example, the 'Reservoirs' node will have a greater impact on our understanding of 'Flood Risk'. Primary research could be targeted at these (darker red) nodes to reduce uncertainty in decisions about 'Flood Risk'. The systematic review and meta-analysis of Carrick et al. (2019) represent a single node in the network ('Tree density') whilst all other nodes come from the literature. (Source: Carrick et al., 2019, © 2018 The Chartered Institution of Water and Environment Management (CIWEM) and John Wiley & Sons Ltd, reproduced with permission from John Wiley & Sons Ltd).

now be developed using JAGs software (Plummer, 2003) through the R package Hydenet (Dalton and Nutter, 2019).

4.7.7 Evidence restatements

Policy decisions often need to be made in areas where the evidence base is both complex and heterogeneous (precluding formal meta-analysis) and highly contested. Evidence 'Restatements' seek to summarise the evidence base in a manner as policy neutral as possible while being clear about uncertainties and evidence gaps. If successful they can clarify the role of economic considerations, value judgements and other considerations in addition to the evidence base in the process of policy-making.

There are different ways to construct restatements and here we describe a model developed by the Oxford Martin School at Oxford University.

1. A topic is chosen after discussion with policy makers.
2. A panel of experts is then convened, typically from universities and research organisations.
3. A draft evidence summary is then prepared which takes the form of a series of numbered paragraphs written to be intelligible to a policy maker who is familiar with the subject but is not a technical expert. The summary is based on a systematic review of the literature and each statement is accompanied by an opinion on the strength of the evidence using a reserved vocabulary that varies between projects.
4. An extensive annotated bibliography is produced, linked to each paragraph, allowing policy makers access to the primary literature should they require further information.
5. The draft restatement is discussed paragraph by paragraph by the expert group and revised over several iterations.
6. This version is then sent out for review to a large group of people including other subject experts but also policy makers and interested parties (for example from the private sector and NGOs). They are asked to comment on the evidence itself, the assessment of the strength of the evidence, and also the aspiration that the restatement is policy neutral.
7. The restatement is revised in the light of these comments and the final version is drawn up after further iterations with the expert panel.
8. The restatement is submitted for publication to a journal (appearing as an appendix to a short 'stub' paper with the annotated bibliography as supplementary material online) where it receives a final review. Restatements are independently funded (by philanthropy).

Topics that have been the subject of restatements include the control of bovine TB in cattle and wildlife (Godfray et al., 2013) and the biological effects of low-dose ionising radiation (McLean

et al., 2017). Other topics covered have included the effects of neonicotinoid insecticides on pollinators, landscape flood management, endocrine-disrupting chemicals, and control of *Campylobacter* infections.

Restatements have been popular with policy makers because (i) they strive to be policy neutral; (ii) they are conducted completely independently of the policy maker (important in areas of great contestation); (iii) they carry the authority of the expert panel and a scientific publication; and (iv) they are relatively brief and written to be understandable by policy makers. Their disadvantages include reliance on expert judgement (the heterogeneous evidence base precluding a more algorithmic approach) and the time and effort required of typically senior experts.

References

- Adams, C.A., Fernández-Juricic, E., Bayne, E.M. et al. 2021. Effects of artificial light on bird movement and distribution: a systematic map. *Environmental Evidence* 10: Article 37, <https://doi.org/10.1186/s13750-021-00246-8>.
- Ausden M., Sutherland W.J. and James R. 2001. The effects of flooding lowland wet grassland on soil macroinvertebrate prey of breeding wading birds. *Journal of Applied Ecology* 38: 320–38, <https://doi.org/10.1046/j.1365-2664.2001.00600.x>.
- Back from the Brink. 2021. *Looking After Spring Speedwell* Veronica verna. Ecology and Conservation Portfolio (Salisbury: Plantlife), https://naturebftb.co.uk/wp-content/uploads/2021/06/Spring_speedwell_FINAL_LORES.pdf.
- Boutin, C. and Harper, J.L. 1991. A comparative study of the population dynamics of five species of Veronica in natural habitats. *Journal of Ecology* 79: 199–221, <https://doi.org/10.2307/2260793>.
- Buzan, T. 1974. *Use Your Head* (London: BBC Books).
- Carrick, J., Abdul Rahim, M.S.A.B., Adjei, C., et al. 2019. Is planting trees the solution to reducing flood risks? *Journal of Flood Risk Management* 12(S2): e12484, <https://doi.org/10.1111/jfr3.12484>.
- Christie, A.P., Amano, T., Martin, P.A., et al. 2020a. The challenge of biased evidence in conservation. *Conservation Biology* 35: 249–62, <https://doi.org/10.1111/cobi.13577>.
- Christie, A.P., Amano, T., Martin, P.A., et al. 2020b. Poor availability of context-specific evidence hampers decision-making in conservation. *Biological Conservation* 248: 108666, <https://doi.org/10.1016/j.biocon.2020.108666>.
- Christie, A.P., Chiaravalloti, R.M., Irvine, R., et al. 2022. Combining and assessing global and local evidence to improve decision-making. OSF, <https://osf.io/wpa4t/>.
- CMP. 2020. Conservation Standards, <https://conservationstandards.org/about/>.
- Cramp, S. and Simmons, K.E.L. 1983. *Handbook of the Birds of Europe, the Middle East and North Africa: The Birds of the Western Palearctic. Vol. III: Waders to Gulls* (Oxford: Oxford University Press).
- Dalton, J.E. and Nutter, B. 2020. HydeNet: Hybrid Bayesian Networks Using R and JAGS. R package version 0.10.11, <https://cran.rproject.org/web/packages/HydeNet/index.html>

- Fedrowitz, K., Koricheva, J., Baker, S.C. 2014. Can retention forestry help conserve biodiversity? A meta-analysis. *Journal of Applied Ecology* 51: 1669–79, <https://doi.org/10.1111/1365-2664.12289>.
- Gigerenzer, G. and Edwards, A. 2003. Simple tools for understanding risks: From innumeracy to insight. *BMJ (Clinical research ed.)* 327: 741–44, <https://doi.org/10.1136/bmj.327.7417.741>.
- Green, R.E. 1986. The management of lowland wet grassland for breeding waders. Unpublished RSPB Report.
- Green, R.E. 1988. Effects of environmental factors on the timing and success of breeding of common snipe *Gallinago gallinago* (Aves: Scolopacidae). *Journal of Applied Ecology* 25: 79–93, <https://doi.org/10.2307/2403611>.
- Green, R.E., Hirons, G.J.M., and Cresswell, B.H. 1990. Foraging habitats of female common snipe *Gallinago gallinago* during the incubation period. *Journal of Applied Ecology* 27: 325–35, <https://doi.org/10.2307/2403589>.
- Godfray, H.C.J., Donnelly, C.A., Kao, R.R., et al. 2013. A restatement of the natural science evidence base relevant to the control of bovine tuberculosis in Great Britain. *Proceedings of the Royal Society B, Biological Sciences* 280: 1634–43, <https://doi.org/10.1098/rspb.2013.1634>.
- Hoodless, A.N., Ewald, J.A., and Baines, D. 2007. Habitat use and diet of common snipe *Gallinago gallinago* breeding on moorland in northern England. *Bird Study* 54: 182–91, <https://doi.org/10.1080/00063650709461474>.
- Holton, N. and Allcorn, R.I. 2006. The effectiveness of opening up rush patches on encouraging breeding common snipe *Gallinago gallinago* at Rogersceugh Farm, Campfield Marsh RSPB reserve, Cumbria, England. *Conservation Evidence* 3: 79–80, <https://conservationevidencejournal.com/reference/pdf/2230>.
- Iacona, G.D., Sutherland, W.J., Mappin, B., et al. 2018. Standardized reporting of the costs of management interventions for biodiversity conservation. *Conservation Biology* 32: 979–88, <https://doi.org/10.1111/cobi.13195>.
- IPCC. 2005. *Guidance Notes for Lead Authors of the IPCC Fourth Assessment Report on Addressing Uncertainties*, July 2005 (Intergovernmental Panel on Climate Change), <https://www.ipcc.ch/site/assets/uploads/2018/02/ar4-uncertaintyguidancenote-1.pdf>.
- Jensen, F.V. 2001. Causal and Bayesian networks. In: *Bayesian Networks and Decision Graphs* (New York: Springer), pp. 3–34.
- Kuhn, M. 2022. Caret: classification and regression training. R package version 6.0-93, <https://CRAN.R-project.org/package=caret>.
- Landuyt, D., Broekx, S., D’hondt, R., et al. 2013. A review of Bayesian belief networks in ecosystem service modelling. *Environmental Modelling & Software* 46: 1–11, <https://doi.org/10.1016/j.envsoft.2013.03.011>.
- Malpas, L.R., Kennerley, R.J., and Hirons, G.J.M., et al. 2013. The use of predator-exclusion fencing as a management tool improves the breeding success of waders on lowland wet grassland. *Journal for Nature Conservation* 21: 37–47, <https://doi.org/10.1016/j.jnc.2012.09.002>.
- Mason, C.F. and MacDonald, S.M. 1976. Aspects of the breeding biology of Snipe. *Bird Study* 23: 33–38, <https://doi.org/10.1080/00063657609476482>.
- McLean A.R., Adlen, E.K., Cardis E., et al. 2017. A restatement of the natural science evidence base concerning the health effects of low-level ionizing radiation. *Proceedings of the Royal Society B* 284, 20171070, <https://doi.org/10.1098/rspb.2017.1070>.

- Nabe-Nielsen, L.I., Reddersen, J., and Nabe-Nielsen, J. 2021. Impacts of soil disturbance on plant diversity in a dry grassland. *Plant Ecology* 222: 1051–63, <https://doi.org/10.1007/s11258-021-01160-2>.
- Nyberg, J. B., Marcot, B.G., and Sulyma, R. 2006. Using Bayesian belief networks in adaptive management. *Canadian Journal of Forest Research* 36: 3104–16, http://wildlifeinfometrics.com/wp-content/uploads/2018/02/Nyberg_et_al_2006_Bayesian_Models_in_AM.pdf.
- Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In: *DSC 2003. Proceedings of the 3rd International Workshop on Distributed Statistical Computing*, ed. by K. Hornik, et al., 124: 1–10, <https://www.r-project.org/nosvn/conferences/DSC-2003/Drafts/Plummer.pdf>.
- Popotnik, G.J. and Giuliano, W.M. 2000. Response of birds to grazing of riparian zones. *Journal of Wildlife Management* 64: 976–82, <https://doi.org/10.2307/3803207>.
- Robson, B. and Allcorn, R.I. 2006. Rush cutting to create nesting patches for lapwings *Vanellus vanellus* and other waders, Lower Lough Erne RSPB reserve, County Fermanagh, Northern Ireland. *Conservation Evidence* 3: 81–83, <https://conservationevidencejournal.com/reference/pdf/2231>.
- Salafsky, N., Boshoven, J., Burivalova, Z., et al. 2019. Defining and using evidence in conservation practice. *Conservation Science and Practice* 1: e27, <https://doi.org/10.1111/csp2.27>.
- Salafsky, N., Irvine, R., Boshoven, J., et al. 2022. A practical approach to assessing existing evidence for specific conservation strategies. *Conservation Science and Practice* 4: e12654, <https://doi.org/10.1111/csp2.12654>.
- Scutari, M. 2010. Learning Bayesian networks with the bnlearn R package. *arXiv Preprint*, <https://doi.org/10.48550/arXiv.0908.3817>.
- Smart, J., Amar, A., O'Brien, M., et al. 2008. Changing land management of lowland wet grasslands of the UK: Impacts on snipe abundance and habitat quality. *Animal Conservation* 11: 339–51, <https://doi.org/10.1111/j.1469-1795.2008.00189.x>.
- Spiegelhalter, D.J., Abrams, K.R., and Myles, J.P. 2004. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*, Vol. XIII (Chichester: John Wiley & Sons), <https://doi.org/10.1002/0470092602>.
- Stewart, G. and Ward, J. 2019. Meta-science urgently needed across the environmental nexus: a comment on Berger-Tal et al. *Behavioural Ecology* 30: 9–10, <https://doi.org/10.1093/beheco/ary155>.
- Stephenson, F., Mill, A.C., Scott, C.L., et al. 2018. Socio-economic, technological and environmental drivers of spatio-temporal changes in fishing pressure. *Marine Policy* 88: 189–203, <https://doi.org/10.1016/j.marpol.2017.11.029>.
- Stewart, G.B., Mengersen, K., and Meader, N. 2014. Potential uses of Bayesian networks as tools for synthesis of systematic reviews of complex interventions. *Research Synthesis Methods* 5: 1–12, <https://doi.org/10.1002/jrsm.1087>.
- Watt, A.S. 1971. Rare species in Breckland; their management for survival. *Journal of Applied Ecology* 8: 593–609, <https://doi.org/10.2307/2402895>.
- White, T., Petrovan, S., Booth, H., et al. 2022. Determining the economic costs and benefits of conservation actions: A framework to support decision making. *Conservation Science & Practice*, <https://doi.org/10.31219/osf.io/kd83v>.
- Wyden, P. 1980. *The Bay of Pigs: The Untold Story* (New York: Simon & Schuster).