Eliciting Expert Knowledge in Conservation Science — Source link

Eliciting Expert Knowledge in Conservation Science

Tara G. Martin[1,2*], Mark A. Burgman[3], Fiona Fidler[3], Petra M. Kuhnert[4], Samantha Low-Choy[5,6], Marissa McBride[3] and Kerrie Mengersen[6]

[1]CSIRO Ecosystem Sciences, Ecoscience Precinct, GPO Box 2583 Brisbane, Queensland 4001 Australia; Tara.martin@csiro.au

[2]ARC Centre of Excellence for Environmental Decisions, University of Queensland, Queensland 4072 Australia

[3]Australian Centre of Excellence for Risk Analysis, School of Botany, University of Melbourne, Parkville, Victoria 3010 Australia

[4]CSIRO Mathematics, Informatics and Statistics, Private Bag 2, Glen Osmond, SA 5064 Australia

[5]Cooperative Research Centre in National Plant Biosecurity, Canberra Australia

[6]Faculty of Science and Technology, Queensland University of Technology, GPO Box 2434 Brisbane, Queensland 4001 Australia

*To whom correspondence should be addressed: Tara.Martin@csiro.au; Ph + 61 7 3833 5727

Word count: 7537

Running head: Elicitation of expert knowledge

Keywords: bias, decision making, expert judgement, expert opinion, elicitation, overconfidence, Bayesian priors

Abstract

Expert knowledge is used widely in the science and practice of conservation because of the relative lack of data and the imminent nature of many conservation decisions. Expert knowledge is substantive information on a particular topic that is not widely known by others. An expert is someone who holds this knowledge and who should be deferred to in its interpretation. When experts use their knowledge to predict what may happen in a particular context, we refer to these predictions as expert judgements, since what will happen is not known for certain. In general an expert-elicitation approach for use in conservation science consists of five steps: decide how information will be used, determining what to elicit, designing the elicitation process, performing the elicitation, and translating the elicited information into quantitative statements that can be used in a model or decision directly. This last step is known as encoding. Some of the considerations in eliciting expert knowledge include determining how to work with multiple experts and combine multiple judgements, minimizing bias in the elicited information, and verifying the accuracy of expert information. We highlight structured elicitation techniques, that if adopted, will improve the accuracy and information content of expert judgement and ensure uncertainty is captured appropriately. Four criteria; study design and context, elicitation design, elicitation method and elicitation output, can be used to assess the comprehensiveness and effectiveness of an elicitation exercise. Just as the reliability of empirical data depends on the rigor with which it was acquired so too does that of expert knowledge.

keep revised abstract to 1 page

Introduction

The growing use of expert knowledge in conservation science is driven by the need to characterize dynamic, complex systems, limited resources to collect new empirical data, and the urgency of conservation decisions (Sutherland 2006; Kuhnert et al. 2010). The utility of expert knowledge depends on the scientific rigor with which it is acquired and its accuracy. Just as observational data and the methods used to collect it are subject to scrutiny, so too should expert knowledge be scrutinized to ensure that uncertainty is quantified and bias in the elicited information is minimized (O'Hagan et al. 2006).

In this review, we defined what expert knowledge is and who qualifies as an expert. We examined how expert knowledge is being used to inform conservation science and practice. We outlined an elicitation approach that consists of five steps; deciding how information will be used, determining what to elicit, designing the elicitation process, performing the elicitation, and encoding the elicited information for use in a model or to inform a decision directly. We focussed on the elicitation of quantities such as population sizes, the likelihood of extinction of a population or the prevalence of a species or pest. We discussed ways of minimising bias, combining multiple judgements, dealing with uncertainty, and increasing the accuracy of elicited information. Finally we outlined criteria for assessing how comprehensive and informative an elicitation exercise has been.

Definition of expert knowledge

Expert knowledge is substantive information on a particular topic that is not widely known by others. An expert is generally considered someone who holds information about a given topic and who should be deferred to in its interpretation (Barley & Kunda 2006). This knowledge may be the result of training, research, and skills, but could also be based on personal experience (Burgman et al. 2011a). When experts use their knowledge to predict what may happen in a particular context, we refer to these predictions as expert judgements, since what will happen is not known for certain. Experts exist, are unequally distributed among the human population, and are not created only through formal education or professional experience (Evans 2008). There are different types of expertise: substantive, which reflects the expert's knowledge of their domain; normative, which is the expert's ability to accurately and clearly communicate their judgements in a particular format, such as probabilities; and adaptive, which describes the degree to which experts are able to extrapolate or adapt to new circumstances (McBride & Burgman 2011).

These types of expertise may be unrelated, but they are all integral to the effective use of expert information.

The quality of expert judgements is reflected in the calibration and informativeness of the judgements (Cooke 1991; O'Hagan et al. 2006). Calibration of a judgement indicates how closely a judgement corresponds to reality (e.g., the amount of agreement between an expert's judgement in the form of, for example, probabilities and what is observed in reality) (O'Hagan et al. 2006). The informativeness of an expert's judgement is reflected in the precision and confidence (e.g., uncertainty of an estimate). Calibration of judgements occurs through observation of the outcomes of predictions or through formal evaluation of an expert's knowledge (tests) or use of knowledge in scenario analyses (Cooke 1991; Burgman et al. 2011b).

Value of expert knowledge

Data on many conservation problems are typically scarce; nevertheless, management decisions must be made (Cook et al. 2009). Expert judgements can provide information about model parameters and help characterize uncertainty in models, the intent of which often is to confront these judgments with data as it becomes available. Where decisions are required urgently, the expert judgments may be the basis for the decision, without additional empirical evidence.

Expert judgement is commonly used to develop and evaluate projects at the stages of hypothesis generation, sample design, model development, and interpretation of results (Fazey et al. 2005; Runge et al. 2011; Sutherland 2011). Expert judgments may be the only, or the most, credible source of information available for making management decisions (Martin et al. 2005; Johnson et al. 2010b), for modeling species distributions (Langhammer 2007), for assessing the risk of colonization or expansion of non-native species (Kuhnert 2011), and for threat- management decision analyses (Joseph et al. 2009; Carwardine et al. 2011). Despite this, there remains ongoing controversy about use of expert judgements (Ludwig et al. 2001; Kuhnert 2011). Reservations reflect concerns that expert judgements may be biased, poorly calibrated, or self-serving (Tversky & Kahneman 1974; Krinitzsky 1993) and thus lead to poor inference and decision making. Substantial research has concentrated on methods to overcome these problems (Kynn 2008).

Conventional definitions of expertise that depend on qualifications and experience may not correspond to the reliability of an expert's judgements (Cooke & Goossens 2008). To provide accurate judgements requires what psychologists call *deliberate practice* (Ericsson 1996), which involves the structured repetition of tasks with immediate and unambiguous feedback about accuracy. In a wide range of domains, a minimum of 10,000 hours of deliberate practice, especially feedback, is required to reach expert performance (Ericsson 1996). Few experts reach highest levels of competence in less than a decade. We speculate that the conditions of many conservation projects rarely provide the opportunity for deliberate practice by individuals. In our view, the prominence of expert judgement in conservation and the potential for its misuse create an imperative to adopt explicit structured and robust procedures to gather expert judgements. These procedures include recording elicitation design and protocols, verifying expert accuracy independently, and training experts by providing them with feedback on their judgements.

Use of expert knowledge in conservation

Examples of the use of expert knowledge to inform decision making can be found in almost all areas of conservation science and practice, and we highlight only a few here. Expert knowledge has been used to inform different types of models used to characterize relations between species and abiotic or biotic covariates: generalized linear regression models (GLMs), classification trees, and Bayesian networks. For example, Smith et al. (2007) used expert judgements about habitat in a Bayesian network for the Julia Creek dunnart (*Sminthopsis douglasi*). For a model of brush-tailed rock wallaby (*Petrogale penicillatus*) distribution, expert knowledge was used to fill information gaps associated with species occupancy in inaccessible sites (Murray et al. 2009; O'Leary et al. 2009).

Population management often depends on expert judgements of population sizes, life-history parameters, and responses of populations to management. Johnson et al. (2010b) used expert judgements to construct a Bayesian network to evaluate the viability of cheetahs (*Acinonyx jubatus*) in southern Africa following translocation. Runge et al. (2011) elicited expert judgements on the responses of endangered (U.S. Endangered Species Act) Whooping Cranes

(*Grus americana*) to management to evaluate what uncertainty, if known, would lead to a different management decision. O'Neill et al. (2008) quantified the trends and variance of possible effects of climate change on polar bears (*Ursus maritimus*) on the basis of judgements of 10 experts. Martin et al. (2005) and Kuhnert et al. (2005) investigated the effect of different intensities of livestock grazing on Australian woodland birds with a Bayesian GLM that was built with information from 20 ecologists . In many cases, the inclusion of expert information substantially improved the power of inferences, whereas empirical data alone had insufficient power.

Sugiura and Murray (2011) noted that assessments of risk of colonization and expansion of non-native invasive species, used data and expert knowledge from many disciplines. Quantitative approaches to invasive species management often rely on expert knowledge to quantify input parameters such as detection probability, prevalence, and risk of establishment (Kuhnert 2011; Low-Choy et al. 2011b) or to specify conditional probabilities in Bayesian networks (Johnson et al. 2010a; Smith et al. 2011). Hayes (2002a, b) and Hayes et al. (2004) presented fault-tree analyses, infection models, and hierarchical holographic models to groups of experts to identify potential undesirable effects associated with the release of various organisms.

Expert knowledge also has been used in models of managed systems for which many relevant parameters (e.g., optimal harvesting levels, effects of harvest, future demand) cannot be assessed directly (Crome et al. 1996; Marcot 2006; Griffiths et al. 2007; Rothlisberger et al. 2010).

Eliciting expert information

Typically, an expert-judgement elicitation team includes the problem owner (person who specifies the problem), facilitator, analyst, and one or more experts. One person, in theory, could have several roles. Generally, definition of the problem and selection of experts is the domain of the problem owner. The facilitator manages the interactions among experts and oversees the judgement-elicitation process, and the analyst handles calibration, elicitation procedures, processing of responses, and analysis of elicited information.

A general elicitation approach includes five steps: deciding how information will be used, determining what to elicit, designing the process of eliciting judgements, performing the elicitation, and translating the elicited information for use in a model, otherwise known as encoding the elicited information. Examples of the elicitation approach used in conservation projects are provided in Supporting Information.

Deciding how expert knowledge will be used

Before undertaking expert elicitation, the problem owner, facilitator, and analyst should have a clear understanding of how the elicited information will be used. Will the elicited judgements be incorporated into a model or form the basis of a decision directly? Bayesian models are particularly useful for incorporating expert judgements because they provide a formal mechanism to include judgements through prior probability distributions of model parameters (Gelman et al. 2004; Kuhnert et al. 2010; Low-Choy 2011). These so-called subjective priors reflect the judgement held by an expert concerning a particular model parameter. Frequentist approaches have also been developed to incorporate expert judgements into GLMs (Lele & Allen 2006).

Determining what to elicit

To identify the variables about which to elicit information, we suggest considering which variables most strongly affect the decision or predictions to be made. Determine what lack of knowledge surrounding parameters impedes making inferences or decisions, and focus judgement elicitation on these parameters and their uncertainty. Elicitation should reveal relevant information about these parameters and their uncertainty, and the format in which questions are posed to the experts should allow experts to express their knowledge easily. To resolve language-based misunderstandings and different interpretations of the decisions or predictions to be made, most elicitation exercises commence with discussion of the questions themselves. This discussion is aided by an awareness of the relevant types of uncertainty (Regan et al. 2002).

Designing the elicitation process

There are several ways to elicit expert knowledge (e.g. Morgan & Henrion 1990; Cooke 1991). Recent publications in the conservation science and ecological literature detail the application of these methods (Burgman 2005; Low-Choy et al. 2009; Kuhnert et al. 2010; McBride & Burgman 2011). We synthesized the common details of the process described in these publications.

During the design phase the steps in the elicitation process are delineated and how to manage bias is established. The elicitation format (e.g., email survey, telephone interview, face-to-face interview, group meeting) is determined. experts are identified; background materials are compiled (e.g., reports, journal articles, datasets), questions are tested and finalised, and scenarios to help the experts understand the questions are developed; logistics of acquisition of and interactions with experts are determined; methods of analysis of the expert data, including methods to address uncertainty, are determined; and roles of the elicitation team are identified (team members described above) (McBride & Burgman 2011).

The design process includes expert training. Training may involve having experts answer practice questions and develop familiarity with the elicitation style and procedure. For example, if probabilistic information is to be elicited, experts could be asked to estimate probabilities or frequencies through a variety of methods, including natural-frequency formats, cumulative-density functions, or probability wheels (Morgan & Henrion 1990; Caponecchia 2009; James et al. 2010; Kuhnert et al. 2010). This phase is particularly useful when detailed information is to be elicited in a format the expert may be unfamiliar with, such as probability distributions and their statistical summaries.

Performing the elicitation

Information may be elicited directly or indirectly (Low-Choy et al. 2009; Kuhnert et al. 2010). Direct elicitation requires the expert to express the knowledge in terms of the quantities required by the analyst. For example, the expert may be asked to provide statistical summaries (e.g., a lower and upper bound and a best estimate), quantiles or cumulative probabilities, or a full parametric probability distribution (see Jose et al. 2009; Kuhnert et al. 2010; Low-Choy et al. 2011a). Indirect elicitation requires experts to answer questions that relate to their experiences. Their responses are then encoded into the quantities required by the analyst. For example, the expert may be asked about expected site occupancy given different habitats, which the analyst then translates into an appropriate probability distribution for a model parameter.

If the elicitation process involves multiple experts, information can be elicited independently and then combined, or a group opinion can be sought. Common group approaches include expert panels and Delphi methods (e.g., Crance 1987; MacMillan & Marshall 2006). Although expert panels foster pooling of knowledge among experts and agreement on the problem and questions at hand, the full diversity of opinions are lost and responses are subject to biases, including dominance of one or more members of the group, polarization among subsets of members, and groupthink, a mode of thinking that occurs when the desire for harmony in a decision-making group overrides a realistic appraisal of alternatives (Janis 1972; Kerr & Tindale 2011). To overcome these limitations, structured interactions such as the Delphi method elicit individual estimates from experts and then allow each expert to adjust their estimates in light of the responses of others while maintaining anonymity (Linstone & Turoff 1975). In a variant of this method experts make initial individual estimates, discuss their responses, and then make a final, individual estimate. This procedure generates group estimates for ecological parameters that usually are more accurate than the estimates of the best-regarded expert in a group (Burgman et al. 2011b).

Another approach to expert elicitation that is gaining recognition is Cooke's method (Cooke 1991; Cooke & Goossens 2004; Cooke et al. 2008; Goossens et al. 2008; Aspinall 2010). In this method the opinion of each expert is weighted on the basis of its accuracy. Experts are brought together to discuss a particular topic (e.g., migration arrival time of particular species) under the guidance of a facilitator. Following group discussion, experts are asked individually to give their judgement (e.g., migration arrival time of particular species over last 5 years). To weight each expert based in their accuracy, each expert is also asked a set of test questions for which the answers are known. Accuracy of answers to the test questions is used to weight their judgement, and the weighted judgement of all experts are pooled to provide a consensus judgement (Cooke 1991; Aspinall 2010). The challenge is to identify test questions with known responses that are closely related to the questions for which answers are unknown.

Encoding the elicited information

Encoding is the process of translating information that has been elicited indirectly into quantitative statements that can be used in a model. For example, Martin et al. (2005) used indirect elicitation to assess the effects of livestock-grazing practices on native birds. For a given level of grazing, experts assessed whether a species' relative abundance would increase (score +1), decrease (–1), or exhibit no change (0). The mean and variance of all expert assessments, by bird species and grazing level, were encoded as priors in a Bayesian GLM. In another indirect approach, (Low-Choy et al. 2010) asked experts to consider sites with different characteristics and estimate the probability a site would be occupied by an endangered mammal.

Selection of elicitation formats and techniques depends on the number and types of experts, accuracy required, and time and resources available to conduct the elicitation. Additionally, there is a trade-off between the number of judgements that can be elicited with accuracy and the need to retain experts' attention throughout the process and to complete the elicitation efficiently (Shephard & Kirkwood 1994).

Software can be used to automate and manage computational tasks, help experts express quantities, provide immediate feedback to experts about the elicited values, and encode elicited information. Packages designed to facilitate the elicitation of expert knowledge include SHELF

(Oakley & O'Hagan 2010), Elicitator (James et al. 2010), Excalibur (Goossens et al. 2008), and ET (Speirs-Bridge et al. 2010).

Additional considerations

Interplay between expert and empirical information

Expert knowledge is often portrayed as subjective and is contrasted with objective empirical data. However, empirical data may reflect biases, inadequacies, and errors in study design, collection, and transcription. We believe expert knowledge and empirical data exist on a continuum of subjectivity and, depending on the particular case, one may be a better proxy for the truth. Both expert knowledge and empirical data require validation. Expert knowledge should be regarded only as a snapshot of the expert's judgements in time, and expert assumptions and reasoning should be documented in such a way that they can be updated as new empirical knowledge comes to light.

In our experience, Bayesian methods best accommodate updating judgements in light of new empirical information because they broadly define subjective probability. Prior information from either empirical data or expert knowledge can be incorporated into Bayesian analyses by specifying appropriate prior probabilities for parameters (McCarthy 2007). Bayesian methods are being used increasingly to augment empirical data with priors elicited from experts and vice versa (Kuhnert et al. 2005; Martin et al. 2005; Murray et al. 2009). Thus, the methods provide a natural platform for learning and managing adaptively (Keith et al. 2011; McDonald-Madden et al. 2011).

Combining expert judgements

Multiple expert judgements can be combined mathematically with either opinion pooling (most common is equal-weighted linear opinion pool [i.e., group average]) or Bayesian approaches, which can incorporate dependencies between experts (expert judgements that vary as a function of others' judgements) (Clemen & Winkler 1999; O'Hagan et al. 2006). The equal-weighted group average is simple and delivers accurate judgements compared with more complex methods (Armstrong 2001). If there are considerable measurable differences in the accuracy of expert judgements, then use of unequal expert weights in opinion pooling or Bayesian approaches will improve estimation (Cooke 1991; Soll & Larrick 2009; Aspinall 2010).

Although generating an expert consensus may be important for modeling and decision-making, it is important that differences in judgement be retained and communicated to decision makers (Keith 1996; Morgan et al. 2001). In many cases, the considerable benefit of enlisting multiple experts lies in the additional questioning of reasoning and assumptions that arises when examining differences in expert judgements (Morgan & Henrion 1990).

Accounting for bias

Humans are susceptible to a range of subjective and psychological biases (overview in Supporting Information), often unknowingly (Slovic 1999; Kynn 2008; McBride & Burgman 2011). Motivational biases arise from the context of the expert, personal beliefs, and from the personal stake one might have in a decision. Accessibility biases arise when information that comes more easily to the mind of an expert exerts a disproportionate influence on an expert's judgements. Anchoring and adjustment biases occur when an expert anchors an estimate on a benchmark and then is unable to adjust this estimate much above or below the benchmark.

Overconfidence bias arises when the confidence of experts in their judgements is higher than is warranted by the accuracy of their estimates (McKenzie et al. 2008). This bias sometimes results in systematic underestimation, in which experts fail to express the extent of uncertainty (O'Hagan et al. 2006).

Although it is important to be aware of the potential for bias, not all experts in all elicitation processes will be biased. Forty years after Tversky and Kanheman's (1974) seminal work, much more is known about the conditions that exacerbate or minimize cognitive biases. In particular, the following may mitigate bias: set tasks that allow for deliberate practice, including unambiguous feedback; compose questions posed to experts in such that they are aligned with an expert's knowledge. Several authors provide more extensive advice on managing elicitation bias (Meyer & Booker 1991; O'Hagan et al. 2006; Kynn 2008; Low-Choy et al. 2009).

Some biases, such as overconfidence, are more resistant to mitigation (Moore & Healy 2008). Overconfidence may increase as availability of information increases (Oskamp 1965; Tsai et al. 2008) and in the absence of regular systematic feedback (Lichtenstein et al. 1982; Dawes 1994). Overconfidence has also been shown to be high when the predictability of the future becomes low (Lichtenstein & Fischhoff 1977; Griffin & Tversky 1992). Unfortunately, this is when we are likely to need expert judgement the most. Overconfidence may also be influenced by the expert's "cognitive style" (Tetlock 2005). Suggested remedies for overconfidence come from information-sampling theory (e.g. Klayman et al. 2006) and include asking the same question more than once or with alternative wording.

Building on the work of Soll and Klayman (2004), Speirs-Bridge et al. (2010) developed a four-step procedure to mitigate overconfidence that elicits a lower bound, upper bound, best estimate, and a level of confidence that the true estimate lies within the nominated lower and upper bounds. For example, to estimate the mean number of native bird species in a particular land-management scenario, one would ask the following: Realistically, what do you think could be the lowest mean number of species? Realistically, what do you think could be the highest mean number of species? What is your best estimate of the mean number of species? For the interval created (lower and upper bound), what is the probability between 0-100% that the mean number of species observed in the study will fall within this interval? The first three steps require the expert to produce an interval, whereas the last step requires the expert to evaluate an interval. The addition of this last step takes advantage of the fact that experts are much better at evaluating intervals than producing intervals (Teigen & Jorgensen 2005; Speirs-Bridge et al. 2010).

Dealing with uncertainty

Expert elicitation is used to capture an expert's best estimate and the uncertainty around this estimate. Eliciting the uncertainty around an estimate may lead to different responses depending on the way in which the question is asked. For this reason, it is useful to distinguish between epistemic (knowledge) uncertainty and natural (aleatory) uncertainty (Regan et al. 2002). The former can be reduced by studying the system and acquiring additional knowledge. The latter can be better understood, but not reduced, by collecting additional data. For example, consider a question about the juvenile dispersal rate of a small mammal: What is the average proportion of juvenile males that disperse from a particular patch each year, and what are the 5^{th} and 95^{th} quantiles for the proportion? An expert may estimate the range of variation expected from year to year, may estimate her or his personal uncertainty about the average proportion that disperse over all years, or may include both elements in the estimate. The question should be posed so as to clarify which elements of uncertainty are sought and to partition them into separate questions.

The question, revised to control epistemic uncertainty, could be broken into two parts: 1) what is the average proportion of juvenile males that disperse from this patch? 2) what are the bounds on the estimate such that you are 90% certain the interval includes the true mean dispersal proportion, averaged over all years? For aleatory (natural variation) uncertainty, questions should allow encoding of variation and skew of the distribution of dispersal proportions from year to year. For example, the question might be Given a true mean dispersal rate equal to the rate you have just estimated, by how much do you expect the proportion to deviate from the underlying true mean, from year to year?

It is not always possible to separate epistemic and aleatory uncertainty in an elicitation. However, the risk of failing to consider these sources of uncertainty is that experts may confound them, and it is not generally possible for an analyst to partition them in retrospect. Questions almost always involve language-based misunderstandings. Pilot elicitations, particularly discussion among expert participants, can often resolve most instances of vagueness, ambiguity, context dependence, and underspecificity (Regan et al. 2002; Burgman 2005) that emerge when questions are first tested.

Accuracy of elicited information

A fundamental question in expert elicitation is how to evaluate the accuracy of the elicited information. The judgements elicited from experts can be viewed as accurate if the expert judgements correspond with the truth. But accuracy can also reflect how well the elicited information corresponds to the experts' true belief (O'Hagen et al 2006). Poor accuracy of an expert's judgements could have very different causes. For example, an expert may hold judgements that are well calibrated to the truth, but may fail to express these judgements accurately. In this case the poor accuracy is a result of poor elicitation. Conversely, an expert may express his or her judgements accurately, but those judgements correspond poorly with the truth. In this case, poor accuracy is due to inaccurate knowledge (O'Hagen et al 2006). Only through the use of calibration and feedback can these sources of inaccuracy be separated. In general, consistent bias across a range of experts and knowledge areas indicates poor elicitation (O'Hagen et al 2006).

The future of expert elicitation

The benefits of incorporating expert knowledge in decision making are real and established. Despite the potential of expert knowledge to contribute to decision making, however, formal methods for eliciting and combining judgements only recently have been adopted and tested for application to conservation science and practice. In our experience, it is often not possible to elicit the required quantities directly; hence, we focus our own research on indirect elicitation techniques (Martin et al. 2005; Low-Choy et al. 2010). Indirect elicitation reduces the cognitive burden on the expert because questions target what the expert knows. With this approach, the amount of work required from the analyst is greater because sophisticated statistical modeling is required.

Methods of expert elicitation is a growing domain in conservation science, and many issues remain open for debate and research, for example, the number of experts needed, identification of experts, validity of aggregation methods for combining judgements, assessment of reliability of experts, need for training and feedback, and independent verification of the accuracy of expert judgements with test questions.

We have identified some commonalities in expert elicitation procedures and general suggestions that will improve elicitation methods, including developing a structured procedure that matches the questions to be posed to the export to what experts know, encoding the elicited information to fit the modelling framework, mitigating the most pervasive and predictable cognitive biases, encouraging experts to make independent assessments, and eliciting uncertainties together with best estimates. We suggest the problem owner and analyst anticipate and work to minimize overconfidence and frame questions to suit the experts' experience, skills, and limitations. The potential that questions will be understood should be tested and questions revised accordingly to clarify meaning. When possible, it is beneficial to clearly distinguish between different types of uncertainty when eliciting bounds on estimates. When eliciting information from multiple experts, identify a method of weighting and combining different judgements.

Finally, it is crucial to provide feedback to experts throughout the elicitation process to ensure expert knowledge is captured accurately. One can ensure expert knowledge is calibrated by providing feedback on the basis of responses to questions for which empirical answers exist. The accuracy of expert knowledge can be addressed by managing over- and underconfidence (and other biases) and by designing elicitation and encoding to target expert knowledge more effectively.

Low-Choy et al. (2011b) devised a checklist of attributes for assessing how comprehensive and effective an elicitation process has been. They based the checklist on four criteria:(1) study context and justification (including study location and topic; singularity of expert knowledge (expert knowledge supplemented, complemented, or sole source of information); (2) elicitation design (number of experts invited and that participated; expert category [sage (holder of highest level of expertise), practitioner, scientist, or stakeholder]; elicitation process piloted with test subjects or reviewed by an elicitation specialist; training provided to standardize terms and mitigate biases); (3) elicitation method (knowledge elicited individually, in groups, or both; knowledge elicited in person, remotely, or both; expert metadata collected; an objective was consensus; and elicited information was qualitative, quantitative, or both); and (4) elicitation output (expert metadata used to weight or interpret results; representation of uncertainty in final output; validation of the experts' knowledge and validation with independent data or expert review). These criteria may motivate ongoing scientific investigation, validation and use of expert knowledge in conservation science and practice.

Given limited resources, complexity of conservation problems and imminent nature of decisions, expert knowledge will continue to play a pivotal role in informing models and decisions. We take the position that independent validation of expert judgements is essential because there is a potential for motivational biases and the propensity for highly divisive conservation issues to influence judgements. We hope that the rigor applied to the elicitation and use of expert knowledge will be the same as is applied to the collection and use of empirical data to ensure the validity of expert knowledge in informing future conservation science and decisions.

Supporting Information

Steps in the expert elicitation process illustrated with examples used in conservation science (Appendix S1) and a summary of subjective biases encountered in expert elicitation (Appendix S2) are available online. The authors are solely responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

Acknowledgments

We are grateful to E. Fleishman for her encouragement to write this paper and her and E. Main's insightful and skilful editing. We are also grateful to W. Aspinall, M. Runge, K. Hayes, C. Cook, and an anonymous reviewer for their thoughtful and thorough reviews. T.G.M. and P.M.K. acknowledge the support of a CSIRO (Commonwealth Scientific and Industry Research Organisation) Julius Career Award, and M.B. was supported by ACERA (Australian Centre of Excellence in Risk Analysis) Project 0611 and a National Science Foundation Award (SES 0725025). T.G.M and M.B acknowledge the support of the National Environmental Research Program, Research Hub for Environmental Decisions. S.L.C. and K.M. gratefully acknowledge Commonwealth of Australia funding for the CRC (Cooperative Research Centre) program.

EM to do Lit. Cited later

Literature Cited

Armstrong, J. S. 2001. Combining forecasts. Pages 417-440 in J. S. Armstrong, editor. Principles of forecasting: A handbook for researchers and practitioners. Kluwer Academic Publishers, Norwell.

Aspinall, W. 2010. A route to more tractable expert advice. Nature **463**:294-295.

Barley, S. R., and G. Kunda. 2006. Contracting: a new form of professional practice. Academy of Management Perspectives **20**:45-66.

Burgman, M., A. Carr, L. Godden, R. Gregory, M. McBride, L. Flander, and L. Maguire. 2011a. Redefining expertise and improving ecological judgment. Conservation Letters **4**:81-87.

Burgman, M. A. 2005. Risks and decisions for conservation and environmental management. Cambridge University Press, UK, Cambridge

Burgman, M. A., M. F. McBride, R. Ashton, A. Speirs-Bridge, L. Flander, B. Wintle, F. Fidler, L. Rumpff, and C. Twardy. 2011b. Expert status and performance. PLoS One **6**:e22998.

Caponecchia, C. 2009. Strategies to improve the communication of probability information in risk analyses. International Journal of Risk Assessment and Management **12**:380-395.

Carwardine, J., T. O'Connor, S. Legge, B. Mackey, H. P. Possingham, and T. G. Martin. 2011. Priority threat management to protect Kimberley wildlife CSIRO Ecosystem Sciences, Brisbane.

Cook, C. N., M. Hockings, and R. W. Carter. 2009. Conservation in the dark? The information used to support management decisions. Frontiers in Ecology and the Environment **8**:181-186.

Cooke, R. M. 1991. Experts in uncertainty: opinion and subjective probability in science. Oxford University Press, New York.

Cooke, R. M., S. ElSaadany, and X. Huang. 2008. On the performance of social network and likelihood based expert weighting schemes Journal of Reliability Engineering & System Safety **93**:745-756.

Cooke, R. M., and L. H. J. Goossens. 2004. Expert judgement elicitation for risk assessments of critical infrastructures. Journal of Risk Research **7**:643-656.

Cooke, R. M., and L. H. J. Goossens. 2008. TU Delft Expert Judgment Data Base. Reliability Engineering & System Safety **93**:657-674.

Crance, J. H. 1987. Guidelines for using the Delphi Technique to Develop Habitat Suitability Index Curves. Page 21. Biological Report 82(10.134). U. S. Fish Wildlife Service.

Crome, F. H. J., M. R. Thomas, and L. A. Moore. 1996. A novel Bayesian approach to assessing impacts of rain forest logging. Ecological Applications **6**:1104-1123.

Dawes, R. M. 1994. House of cards: Psychology and psychotherapy built on myth. Free Press, New York.

Ericsson, K. A. 1996. The acquisition of expert performance : an introduction to some of the issues. Pages 1-50 in K. A. Ericsson, editor. The road to excellence: the acquisition of expert performance in the arts and sciences, sports, and games, Erlbaum, Mahwah, NJ.

Evans, R. 2008. The sociology of expertise: the distribution of social fluency. Sociology Compass **2**:281-298.

Fazey, I., J. A. Fazey, and D. M. A. Fazey. 2005. Learning more effectively from experience. Ecology and Society **10(2): 4**:[online] URL: http://www.ecologyandsociety.org/vol10/iss12/art14/.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin 2004. Bayesian Data Analysis. Chapman and Hall/CRC, USA.

Goossens, L. H. J., R. M. Cooke, A. R. Hale, and L. Rodic-Wiersma. 2008. Fifteen years of expert judgement at TUDelft. Safety Science **46**:234-244.

Griffin, D., and A. Tversky. 1992. The weighing of evidence and the determinants of confidence. Cognitive Psychology **24**:411-435.

Griffiths, S. P., P. M. Kuhnert, W. N. Venables, and S. J. M. Blaber. 2007. Estimating abundance of pelagic fishes using gillnet catch data in data-limited fisheries: a Bayesian approach. Canadian Journal of Fisheries and Aquatic Sciences **64**:1019-1033.

Hayes, K. R. 2002a. Identifying hazards in complex ecological systems. Part 1: Fault-tree analysis for biological invasions. Biological Invasions **4**:235-249.

Hayes, K. R. 2002b. Identifying hazards in complex ecological systems. Part 2: Infection models and effects analysis for biological invasions. Biological Invasions **4**:251-261.

Hayes, K. R., Gregg, P.C., Gupta, V.V.S.R., Jessop, R., Lonsdale, W.M., Sindel, B., Stanley, J., Williams, C.K. 2004. Identifying hazards in complex ecological systems. Part 3: Hierarchical holographic model for herbicide tolerant oilseed rape. Environmental Biosafety Research **3**:109-128.

James, A., S. L. Choy, and K. Mengersen. 2010. Elicitator: an expert elicitation tool for regression in ecology. Environmental Modelling & Software **25**:129-145.

Janis, I. 1972. Victims of Groupthink. Houghton Mifflin, Boston.

Johnson, S., G. Hamilton, F. Fielding, and K. Mengersen. 2010a. An integrated Bayesian Network approach to Lyngbya majuscule bloom initiation. Marine Environmental Research **69**:27-37.

Johnson, S., K. Mengersen, A. de Waal, K. Marnewick, D. Cilliers, A. Houser, and L. Boast. 2010b. Modelling cheetah relocation success in southern Africa using an Iterative Bayesian Network Development Cycle. Ecological Modelling **221**:641-651.

Jose, V. R. R., R. F. Nau, and R. L. Winkler. 2009. Sensitivity to distance and baseline distributions in forecast evaluation. Management Science **55**:582-590.

Joseph, J. N., R. F. Maloney, and H. P. Possingham. 2009. Optimal Allocation of Resources among Threatened Species: a Project Prioritization Protocol. Conservation Biology **23**:328-338.

Keith, D. A., T. G. Martin, E. McDonald-Madden, and C. Walters. 2011. Uncertainty and adaptive management for biodiversity conservation. Biological Conservation **144**:1175-1178.

Keith, D. W. 1996. When is it appropriate to combine expert judgments? Climatic Change **33**:139-143.

Kerr, N. L., and R. S. Tindale. 2011. Group-based forecasting? A social psychological analysis. International Journal of Forecasting **27**.

Klayman, J., J. B. Soll, P. Juslin, and A. Winman. 2006. Subjective Confidence and the Sampling of Knowledge. Pages 153-182 in K. Fiedler, and P. Juslin, editors. Information Sampling and Adaptive Cognition. University of Cambridge Press, Cambridge, U.K.

Krinitzsky, E. L. 1993. Earthquake probability in engineering—part 1: the use and misuse of expert opinion. Engineering Geology **33**:257-288.

Kuhnert, P. M. 2011. Four case studies in using expert opinion to inform priors, Special Issue Paper. Environmetrics **DOI:10.1002/env.1115.**

Kuhnert, P. M., T. G. Martin, and S. P. Griffiths. 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecology Letters **13**:900-914.

Kuhnert, P. M., T. G. Martin, K. Mengersen, and H. P. Possingham. 2005. Assessing the impacts of grazing levels on bird density in woodland habitat: a Bayesian approach using expert opinion. Environmetrics **16**:717-747.
Kynn, M. 2008. The 'heuristics and biases' bias in expert elicitation. Journal of the Royal Statistical Society: Series A (Statistics in Society) **171**:239-264.
Langhammer, P. F., Bakarr, M.I, Bennun, L.A. et al. 2007. Identification and gap analysis of key biodiversity areas: Targets for comprehensive protected area systems Technical Report No 15. IUCN Gland, Available from http://data.iucn.org/dbtw-wpd/edocs/PAG-015.pdf.
Lele, S. R., and K. L. Allen. 2006. On using expert opinion in ecological analyses: a frequentist approach. Environmetrics **17**:683-704.
Lichtenstein, S., and B. Fischhoff. 1977. Do those who know more also know more about how much they know. Organizational Behavior and Human Performance **20**:159-183.
Lichtenstein, S., B. Fischhoff, and L. D. Phillips. 1982. Calibration of probabilities: The state of the art to 1980. Pages 306-334 in D. Kahneman, P. Slovic, and A. Tversky, editors. Judgment under uncertainty: Heuristics and biases. Cambridge University Press, New York.
Linstone, H. A., and M. Turoff 1975. The Delphi Method: Techniques and Applications. Addison-Wesley Publishing, Boston.
Low-Choy, S. 2011. Priors: silent or active partners of Bayesian inference? in K. Mengersen, C. Alston, and A. N. Pettitt, editors. Bayesian Statistics By Example In press.
Low-Choy, S., J. Murray, A. James, and K. Mengersen. 2010. Indirect elicitation from ecological experts: from methods and software to habitat modelling and rock-wallabies in A. O'Hagan, and M. West, editors. Handbook of Applied Bayesian Analysis. Oxford University Press, Oxford, UK.
Low-Choy, S., J. Murray, A. James, and K. Mengersen. 2011a. Elicitator: a user friendly, interactive tool to support elicitation of expert knowledge in A. Perrera, C. A. Drew, and C. Johnson, editors. Expert Knowledge and its Applications in Landscape Ecology. In press.
Low-Choy, S., R. A. O'Leary, and K. L. Mengersen. 2009. Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. Ecology **90**:265-277.
Low-Choy, S., P. Whittle, and C. Anderson. 2011b. Quantitative approaches to designing plant biosecurity surveillance in S. McKirdy, editor. Biosecurity in Agriculture and the Environment, In Press.
Ludwig, D., M. Mangel, and B. Haddad. 2001. Ecology, conservation, and public policy. Annual Review of Ecology and Systematics **32**:481-517.
MacMillan, D. C., and K. Marshall. 2006. The Delphi process - an expert-based approach to ecological modelling in data-poor environments. Animal Conservation **9**:11-19.
Marcot, B. G. 2006. Characterizing Species at Risk I: Modeling Rare Species Under the Northwest Forest Plan. Ecology and Society **11**:10 online.
Martin, T. G., P. M. Kuhnert, K. Mengersen, and H. P. Possingham. 2005. The power of expert opinion in ecological models using Bayesian methods: Impact of grazing on birds. Ecological Applications **15**:266-280.
McBride, M. F., and M. A. Burgman. 2011. What is expert knowledge, how is such knowledge gathered, and how do we use it to address questions in landscape ecology? in A. Perera, C. Johnson, and C. A. Drew, editors. Expert Knowledge and Its Application in Landscape Ecology. Springer.
McCarthy, M. A. 2007. Bayesian Methods for Ecology. Cambridge University Press.

McDonald-Madden, E., M. C. Runge, H. P. Possingham, and T. G. Martin. 2011. Optimal timing for managed relocation of species faced with climate change. Nature Climate Change **1**:261-265.

McKenzie, C. R. M., M. J. Liersch, and I. Yaniv. 2008. Overconfidence in interval estimates: What does expertise buy you? Organizational Behavior and Human Decision Processes **107**:179-191.

Meyer, M., and J. Booker 1991. Eliciting and analyzing expert judgment: a practical guide Academic Press New York.

Moore, D. A., and P. J. Healy. 2008. The trouble with overconfidence. Psychological Review **115**:502.

Morgan, M. G., and M. Henrion 1990. Uncertainty: a guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, New York.

Morgan, M. G., L. F. Pitelka, and E. Shevliakova. 2001. Elicitation of expert judgments of climate change impacts on forest ecosystems. Climatic Change **49**:279-307.

Murray, J. V., A. W. Goldizen, R. A. O'Leary, C. A. McAlpine, H. P. Possingham, and S. L. Choy. 2009. How useful is expert opinion for predicting the distribution of a species within and beyond the region of expertise? A case study using brush-tailed rock-wallabies Petrogale penicillata. Journal of Applied Ecology **46**:842-851.

O'Hagan, A., C. E. Buck, A. Daneshkhah, J. R. Eiser, P. H. Garthwaite, D. J. Jenkinson, J. E. Oakley, and T. Rakow 2006. Uncertain judgements: eliciting experts' probabilities. Wiley, Chichester.

O'Leary, R. A., S. Low-Choy, J. V. Murray, M. Kynn, R. Denham, T. G. Martin, and K. Mengersen. 2009. Comparison of three expert elicitation methods for logistic regression on predicting the presence of the threatened brush-tailed rock-wallaby Petrogale penicillata. Environmetrics **20**:379-398.

O'Neill, S. J., T. J. Osborn, M. Hulme, I. Lorenzoni, and A. R. Watkinson. 2008. Using expert knowledge to assess uncertainties in future polar bear populations under climate change. Journal of Applied Ecology **45**:1649-1659.

Oakley, J. E., and A. O'Hagan. 2010. SHELF: the Sheffield Elicitation Framework (version 2.0). School of Mathematics and Statistics, University of Sheffield, http://tonyohagan.co.uk/shelf.

Oskamp, S. 1965. Overconfidence in case-study judgments. Journal of Consulting Psychology **29**:261-265.

Regan, H. M., M. Colyvan, and M. A. Burgman. 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. Ecological Applications **12**:618-628.

Rothlisberger, J. D., D. M. Lodge, R. M. Cooke, and D. C. Finnoff. 2010. Future declines of the binational Laurentian Great Lakes fisheries: the importance of environmental and cultural change. Frontiers in Ecology and the Environment **8**:239-244.

Runge, M. C., S. J. Converse, and J. E. Lyons. 2011. Which uncertainty? Using expert elicitation and expcetd value of information to design an adaptive program. Biological Conservation **144**:1214-1223.

Shephard, G. G., and C. W. Kirkwood. 1994. Managing the jugmental probability elicitation process: a case study of analyst/manager interaction. IEEE Transactions on Engineering Management **41**:414-425.

Slovic, P. 1999. Trust, emotion, sex, politics, and science: surveying the risk-assessment battlefield. Risk Analysis **19**:689-701.

Smith, C., A. L. Howes, B. Price, and C. A. McAlpine. 2007. Using a Bayesian belief network to predict suitable habitat of an endangered mammal - The Julia Creek dunnart (Sminthopsis douglasi). Biological Conservation **139**:333-347.

Smith, C., R. D. van Klinken, L. Seabrook, C. A. McAlpine, and J. Ryan. 2011. Toward a general framework for modelling the susceptibility of landscapes to plant invasion. Diversity and Distributions **in press**.

Soll, J. B., and J. Klayman. 2004. Overconfidence in interval estimates. Journal of experimental psychology: learning, memory and cognition **30**:299-314.

Soll, J. B., and R. P. Larrick. 2009. Strategies for revising judgment: How (and how well) people use others' opinions. Journal of Experimental Psychology: Learning, Memory, and Cognition **35**:780.

Speirs-Bridge, A., F. Fidler, M. McBride, L. Flander, G. Cumming, and M. A. Burgman. 2010. Reducing overconfidence in the interval judgments of experts. Risk Analysis **30**:512-523.

Sugiura, K., and N. Murray. 2011. Risk analysis and its link with standards of the World Organisation for Animal Health. Rev. sci. tech. Off. int. Epiz **30**:281-288.

Sutherland, W. J. 2006. Predicting the ecological consequences of environmental change: A review of the methods. Journal of Applied Ecology **43**:599-616.

Sutherland, W. J., Bardsley, S., Bennun, L., Clout, M., Côté, I.M., Depledge, M.H., Dicks, L.V., Dobson, A.P., Fellman, L., Fleishman, E., Gibbons, D.W., Impey, A.J., Lawton, J.H., Lickorish, F., Lindenmayer, D.B., Lovejoy, T.E., Mac Nally, R., Madgwick, J., Peck, L.S., Pretty, J., Prior, S.V., Redford, K.H., Scharlemann, J.P.W., Spalding, M. & Andrew R. Watkinson, A.R. . 2011. A horizon scan of global conservation issues for 2011. Trends in Ecology and Evolution **26,** :10-16.

Teigen, K. H., and M. Jorgensen. 2005. When 90% confidence intervals are 50% certain: On the credibility of credible intervals. Applied Cognitive Psychology **19**:455-475.

Tetlock, P. E. 2005. Expert political judgment:how good is it? how can we know? Princeton University Press, Princeton.

Tsai, C. I., J. Klayman, and R. Hastie. 2008. Effects of amount of information on judgment accuracy and confidence. Organizational Behavior and Human Decision Processes **107**:97-105.

Tversky, A., and D. Kahneman. 1974. Judgement under uncertainty: Heuristics and biases. Science **185**:1124-1131.

Supplementary Information

Supplementary Information

Table S1a. Steps in the expert elicitation process illustrated with an example of the impact of livestock grazing on bird species (Kuhnert et al. 2005; Martin et al. 2005).

I. Decide how expert knowledge will be used		Example from Martin et al. (2005) and Kuhnert et al. (2005)
1	Determine the purpose of the elicitation (e.g. how will expert information be used)	Expert data to complement empirical data on the impact of livestock grazing on 31 Australian woodland birds in a Bayesian model
II. Determine what to elicit		
2	Define the research question	Are bird species increasing, decreasing or showing no change in relative abundance in response to different livestock grazing management practices?
III. Designing the elicitation process		
3	Determine the elicitation format (email survey, phone interview, face to face)	Email format was chosen since experts were widely dispersed and resources were not available to bring them all together
4	Identify expert(s) and make contact to determine whether they will participate in the study	30 experts in bird ecology identified and contacted by phone to ask if they would participate in the elicitation exercise; 20 agreed to participate
5	Develop background materials, test questions, research questions, scenarios	Define different levels of grazing management intensity and illustrate with photos. Develop instructions and an example of how to complete the survey
6	Design the elicitation procedure, describe the logistics of expert interactions (if any), and acquisition of the expert judgements	Expert responses gathered independently via email and compiled in a spreadsheet
7	Construct method for the analysis and synthesis of expert data including uncertainty	Expert responses weighted equally when calculating mean and variance
8	Outline elicitation roles: problem owner, facilitator, analyst and experts	In this case, the problem owner, facilitator and analyst were the same person (Martin). Kuhnert provided additional analyst expertise.
9	Expert training and calibration to resolve any misunderstanding regarding the questions and what is expected	Issues of linguistic uncertainty were resolved over the phone or via email prior to the elicitation. No calibration was used.
IV. Perform the elicitation		
10	Determine whether information will be elicited directly or indirectly	It was deemed impossible for most experts to provide estimates of actual bird species abundance with any level of accuracy, hence an indirect elicitation approach was used where experts were asked to estimate for

		each bird species and grazing level, relative changes in abundance as +1 increase, 0 no change, -1 decrease
11	Single or multiple experts	Multiple experts were chosen since data could then be combined to get estimates of the uncertainty around the mean response for each species and grazing level
12	If multiple experts, how will estimates be combined?	Experts provided independent estimates that were combined via linear averaging. Experts were not given the opportunity to revise their estimates in light of responses of others
V. Encode the elicited information		
13	Determine how the elicited information will be encoded in a model	A Bayesian GLM model was used where the expert data formed 'prior' distributions representing the relative change in bird species abundance for each bird species under each grazing level. These prior distribution were then used to update empirical data collected for each of the bird species under each grazing level

Table S1b. Steps in the expert elicitation process illustrated with an example on habitat modelling.

I. Decide how expert knowledge will be used		Ecological case study (Murray et al. 2009), based on statistical method (Low-Choy et al. 2010; Low-Choy 2011) and software (James et al. 2010).
1	Determine the purpose of the elicitation (e.g. how will expert information be used)	Expert data used to address known limitations of empirical data (200 sites in each of two regions) on the habitat associations for site occupancy by brush-tailed rock wallabies in a Bayesian GLM
II. Determine what to elicit		
2	Define the research question	Within the envelope of the species, what is the relative contribution of each habitat factor to the probability of site occupancy by this species? Which habitat factors are the most or least important? Which areas are more or less suitable for the species?
III. Design the elicitation process		
3	Determine the elicitation format (email survey, phone interview, face to face)	Face-to-face intensive interview (standard script of main questions) supported by *Elicitator* software (generally 2-4 hours)
4	Identify expert(s) and make contact to determine whether they will participate in the study	9 experts identified with field experience of at least 10 occupied sites; all agreed to participate. Each expert had experience in one region (southern Queensland or northern New South Wales)
5	Develop background materials, test questions, research questions, scenarios	Identify habitat descriptors from previous smaller study (of 50 sites), obtain/develop corresponding GIS datasets, and illustrate with photos at site and landscape scales. Use stratified design to select 30 elicitation sites in area never visited by any experts. Develop interview script and refine through three pilots.
6	Design the elicitation procedure, describe the logistics of expert interactions (if any), and acquisition of the expert judgements	Expert responses recorded independently and compiled into a relational database using the *Elicitator* software package. Outside-in elicitation of theoretical (100% CrI) then realistic (95% and 50%) bounds, culminating in best estimate. Four forms of feedback provided during elicitation: recording and confirmation of response; reflecting via alternative graphical/textual representations; comparing across sites, and assessing the implications via statistical encoding and model assessment.
7	Construct method for the analysis and synthesis of expert data including uncertainty	Different methods of combining multiple opinions considered. Ecologically, experts and data of same or different region combined (Low-Choy et al. 2011a). Mathematically, averaging (Martin et al. 2005), linear pooling (O'Hagan et al 2006) and Bayesian methods considered (Low Choy et al. 2010). Sensitivity analysis conducted to highlight diversity and assess contribution of individual experts to consensus.
8	Outline elicitation roles: problem owner, facilitator, analyst and experts	Problem owners balanced among funding bodies: PhD supervisors (UQ,QUT), Research Management Plan (EPA), several community groups. Facilitator/Analysts: software designer James and statistical designer Low-Choy ensured elicitation methodology and protocol

		tailored to context. Ecologist Murray undertook elicitations (with comprehensive training by team) and initial analysis. Comprehensive analysis by statistical designer Low-Choy.
9	Expert training and calibration to resolve any misunderstanding regarding the questions and what is expected	Issues of linguistic uncertainty were resolved during pilots and email/questionnaire prior to the elicitation, and during pre-elicitation training (1 site). One site with known presence/absence used to calibrate. Photos and GIS of all habitat attributes available to experts to help clarify covariates.
IV. Perform the elicitation		
10	Determine whether information will be elicited directly or indirectly	It was deemed impossible for most experts to estimate effects of each habitat covariate (as is required with multiple criteria decision analysis). Hence an indirect elicitation approach was used where experts were asked to estimate relative likelihood of site occupancy with same habitat (number of sites in 100 such sites, within the species envelope).
11	Single or multiple experts	Multiple experts were chosen to capture diversity and consensus of knowledge (most experts had specific field experience and an overview was desired). However each expert was interviewed separately, since insufficient funding or ability to bring all experts (often sited in remote locations) together in one sitting.
12	If multiple experts, how will estimates be combined?	Experts provided independent estimates, not updated in light of other experts. Linear pooling was considered to mathematically combine estimates, in addition to others (detailed above in item 7, encoding).
V. Encode the elicited information		
13	Determine how the elicited information will be encoded in a model	A Bayesian GLM model was used where the expert data informed prior distributions on parameters describing relative contribution of each habitat factor to relative likelihood of site occupancy. These priors were obtained by extending a conditional mean priors approach to accommodate uncertainty as well as best estimates in each response. The algorithm is embedded in the tool *Elicitator* (James et al. 2010; Low-Choy et al. 2011a).

Table S3. Subjective biases encountered in expert elicitation

Bias	**Description**	**Example**	**References**
Anchoring	Final estimates are influenced by an initial salient estimate, either generated by the individual or supplied by the environment	People give a higher estimate of the length of the Mississippi River if asked whether it is longer or shorter than 5000 miles, than asked whether it is longer or shorter than 200 miles	(Tversky & Kahneman 1974; Jacowitz & Kahneman 1995)
Availability bias	People's judgments are influenced more heavily by the experiences or evidence that most easily come to mind	Tornadoes are judged as more frequent killers than asthma, even though the latter is 20 times more likely	(Tversky & Kahneman 1973; Lichtenstein et al. 1978)
Confirmation bias	People search for or interpret information (consciously or unconsciously) in a way that confirms their prior beliefs	Scientists may judge research reports that agree with their prior beliefs to be of higher quality than those that disagree	(Lord et al. 1979; Koehler 1993)
Dominance	Social pressures induce group members to conform to the beliefs of a senior or forceful member of the group	Groups spend more of their time addressing the ideas of high-status members than they do exploring ideas put forward by lower-status members	(Maier & Hoffman 1960)
Egocentrism	Individuals tend to give more weight to their own opinions then to the opinions of others than is warranted	Individuals attribute weights of on average 20-30% to advisor opinions in revising their judgments, when higher weights would have been optimal	(Yaniv 2004)
Framing	Individuals draw different conclusions from the same information, depending on how that information is presented	Presenting probabilities as natural frequencies (e.g. 6 subpopulations out of 10) helps people reason with probabilities and reduce biases such as overconfidence	(Gigerenzer & Hoffrage 1995)
Groupthink	When groups become more concerned with achieving concurrence among their members than in arriving at carefully considered decisions	Foreign policy fiascos such as the invasion of North Korea and the Bay of Pigs invasion have been attributed to decision makers becoming more concerned with retaining group approval than making good decisions	(Janis 1972)
Halo effects	When the perception of an attribute for an individual or object is influenced by the perception of another attribute or attributes	Attractive people are ascribed more intelligence then less attractive people	(Nisbett & Wilson 1977)
Overconfidence	The tendency for people to have greater confidence in their judgments then is warranted by their level of knowledge	People frequently provide 90% confidence intervals that contain the truth on average only 50% of the time	(Lichtenstein et al. 1982; Soll & Klayman 2004)

References

Gigerenzer, G., and U. Hoffrage. 1995. How to improve Bayesian reasoning without instruction: frequency formats. Psychological Review **102**:684-704.

Jacowitz, K. E., and D. Kahneman. 1995. Measures of anchoring in estimation tasks. Personality and Social Psychology Bulletin **21**:1161.

James, A., S. L. Choy, and K. Mengersen. 2010. Elicitator: an expert elicitation tool for regression in ecology. Environmental Modelling & Software **25**:129-145.

Janis, I. L. 1972. Victims of groupthink. Houghton Mifflin, New York.

Koehler, J. J. 1993. The influence of prior beliefs on scientific judgments of evidence quality Organizational Behavior and Human Decision Processes **56**:28-55.

Kuhnert, P. M., T. G. Martin, K. Mengersen, and H. P. Possingham. 2005. Assessing the impacts of grazing levels on bird density in woodland habitat: a Bayesian approach using expert opinion. Environmetrics **16**:717-747.

Lichtenstein, S., B. Fischhoff, and L. D. Phillips. 1982. Calibration of probabilities: The state of the art to 1980. Pages 306-334 in D. Kahneman, P. Slovic, and A. Tversky, editors. Judgment under uncertainty: Heuristics and biases. Cambridge University Press, New York.

Lichtenstein, S., P. Slovic, B. Fischhoff, M. Layman, and B. Combs. 1978. Judged frequency of lethal events. Journal of Experimental Psychology-Human Learning and Memory **4**:551-578.

Lord, C. G., L. Ross, and M. R. Lepper. 1979. Biased assimilation and attitude polarization: effects of prior theories on subsequently considered evidence. Journal of Personality and Social Psychology **37**:2098-2109.

Low-Choy, S., J. Murray, A. James, and K. Mengersen. 2010. Indirect elicitation from ecological experts: from methods and software to habitat modelling and rock-wallabies in A. O'Hagan, and M. West, editors. Handbook of Applied Bayesian Analysis. Oxford University Press, Oxford, UK.

Low-Choy, S., Murray, J., James, A. and Mengersen, K. 2011. Elicitator: a user friendly, interactive tool to support elicitation of expert knowledge in A. Perrera, Drew, C.A., Johnson, C., editor. Expert Knowledge and its Applications in Landscape Ecology. In press.

Maier, N. R. F., and L. R. Hoffman. 1960. Quality of first and second solutions in group problem solving. Journal of Applied Psychology **44**:278-283.

Martin, T. G., P. M. Kuhnert, K. Mengersen, and H. P. Possingham. 2005. The power of expert opinion in ecological models using Bayesian methods: Impact of grazing on birds. Ecological Applications **15**:266-280.

Murray, J. V., A. W. Goldizen, R. A. O'Leary, C. A. McAlpine, H. P. Possingham, and S. L. Choy. 2009. How useful is expert opinion for predicting the distribution of a species within and beyond the region of expertise? A case study using brush-tailed rock-wallabies Petrogale penicillata. Journal of Applied Ecology **46**:842-851.

Nisbett, R. E., and T. D. Wilson. 1977. Halo effect: evidence for unconscious alteration of judgments. Journal of Personality and Social Psychology **35**:250-256.

Soll, J. B., and J. Klayman. 2004. Overconfidence in interval estimates. Journal of experimental psychology: learning, memory and cognition **30**:299-314.

Tversky, A., and D. Kahneman. 1973. Availability: A heuristic for judging frequency and probability. Cognitive Psychology **5**:207-232.

Tversky, A., and D. Kahneman. 1974. Judgment under uncertainty: heuristics and biases. Science **185**:1124-1131.

Yaniv, I. 2004. Receiving other people's advice: influence and benefit. Organizational Behavior and Human Decision Processes **93**:1-13.