

REVIEW

Methods for collaboratively identifying research priorities and emerging issues in science and policy

William J. Sutherland^{1*}, Erica Fleishman², Michael B. Mascia³, Jules Pretty⁴
and Murray A. Rudd⁵

¹Conservation Science Group, Department of Zoology, University of Cambridge, Cambridge CB2 3EJ, UK; ²Bren School of Environmental Science & Management, 6832 Ellison Hall, University of California, Santa Barbara, CA 93106-3060, USA; ³Conservation Science Program, World Wildlife Fund, 1250 24th St NW, Washington, DC 20009, USA; ⁴Interdisciplinary Centre for Environment and Society, Department of Biological Sciences, University of Essex, Colchester CO4 3SQ, UK; and ⁵Environment Department, University of York, Heslington, York YO10 5DD, UK

Summary

1. There is a widely recognized gap between the data generated by researchers and the information required by policy makers. In an effort to bridge the gap between conservation policy and science, we have convened in several countries multiple groups of policy makers, practitioners and researchers to identify priority information needs that can be met by new research in the social and natural sciences.

2. The exercises we have coordinated included identification of priority policy-relevant research questions in specific geographies (UK, USA, Canada); questions relating to global conservation; questions relating to global agriculture; policy opportunities in the United Kingdom; and emerging global conservation issues or ‘horizon scanning’.

3. We outline the exercises and describe our methods, which are based on principles of inclusivity, openness and democracy. Methods to maximize inclusiveness and rigour in such exercises include solicitation of questions and priorities from an extensive community, online collation of material, repeated voting and engagement with policy networks to foster uptake and application of the results.

4. These methods are transferable to a wide range of policy or research areas within and beyond the conservation sciences.

Key-words: collaboration, horizon scanning, participation, planning, policy makers, priority setting

Introduction

There are at least three main audiences for priority-setting exercises at the nexus of conservation and environmental science and policy. The first audience is policy makers and practitioners in public, private and nonprofit organizations (hereafter *policy makers*). These individuals are likely to benefit from development of an agenda for scientific inquiry that meets their information needs. The second audience, funders of research, may find it useful to support broad themes that suppliers and users of knowledge jointly have identified as relevant. The third audience, researchers, can better apply the knowledge they create and synthesize by learning which

questions and issues are considered most important by policy makers.

Substantive communication between producers and users of knowledge has been identified repeatedly as essential for developing credible, relevant and legitimate institutional and technological solutions to environmental policy challenges (Cash *et al.* 2003; Sarewitz & Pielke 2007). In a review of environmental issues for which avoidable, undesirable outcomes resulted from delay between identification of the source of a problem (e.g. asbestos, Bovine Spongiform Encephalopathy, sulphur dioxide) and action, the European Environment Agency (2001) identified two key lessons: ‘research and monitor for early warnings’ and ‘search out and address blind spots and gaps in scientific knowledge’. Holmes & Clark (2008) concluded that science generally was not used to inform policy at a sufficiently early stage to have a positive effect on

*Correspondence author. E-mail: w.sutherland@zoo.cam.ac.uk
Correspondence site: <http://www.respond2articles.com/MEE/>

outcomes. Moreover, public science and policy institutions often are responsive rather than proactive (Pretty 2009), as illustrated by the UK government's response to the 2001 outbreak of foot and mouth disease (Anderson 2007). Thus, our aim is not to develop policy but encourage more timely research in order to aid the development of evidence-based policy (Lawton 2007).

The persistent gap between the information generated by natural and social scientists and the information desired by conservation policy makers has prompted calls and requirements by funders for science to be incorporated more comprehensively within policy processes (Lawton 2007; Sutherland *et al.* 2004). For example, the Canadian Social Science and Humanities Research Council's 2008 Environmental Issues special call for proposals explicitly stated that funded research must advance the objectives of the Canadian science and technology strategy (Government of Canada 2007) by focusing on government priorities that included environmental science. Researchers may wish to inform policy decisions through their research, but often lack knowledge of policy makers' immediate information needs and longer-term priorities. One potential solution is to promote collaborations between policy makers and researchers to identify by consensus a set of priorities for scientific inquiry.

To encourage greater collaboration between policy makers and scientists affiliated with academic, public and nongovernmental organizations, and to catalyse the development of policy-relevant conservation science, we have facilitated several participatory exercises to identify research priorities. We also conducted a participatory exercise to identify topics for which opportunities exist to create new policies and another exercise to identify emerging issues in conservation. This suite of participatory exercises has encouraged others with similar goals to adopt analogous processes (e.g. Morton *et al.* 2009; Armsworth *et al.* 2010). These exercises can establish the necessary foundations for joint fact-finding, a process through which parties with potentially diverse interests collaborate to identify, define and answer critical scientific questions that hinder development of effective policies (Karl, Susskind, & Wallace 2007). However, these exercises are more challenging to orchestrate successfully than may be superficially apparent. Our aim is to foster further collaborative priority-setting among researchers and policy makers by outlining our transferable methods and insights. Our work was founded on the understanding that rigorous, relevant science is necessary to inform societal decisions regarding conservation policy and practice.

Materials and methods

Each exercise we organized (Table 1) had one of three major objectives: (i) to identify a set of research questions that, if answered, would increase the effectiveness of policies related to a given topic (Sutherland *et al.* 2006, 2009; Fleishman *et al.* in press, Pretty *et al.* in press, Rudd *et al.* in press), (ii) to identify potential policy developments that would support conservation in the light of technological advances, changes in knowledge or environmental change (Sutherland

et al. in press), (iii) to conduct horizon scanning (Sutherland & Woodroof 2009), the systematic search for potential threats and opportunities that are currently poorly recognized (Sutherland *et al.* 2008, 2010a, Sutherland *et al.* 2011).

The exercises had a common structure but individual differences. Organisers identified the objectives and solicited and collated suggested questions, areas for policy development or emerging issues (horizons) from a large, diverse group of individuals. Organizers and participants reduced the set of suggestions by an iterative process of voting and discussion to produce a final list. This list was presented as part of a manuscript submitted to a peer-reviewed journal.

The key decisions faced by organizers of each exercise were the scope of issues, target number of issues, selection of participants and selection of voting methods.

For exercises in which questions were identified, we established the following criteria for questions: (i) answerable through a realistic research design, (ii) that have a factual answer that does not depend on value judgments, (iii) that address important gaps in knowledge, (iv) of a spatial and temporal scope that reasonably could be addressed by a research team, (v) not formulated as a general topic area, (vi) not answerable with *it all depends*, (vii) except if questioning a precise statement ('does the earth go round the sun?') should not be answerable by *yes* or *no* (i.e. not 'is X better for biodiversity than Y'), (viii) if related to impact and interventions, contains a subject, an intervention, and a measurable outcome. An ideal question suggests the design of research that is required to answer it or can be envisioned as translating the question (Pullin, Knight, & Watkinson 2009) into directly testable research hypotheses.

PREVIOUS EXERCISES

The various exercises differed in the methods used (see Table 1). Some of this variation was because of different circumstances and experimenting with different approaches. However, there has been a gradual standardization and improving of the methods through the different exercises.

One hundred ecological questions of highest relevance to policy makers in the United Kingdom

A team of two led the *UK questions* exercise, which was designed to identify the highest priority ecological questions that policy makers wish to have answered (Sutherland *et al.* 2006). Questions compiled by participants prior to the workshop were classified into 12 general themes. On the first day of the workshop, facilitators organized four sets of three concurrent breakout sessions corresponding to each of the themes. In each breakout session, policy makers started the process of culling questions in their theme by collectively affixing gold (highest priority – approximately top 10% of questions; two points) and silver (high priority – approximately next 10% of questions; one point) stars to a single set of questions listed on posters. Subsequent discussion and modification were restricted to the set of questions with at least one star. This process resulted in 188 questions across the 12 themes. On the second day, the organisers divided the 188 questions into three approximately equal sets by combining related themes. Each set was subsequently reduced by concurrent breakout groups to the final target number (33 for two groups and 34 for one group, for a total of 100). There was no plenary session. The aim of this exercise (unlike others described below) was to determine the priorities of policy makers; academics did not vote and it was made clear that they were present at the workshop to assist the policy makers with modifying questions and identifying questions that already have

Table 1. Features of collaborative exercises conducted between 2006 and 2010 to identify priority research questions, opportunities for developing new policies and emerging issues in conservation

Focal topic	Region	Solicitation method	Number of candidate questions, priorities, or issues	Number of individuals involved in producing candidate material	Final number of questions or issues	Methods for identification and refinement of questions	Characteristics of workshop participants	Reference
Questions of ecological importance	UK	Workshop participants canvassed their professional networks and compiled questions	1003	654	100	Voting by policy makers to narrow set of questions, then two stages of consensus-based subgroup discussions to select final set of questions and refine wording	Purposive sample ($n = 39$) of those in the UK concerned with the application of ecology and affiliated with government agencies ($n = 13$), nongovernmental organizations ($n = 15$), and academia ($n = 11$)	Sutherland <i>et al.</i> (2006)
Questions of importance to the conservation of biological diversity	Global	Workshop participants canvassed their professional networks and compiled questions	2291	761	100	Pre-workshop voting by participants to narrow set of questions, then workshop with consensus-based discussions by subgroups of topic-area specialists, followed by consensus-based full-group discussions to select final set of questions. Consensus-based post-workshop email dialogue to refine wording	A purposive sample of government or intergovernmental agencies ($n = 9$), multinational nongovernmental organizations ($n = 21$), and academia ($n = 14$) that included representatives from all continents except Antarctica	Sutherland <i>et al.</i> (2009)
Emerging issues in conservation	Global	Workshop participants searched literature and canvassed their professional networks	71	158	15	Pre-workshop voting on alignment between issues and criteria and whether participant was aware of the issue. Discussion at workshop followed by voting. Consensus-based post-workshop email dialogue to refine wording	A purposive sample of professional horizon scanners ($n = 4$), representatives from organizations with wide responsibilities ($n = 8$), individuals with specific expertise (6) and academics (7)	Sutherland <i>et al.</i> (2010b, in press). Numbers given in previous columns relate to the second exercise
Priority opportunities for nature conservation	UK	Workshop participants canvassed their professional networks and compiled questions	117	31	25	Pre-workshop voting by participants to narrow and prioritize list of questions, then discussion during workshop to shorten list. Discussion and voting to produce final list	Purposive sample of 31 representatives from UK governmental bodies ($n = 10$), nongovernmental organizations ($n = 10$), and academia ($n = 11$)	Sutherland <i>et al.</i> (2010)

Table 1. (Continued)

Focal topic	Region	Solicitation method	Number of candidate questions, priorities, or issues	Number of individuals involved in producing candidate material	Final number of questions or issues	Methods for identification and refinement of questions	Characteristics of workshop participants	Reference
Priority questions for informing conservation policy	USA	Workshop participants canvassed their professional networks; questions compiled via dedicated website	531	374	40	Criteria-based pre-workshop voting to narrow set of possible questions, then workshop with consensus-based discussions by subgroups of topic-area specialists, followed by consensus-based full-group discussions to select final set. Consensus-based post-workshop email dialogue to refine wording	Purposive sample of USA conservation thought leaders ($n = 42$) from government agencies ($n = 12$), nongovernmental organizations ($n = 9$), private sector ($n = 5$), professional societies ($n = 6$), and academia ($n = 3$). Seven organizers also attended and facilitated the workshop	Fleishman <i>et al.</i> (unpublished manuscript)
Questions relevant to policies for agriculture and food systems	Global	Core group of 54 participants canvassed their professional networks and compiled questions	618	~ 200	100	Fourteen subgroups prioritized questions. Then full voted on final questions	No workshop. Electronic interactions among participants were mediated by secretariat	Pretty <i>et al.</i> (in press)
Priority questions for informing conservation policy	Canada	Workshop participants canvassed their professional networks; questions compiled via dedicated website	396	271	40	Criteria-based pre-workshop voting to narrow set of questions, then workshop with consensus-based discussions by subgroups of topic-area specialists, followed by consensus-based full-group discussions to select final set. Consensus-based post-workshop email dialogue to refine wording	Purposive sample of Canadian conservation thought leaders ($n = 27$) from government agencies ($n = 10$), nongovernmental organizations ($n = 3$), private sector ($n = 2$), policy sector (5), and academia ($n = 7$). Eight organizers and four students also attended and facilitated the workshop	Rudd <i>et al.</i> (in press)
Major emerging issues in conservation	UK	Workshop participants searched the literature and canvassed their professional networks and compiled issues	195	452	25	Pre-workshop voting by participants, full-group discussion and voting during workshop, final list determined by voting	Purposive sample of experts ($n = 32$) from major governmental bodies ($n = 13$), nongovernmental organizations ($n = 6$), universities ($n = 11$), and industry ($n = 1$). One science journalist	Sutherland <i>et al.</i> (2008)

been answered. The organizers led the drafting of a manuscript summarizing this process and the resulting list of 100 questions, which was circulated to all participants for comments. In addition to the manuscript, which was published in a peer-reviewed journal, results were disseminated through a press release and presentation to the funders.

One hundred priority global conservation questions

The *global questions* exercise was designed to identify questions that, if answered, would have a high probability of increasing the success of global conservation actions (Sutherland *et al.* 2009). Workshop participants solicited questions widely, using diverse methods (including email announcements, workshops and personal communication). Most of the individuals who contributed questions provided their name and affiliation. Names and affiliations were used for reporting purposes only. The questions were classified *post hoc* into 15 general themes (e.g. forest) and various numbers of subthemes (e.g. forest: carbon) for ease of discussion and prioritization.

In advance of the workshop, the full set of contributed questions was circulated in an Excel spreadsheet to each workshop participant. Participants were asked to select questions that they believed should be considered for the final 100 within any themes of which they had sufficient knowledge, retaining roughly 5% of the questions (corresponding to 100 questions from the 2291 submitted) within the themes they reviewed. They were encouraged to engage multiple colleagues in selection of questions and were invited to rephrase questions or contribute additional questions to fill any noticeable gaps. The organizer compiled participants' votes and – prior to the workshop – circulated to all participants the resulting list of priority questions, the score (summed votes for retention) for each, and any suggestions for rephrasing. Of the 2291 original questions, 1655 received at least one such 'priority' vote from the workshop participants. Many participants retained more than 100 questions.

This 2-day workshop focused on winnowing these 1655 questions into a core set of 100 questions. During the first day of the workshop, expert subgroups addressed each of the 15 themes, with 3–4 subgroups meeting in concurrent breakout sessions to winnow and refine questions. This process reduced the list of 1655 questions to 258. During the second day, three concurrent subgroups each addressed 3–5 pooled themes, winnowing the remaining set of questions further, until each subgroup identified its 30 primary priority and 10 secondary priority questions. At the end of the second day, the organizer guided a plenary discussion to address overlaps, gaps, awkward phrasing and other concerns with the 90 highest priority questions. Decisions on whether to remove or merge thematically overlapping questions were made by majority vote. Eight questions were removed, leaving 82 questions. Participants then voted for 10 questions from among the 30 second-priority questions; the 18 questions with the greatest number of votes were added to the existing 82 for a total of 100 questions. One participant edited the questions for each theme. The organizer then inserted the 100 questions into a draft manuscript and circulated it to all participants to edit, resulting in eventual publication of a peer-reviewed manuscript.

UK policy options in conservation

The aim of the *UK policy priorities* exercise (Sutherland *et al.* in press) was to identify opportunities for new policies for the United Kingdom presented by new technologies (such as nanotechnology), new issues (such as effectiveness of protected areas as climate changes) or opportunities to modify and increase the effectiveness of current poli-

cies. The organizer invited participants to submit briefs on a maximum of ten policy issues. Briefs outlined the background, policy options, and research needs for each issue. The submitted issues were classified into 12 general themes. Prior to the workshop, participants scored the issues in each theme for which they were knowledgeable on a scale from 1 (low priority) to 10 (high priority). Participants were also invited to identify issues or suggest changes to the issues. Before the workshop, the mean score was calculated for each issue and these scores and collated comments were circulated to workshop participants.

During the first day of the workshop, four sets of three concurrent subgroups reduced the list of 117 issues to 42 and ranked the issues retained by their subgroup. On the second day, all participants met in plenary. Participants discussed and modified each of the 42 issues; each participant then independently scored each issue on scale from 1 (low priority) to 10 (high priority). The mean scores for each issue were reviewed and discussed, and the top 25 issues were selected. Participants then refined the policy options and research needs for the 25 issues. After the meeting, the issue briefs were rewritten by small teams. Options and research needs were further refined by all participants.

Horizon scanning for forthcoming issues in the United Kingdom

The aim of the *UK horizon scanning* exercise (Sutherland *et al.* 2008) was to identify opportunities for new environmental policies within the UK. Participants were asked to suggest issues that could affect the probability of achieving social, economic and ecological objectives in the future but currently seem to receive little attention and to write a 100–200 word summary of each issue they contributed. The submitted issues were classified into 12 general themes and sent to each participant.

Prior to the workshop, participants scored those issues for which they had sufficient expertise (many people scored all issues) on four scales, each from 1 (low priority) to 9 (high priority): likelihood that the suggested issue may be imminent; likely effect on society's social, economic or ecological objectives; novelty from the perspective of the participants; and priority for inclusion among the 25 top issues. Participants worked independently of each other to score issues but often worked in collaboration with other colleagues. During the first day of the workshop, four sets of three concurrent sessions prioritized and ranked three to four issues per theme, resulting in a total of 41 issues. During the second day of the workshop, each issue was described by the session chair, discussed by the full group of participants and then given a score from 1 (low priority) to 9 (high priority) by each participant. Issues were ranked by mean score, and the final list of 25 issues was selected by consensus in a final session.

Three concurrent subgroups then assessed the imminence and the opportunity and threat presented by each issue as low, medium or high on the basis of the initial scores by all participants and subsequent discussion. For each issue, they also identified research questions. Participants took responsibility for rewriting individual issues, which were circulated in a draft manuscript to all participants for editing.

Global horizon scanning for conservation issues

The global horizon scanning exercise (Sutherland *et al.* 2010b, in press) was aimed at identifying issues that could be important yet have attracted relatively little attention given its importance. The horizon scanning has been conducted in two successive years and is

expected to become an annual exercise. The methods were similar in the 2009 and 2010 workshops but the details for 2010 are described here. Participants identified issues through reading, professional contacts and, in some cases, seeking suggestions from other members of their organizations. Participants were asked to identify 1–3 issues and prepare for each an approximately 200 word brief with references.

The briefs were circulated to all participants, who were asked to vote on a 1–10 scale (where 10 is the most suitable) whether the issue was important but had attracted little attention. The 35 issues with the highest mean scores retained. Participants were also asked whether they already were aware of the issue. The list of issues that did not receive one of the 35 highest mean scores was circulated, and participants were invited to retain any of those issues. Two participants were asked to research each issue and consider whether it was likely to be important, whether it had received little attention and whether the text describing the issue should be modified.

Each issue was discussed by the full group during a 1-day workshop. Participants were given the mean scores, the percentage of the group who had heard of the issue and any comments provided on any of the issues. The identities of the person who suggested the issue and the two researchers for each issue were only known by the organizer. The proposer of the issue was asked not to speak until two others had spoken. After discussion, each individual scored each issue on a 1–100 scale (where 100 is the most suitable). The scores were converted to ranks and the mean rank calculated. The final list of 15 issues was selected by consensus in a final session.

One person edited the descriptions of the issues, and a draft manuscript was circulated to all participants for further editing.

Identification of 100 priority questions in global agriculture

The *agriculture questions* exercise (Pretty *et al.* 2010) was aimed at identifying issues related to global food security and sustainable agriculture. Some members of the core group of 55 participants convened workshops to generate questions, whereas others solicited questions from their networks of contacts. Questions were classified into 14 general themes by the coordinators. The exercise coordinators removed or combined overlapping questions. The coordinator appointed participants to revise and prioritize the questions in each theme. An expert group of 3–5 of the participants (participants could join more than one group) refined and reduced the number of questions in each theme, added new questions if there were obvious gaps and sorted questions into three groups: essential to retain (five questions), possibly retain (approximately 10 questions) and reject. The questions identified as essential and possible to retain were compiled by the organizer to produce a list of 70 essential questions and 146 possible questions. Each participant identified their 30 highest ranking questions from the set of 286 to produce a final consensus list of 100 questions. The process of revising questions and voting was conducted electronically.

Forty USA questions in conservation

In the *USA questions* exercise, a seven-member organizing team led a participatory process to identify research questions in conservation science with high relevance for decision-making in the United States by 2020. To refine the project scope and to identify mechanisms for improving uptake, the organizers first conducted informal interviews with nine current and former senior public sector policy makers and science advisors to public sector policy makers.

Participants were specifically invited as individuals rather than as formal representatives of their organizations. The workshop organizers and participants solicited questions from within their organizations, from other colleagues and in public forums (e.g. email lists). Questions were compiled via a simple website over 6 weeks; responses were anonymous unless the respondent chose to provide an organizational affiliation or name, and personal information was kept confidential.

In advance of the workshop, participants screened the list of questions and noted any that did not meet the seven criteria listed earlier. The workshop organizer compiled these responses. If a simple majority of participants noted that a given question did not meet the criteria, the question tentatively was discarded. The list of discarded questions was circulated to all participants to provide an opportunity for reconsideration.

During the workshop, participants in three sequential sets of three concurrent small-group discussions winnowed the list of 271 retained questions to 36 priority questions and 18 possible alternates. At this stage, the criteria for questions were treated as aspirational rather than strictly enforced. A plenary session refined the 36 proposed questions and filled gaps from the list of alternates to reach the target of 40 priority research questions. Those who attended the workshop in person or remotely subsequently refined the questions and draft manuscript via email dialogue.

Forty Canadian questions

The objective of the *Canadian questions* exercise was to identify questions that, if answered, would increase the effectiveness of policies related to conservation and resource management in Canada. A core team of eight designed the project. *Canadian questions* and *USA questions* were run in parallel and shared methods. Workshop organizers and participants solicited questions from within their organizations, from other colleagues and in public forums (e.g. email lists, social networking sites). Participants were asked to submit questions in either English or French via a website over a period of 5 weeks. French questions were translated into English. A total of 396 questions were collected.

Of the 39 participants, 28 were able to attend a workshop; those who were not able to attend were actively engaged in post-workshop question refinement. Prior to the workshop, the core team combined and refined questions to reduce the list to 242. During the workshop, attendees chose the highest 40 following the same general subgroup and plenary structure as in *USA questions*. Following the workshop, questions and a draft manuscript were refined by email with participation of the full group, including those unable to attend the workshop.

PRINCIPLES AND LESSONS ON METHODS

A number of guiding principles have emerged and been applied during the facilitation of the above exercises. The principles can be categorized with respect to (i) defining the project, (ii) organizing the participants, (iii) soliciting and managing questions or issues and (iv) disseminating results.

Defining the project

Vision. A clear vision makes it easier to devise methods, identify potential participants and design outputs. Elements of the vision include the aim and audience of the exercise. In many cases, the aim may be to convene a representative subset of conservation or resource

professionals, including policy makers to identify research priorities. In such cases, the target audience is the research community as well as funding bodies. The aim and audience inform decisions about the most appropriate means of communicating results. An article in a peer-reviewed journal is the traditional output for researchers and serves as a deliverable to demonstrate accountability to funders. Nevertheless, peer-reviewed manuscripts are usually not sufficient to effect change among conservation policy makers. Policy briefs, media engagement (e.g. press releases, op-eds) and working through policy specialists in professional societies are perhaps the most effective actions researchers can take to ensure that the aims of a priority-setting exercise are met. Funders may be interested in supporting outreach efforts and discussion sessions, especially those identified by policy makers.

Scope. A clear vision enables definition of a feasible scope for the exercise. For example, the *UK questions* were constrained to ecological topics, whereas the *global questions* were constrained to conservation of biological diversity. Opportunities included in the final version of the *UK policy priorities* exercise were required to be new and to have an apparent contemporary application to policy. It is important to define the precise geographical boundaries, both terrestrial and aquatic, for regional exercises.

Number of priorities. The number of priority questions or issues is partly related to the breadth of the topic. For some narrow topics, such as a particular threatened species or invasive species, we think it might be useful to identify, say, the five priority questions. A larger number of priorities may minimize the tendency to contribute questions that are extremely broad and inclusive. The target number of priorities should be tractable to select given the time available for in person or remote discussions and, if applicable, should be appropriate for future exploration of relative priorities among sectors on the basis of surveys. For example, in *Canadian questions* and *USA questions*, 40 questions approaches the maximum that can be accommodated in a best–worst scaling (Flynn *et al.* 2007) survey that ranks all questions.

Organizing the participants

Organizing team. The exercises described here have been managed by a single individual or by a team of 3–8 people with diverse skills, expertise and affiliations. Including or consulting an organizer who has participated in a previous priority-setting exercise ensures that new exercises benefit from experience rather than repeating mistakes. Similarly, familiarity with website development and data base management dramatically simplifies the process of compiling and organizing initial questions. The process of selecting and engaging a diverse set of workshop participants requires either a team or, if run by an individual, considerable consultation with experts.

Composition of participants. We typically used purposive sampling (subjective sampling with a purpose – in this case stratification) to invite a diverse set of suitable participants stratified by geography, disciplinary and subject matter expertise, organizations and other domains of interest. It is usually unrealistic to convene participants who represent all combinations of domains (for example each combination of region, topic and organizational type) so a challenge is to ensure that each domain is sufficiently represented by the people invited. Some exercises have involved funders of scientific research, industry representatives and members of the policy community. It is a serious challenge to obtain comprehensive coverage of all domains, especially if individuals withdraw from the exercise late in the process.

We have found that interaction among policy makers and academics helps to identify which questions are important, answerable by research, and for which substantial knowledge does not already exist. As the number of conservation priority-setting exercises has increased, we have aimed to engage new individuals to encourage greater inclusivity. If few individuals have expertise on a given topic, however, it may be necessary to engage individuals who have participated in one or more previous exercises. We have aimed to invite participants who have broad interests and will look beyond their organization's immediate priorities.

Any exercise requires distinct attention to diversity and inclusivity. For *USA questions*, we sought collective expertise in policy formulation, application of science to policy and funding of scientific research at different levels of government and different types of public and private organizations. For this exercise, we specifically invited participants as individuals, not as representatives of their organizations. We aimed for a mix of social and natural scientists, and for collective expertise in different biomes. Special efforts were made in *Canadian questions* to engage experts on Aboriginal issues because Aboriginal peoples are strongly dependent on the environment for subsistence and livelihood, and because their territorial and traditional lands are often the most vulnerable to climate change.

More formal techniques such as stakeholder mapping (Reed *et al.* 2009) may ensure comprehensive representation.

Number of participants. Larger groups result in a greater number of ideas, increased recognition that the exercise is a community-wide involvement. However, discussion among large groups may be relatively ineffective in generating output and undemocratic as many individuals become frustrated if they do not have opportunities to speak. We typically have aimed for a number of participants who does not prohibit plenary discussion. Approximately 40 participants seem, in our experience, to be in a comfortable upper limit. During the workshops, we typically divided into smaller groups for in-depth discussion of particular issues, but returned to plenary session to address cross-cutting issues and encourage exchange of ideas. Breakout groups of about 8–12 people generally have sufficient collective knowledge and diverse perspectives, yet are small enough to allow all individuals to participate in the discussion. We believe it is helpful to schedule ample time for socializing and networking during breaks and meals and in the evenings.

Gaps in expertise resulting from declined invitations to participate in the exercise can usually be addressed by issuing further invitations, but last-minute cancellations may lead to gaps that cannot be filled. Policy makers seem especially prone to conflicts that result in late cancellations. A solution for *Canadian questions* and *USA questions* was to contact key policy makers after the workshop and ensure they remained engaged in the process of question refinement. Our experience is that local participants have few travel costs but are more likely to be distracted by other commitments. The input of individuals who are unable to participate in all stages of the exercise is much less useful.

Facilitation of breakout sessions. In advance of the workshop, the organizing team typically designated either a member of the organizing team or a workshop participant to facilitate each breakout session. The facilitator is responsible for leading discussions to cull and refine questions that fall within a specific topic area or theme. For exercises that required winnowing a large number of questions to a much smaller set, we found it productive for the facilitators to have already identified for potential of removal of those questions that were redundant, received few votes for retention, or diverged considerably from the aspirational criteria for tractable questions. A

challenge is to balance the efficiency of strong leadership with the requirements for inclusivity and deliberative discussion.

Experienced facilitators with a knowledge of the subject are preferred, especially those who have previously participated in a collaborative priority-setting exercise. We have never used professional facilitators who do not have the subject-specific knowledge as we believe that knowledge is essential. In the USA and Canadian exercises, facilitators were asked to clarify when they were offering personal opinion.

We found it helpful to assign a recorder to work with the facilitator in each breakout session. The recorder is responsible for tracking in real time the breakout group's decisions regarding which questions to eliminate, combine and refine. The most effective recorders have been fairly experienced scientists in the topic area; they understand the group discussion, ask clarifying questions and translate the group dialogue into a rigorous research question. Effective recorders also are fast, accurate typists who are familiar with the software and hardware used during the workshop. In the ideal situation, the recorder often serves as a *de facto* co-facilitator.

For *Canadian questions* and *USA questions*, it was effective to print out lists of questions for each breakout session and to allow participants 10 min at the start of a breakout session to read over the questions and note which to retain or eliminate. We encourage storing and managing all questions in Excel, a simple and familiar program. The facilitator and recorder of each thematic breakout group can be provided with an Excel spreadsheet with the questions that must be winnowed and refined. Adding a column between the question number and the question itself allows the breakout group to designate priority questions or ideas with a simple notation. Projection of the Excel spreadsheet onto a large screen helps the group to collectively designate priority questions and, as necessary, refine the phrasing for each. Sorting and filtering the spreadsheet via the 'designated priority' column allows the recorder and other participants to track the emerging set of priorities.

Soliciting and managing questions or issues

Generating questions and issues. Different methods can be applied to generate questions or issues. We encourage workshop participants to think widely and consult with those outside their particular expertise. We suggest seven means of generating material that are neither mutually exclusive nor exhaustive: (i) reflection by individual workshop participants, (ii) reviews of the peer-reviewed and gray literature by individual workshop participants, (iii) informal discussions between workshop participants and colleagues, (iv) use of email, blogs, tweets, Facebook, and other electronic mechanisms for social networking, (v) facilitating a workshop with colleagues, (vi) assigning students to generate material as a class assignment, and (vii) an interactive website. A range of other possibilities for generating issues, such as the use of scenarios (Sutherland 2006), might merit further consideration.

The range of possible foci for questions includes underlying scientific understanding, scientific methods, the nature or magnitude of effects of a given driver on a given response, effectiveness of human interventions and optimization of human interventions.

Scope of the questions. In our experience defining the scope of questions is quite challenging. Broader questions can attract support from a greater proportion of participants. There is also a group tendency to merge or expand questions. However, such broader questions are likely to be of less interest to researchers and policy makers. This challenge can be overcome in a number of ways. One way is to attempt to define level of the effort needed to answer a question. For example,

we have used as a criterion the concept that a research team or program supported by a funder might have a reasonable probability of answering the question.

In the *USA questions*, it was decided to give concise questions followed by brief explanations, caveats or examples.

Winnowing questions. We have used diverse methods to identify the set of priority questions or issues from among a set that often is much larger. In *UK questions*, we winnowed questions by consensus after individual participants identified priority issues. We have also used multiple rounds of formal voting and refinement. In each exercise, we have allowed participants to reinstate questions or issues that were eliminated in earlier rounds.

It can be helpful to set a target for the number or proportion of submitted questions that will be discussed during the workshop. Time at the workshop is more effectively spent on discussing and refining a relatively small group of high quality questions than on removing a relatively large number of questions that do not meet the established criteria. For example, participants in the *Global question* exercise were allowed to retain more than 100 questions in advance of the workshop and thus the workshop started with 1655 questions.

In *UK questions*, we wished to identify the questions that policy makers considered as highest priority. Accordingly, only the policy makers voted. In subsequent exercises, both policy makers and researchers voted.

Using the web. Both *USA questions* and *Canadian questions* used a simple web interface to collect candidate questions. We encouraged all contributors of questions who voluntarily provided their email address to further distribute the solicitation opportunistically to their colleagues.

Web-based question collection has three potential advantages relative to email collection. First, web-based collection reduces manual labour for the research team. Questions could be downloaded, edited slightly as needed, and posted on a website for viewing with < 20 min of work every several days. Second, access codes embedded in web-links in calls for questions that are sent to participants via email can be used to track the number of questions derived from specific invitations. Such tracking allowed us to conduct targeted follow-up when we suspected our key contacts had not distributed the solicitation. Third, the website can provide supplemental information about the exercise and collect basic demographic information (e.g. organizational affiliation, years of professional experience) from contributors.

As with all social survey methods, organizers must be attentive to issues of participant risk and research ethics. Data collection websites should clearly specify the intent of the research and how respondent contributions will be used. The website design should ensure that anonymity can be maintained and that personal information will not be collected from the respondent's computer or associated with any of the questions. The data collection website for *USA questions* and *Canadian questions* was hosted at Memorial University of Newfoundland and the question solicitation 'survey' approved by the University's Research Ethics Board.

Voting systems

The major change we made in running these exercises as a result of our experience was to improve the rigour of the voting systems. In our first exercise, *UK questions*, participants applied coloured stars to a printed list of questions in the presence of others during the start of each workshop session. In some later exercises, the first round of voting was conducted independently by the participants before the

workshop and individuals assigned each question a score from, say, 10 (definitely retain) to 1 (definitely discard). Such approaches are both more rigorous than applying stars and give more time for discussion at the meeting. We have used two different methods to reduce the set of questions during workshops. The first is to give each issue a score (often 1–10 scale). The second is to allow each participant to indicate whether each issue should be removed or retained. In either case, the number of votes to retain was summed. If there is a very large number of issues of varying quality then an advantage of voting to retain issues is that all issues with no votes can be removed. Decisions about which issues to retain typically were based on the mean scores. Nevertheless, we think the mean rank used in Sutherland *et al.* (2011) is probably a better measure as it is less influenced by variation in voting behaviour. During face-to-face workshops, a show of hands was sometimes used to decide whether to retain a given issue.

Electronic voting was always confidential. A show of hands has the considerable advantage of typically taking about 30 s whereas collating scores takes tens of minutes. It might be possible to implement automated voting systems.

Disseminating results

Transparency and democracy. Aiming for openness and transparency among the participants gives the final output more legitimacy. Mechanisms to increase transparency include ensuring every question with one or more votes for retention is considered at the workshop (*Global questions*) and circulating the questions or issues that have been tentatively excluded and allowing some to be reconsidered. To enhance participation and transparency, we typically compile workshop outputs within the outline of a manuscript (using Microsoft Word) and then circulate this draft manuscript to all participants via email. Each participant then has the opportunity to refine workshop outputs and build the manuscript by contributing their thoughts, with additions, revisions and comments marked using the 'track changes' function. To maintain version-control, individuals must announce to all others if they are about to edit and then circulate the revised manuscript to all participants via email once their edits are complete. The main organizer occasionally incorporates the edits and sometimes suggests solutions or ask someone to research and resolve an issue. This process means that all changes are visible to all. Major changes are usually stated in the email accompanying an individual's edits so that others may respond if they wish.

Participant ownership. The expectation that participants will co-author major products (e.g. publications in high profile peer-reviewed journals) has several advantages. Authorship can promote ownership in the exercise, which enhances the quality of the outputs and increases the likelihood that these outputs will be pursued and answers used to inform policy. Diverse authorship demonstrates to the research and policy communities that the work is genuinely collaborative. Furthermore, diverse authorship can enhance the credibility of the exercise by illustrating the breadth of participants.

Discussion

ARE PRIORITY-SETTING EXERCISES EFFECTIVE?

Several of the exercises have attracted considerable interest among researchers and the media. The *UK questions* article was the most downloaded paper ever from any British Ecological Society journal and the third most downloaded paper from

any of Blackwell's 850 journals in 2006. It attracted substantial publicity (for example a full page in the *Guardian*, one of the four mainstream UK newspapers). The *Global questions* article was the most downloaded paper in *Conservation Biology* in 2009. The UK Marine Science Strategy <http://www.defra.gov.uk/environment/marine/documents/science/mscc/msccstrategy.pdf> identified priority research questions in each of its three themes that it acknowledged were based on *UK questions* (Sutherland 2006).

It is premature to assess the effectiveness of these exercises in terms of specific research and policy and it is notoriously difficult to assess the real impact of papers (Lane 2010), especially as the conversion from science to specific policy initiatives may be slow (Cash *et al.* 2003). It is necessary to consider the different types of research impact and the relationship between them. 'Conceptual impact' is indirect, influencing the way policy makers think (Weiss 1977), and a necessary precursor to 'instrumental impact', the direct and longer-term policy changes arising from research. Research that is successfully transferred to decision makers informs the debate on policy alternatives with a view to concrete action (Albaek 1995). It can help policy makers discuss 'what's known to be true' and filter out policy options that have a low likelihood of success (Rigby 2005). As a consequence, the visible manifestation of research in policy can be limited while the contribution of science to policy is potentially more important, but less concrete, than commonly thought (Albaek 1995).

Potential impacts on policy should be apparent from the definition and framing of issues, the articulation of possible policy alternatives, and the choices of policy makers among various policy alternatives (Cash *et al.* 2003). Shulha & Cousins (1997) found that quality of methods, the credibility of researchers and organizations, political context and the involvement of decision makers in the research process positively influenced research uptake and impact. The participatory nature of our exercises may help increase their impact.

The initial interest in the exercises suggests that they are encouraging the discussion and thought about conservation issues and may be having quite rapid conceptual impact. Although it is unclear whether this interest will translate into instrumental impacts, we are hopeful that the interest indicates the potential for specific and direct impact on conservation science and policy over time.

CAVEATS

Any priority-setting exercise is the product of the people who participate. The results can be influenced by the interests present. Furthermore, individuals may have agendas. However, a diverse and moderately large group, clear criteria, and a democratic process all help reduce the impact of any individual. In *UK questions*, we avoided pressures to rank the final set of issues; we believed ranking would increase the pressure to create broad questions. In *Canadian questions* and *USA questions*, however, we are conducting rankings after publication of the lists of questions. We are testing whether there are differences in priority within and among sectors (e.g. potential differences

in ranking between social and natural scientists or between operational and senior level civil servants).

MECHANISMS FOR COMMUNICATION

Electronic communication seems to be effective for initial solicitation of research questions and for refining questions after the face to face workshop. The human contact facilitated by the workshop itself, however, has proven important for identifying core issues, initiating the process of refining questions, sharing perspectives, building trust and understanding and providing a sense of ownership in the questions and products.

POSSIBLE DEVELOPMENTS

Our experience is that the priority-setting process succeeds in engaging policy makers. The aim of many exercises has been to identify the research needs of policy makers and they are typically pleased to be invited to participate. We believe that it would be useful to apply such methods to policy areas outside conservation, such as social policy or poverty alleviation. We have typically carried out exercises linking policy and research, but similar methods could also be applied among researchers or among policy makers to identify their priorities or to identify possibly interventions to deal with problems (Jacquet *et al.* in press).

Our exercises have been either global or based in the United Kingdom on North America. Exercises are being carried out or planned in other parts of the world, such as the Alps, Galapagos Islands and Israel and for other subjects, such as mineralogy and the study of the science-policy interface. We thus expect that our methods can be applied widely, with appropriate modification, across different communities and for different subjects.

Acknowledgements

Thanks to British Trust for Ornithology, Defra, Kresge Foundation, Natural England, NERC, UK Foresight, and Social Science and Humanities Research Council (Canada) for funding these exercises. We thank the numerous individuals who have participated in the various exercises, many of whom provided feedback on the process. William Sutherland is funded by Arcadia. Murray Rudd was funded by the Canada Research Chairs program. Mark Burgman, an anonymous referee and an associate editor all made very useful comments.

References

Albaek, E. (1995) Between knowledge and power: utilization of social science in public policy making. *Policy Sciences*, **28**, 79–100.

Anderson, I. (2007) *Foot and Mouth Disease. A Review and Lessons Learned*. The Stationary Office, London.

Armstrong, P.R., Armstrong, A.N., Compton, N., Cottle, P. & Davies, I. (2010) The ecological research needs of business. *Journal of Applied Ecology*, **2**, 235–243.

Cash, D.W., Clark, W.C., Alcock, F., Dickson, N.M., Eckley, N., Guston, D.H., Jager, J. & Mitchell, R.B. (2003) Knowledge systems for sustainable development. *Proceedings of the National Academy of Sciences of the United States of America*, **100**, 8086–8091.

European Environment Agency (2001) *Late Lessons from Early Warnings: The Precautionary Principle 1896–2000*, Environmental Issue Report No 22. Office for Official Publications of the European Communities, Copenhagen.

Flynn, T.N., Louviere, J.J., Peters, T.J. & Coast, J. (2007) Best-worst scaling: what it can do for health care research and how to do it. *Journal of Health Economics*, **26**, 171–189.

Government of Canada (2007) *Mobilizing Science and Technology to Canada's Advantage*. Industry Canada, Ottawa.

Holmes, J. & Clark, E. (2008) Enhancing the use of science in environmental policy-making and regulation. *Environmental Science & Policy*, **11**, 702–711.

Jacquet, J., Boyd, I., Carlton, J.T., Fox, H., Johnson, A.E., Mee, L., Roman, J., Spalding, M. & Sutherland, W.J. (in press) Scanning the oceans for solutions. *Solutions*.

Karl, H.A., Susskind, L.E. & Wallace, K.H. (2007) A dialogue, not a diatribe: effective integration of science and policy through joint fact finding. *Environment*, **49**, 20–34.

Lane, J. (2010) Let's make scientific metrics more scientific. *Nature*, **464**, 488–489.

Lawton, J.L. (2007) Ecology, politics and policy. *Journal of Applied Ecology*, **44**, 465–474.

Morton, S.R., Hoegh-Guldberg, O., Lindenmayer, D.B., Harriss Olson, M., Hughes, L., McCulloch, M.T. *et al.* (2009) The big ecological questions inhibiting effective environmental management in Australia. *Austral Ecology*, **34**, 1–9.

Pretty, J. (2009) Speaking truth to power: foot and mouth and the future of agriculture and its communities. *The Social and Cultural Impact of Foot-and-Mouth Disease in the UK in 2001* (eds M. Döring & B. Nerlich), pp. 245–257. Manchester University Press, Manchester.

Pretty, J., Sutherland, W.J., Ashby, J., Auburn, J., Baulcombe, D., Bell, M. *et al.* (2010) The top 100 questions for global agriculture and food. *International Journal of Agricultural Sustainability*, **8**, 219–236.

Pullin, A.S., Knight, T.M. & Watkinson, A.R. (2009) Linking reductionist science and holistic policy using systematic reviews: unpacking environmental policy questions to construct an evidence-based framework. *Journal of Applied Ecology*, **46**, 970–975.

Reed, M.S., Graves, A.R., Dandy, N., Posthumus, H., Hubacek, K., Morris, J., Prell, C., Quinn, C.H. & Stringer, L.C. (2009) Who's in and why? Stakeholder analysis as a prerequisite for sustainable natural resource management. *Journal of Environmental Management*, **90**, 1933–1949.

Rigby, E. (2005) Linking research and policy on Capitol Hill: insights from research brokers. *Evidence & Policy*, **1**, 195–214.

Rudd, M.A., Beazley, K., Cooke, S.J., Fleishman, E., Lane, D.E., Mascia, M.B. *et al.* (in press) Generation of priority research questions to inform conservation policy and management at a national level. *Conservation Biology*.

Sarewitz, D. & Pielke, J.R.A. (2007) The neglected heart of science policy: reconciling supply of and demand for science. *Environmental Science & Policy*, **10**, 5–16.

Shulha, L.M. & Cousins, J.B. (1997) Evaluation use: theory, research, and practice since 1986. *Evaluation Practice*, **18**, 195–208.

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. & Woodroof, H.J. (2009) The need for environmental horizon scanning. *Trends in Ecology and Evolution*, **24**, 523–527.

Sutherland, W.J., Pullin, A.S., Dolman, P.M. & Knight, T.M. (2004) Mismatches between conservation science and practice. *Trends in Ecology and Evolution*, **19**, 565–566.

Sutherland, W.J., Armstrong-Brown, S., Armsworth, P.R., Brereton, T., Brickland, J., Campbell, C.D. *et al.* (2006) The identification of one hundred ecological questions of high policy relevance in the UK. *Journal of Applied Ecology*, **43**, 617–627.

Sutherland, W.J., Bailey, M.J., Bainbridge, I.P., Brereton, T., Dick, J.T.A., Drewitt, J. *et al.* (2008) Future novel threats and opportunities facing UK biodiversity identified by horizon scanning. *Journal of Applied Ecology*, **45**, 821–833.

Sutherland, W.J., Adams, W.M., Aronson, R.B., Aveling, R., Blackburn, T.M., Broad, S. *et al.* (2009) An assessment of the 100 questions of greatest importance to the conservation of global biological diversity. *Conservation Biology*, **23**, 557–567.

Sutherland, W.J., Albon, S.D., Allison, H., Armstrong-Brown, S., Bailey, M.J., Bereton, T. *et al.* (2010a) The identification of priority opportunities for UK nature conservation policy. *Journal of Applied Ecology*, **47**, 955–965.

Sutherland, W.J., Clout, M., Côté, I.M., Daszak, P., Depledge, M.H. *et al.* (2010b) A horizon scan of global conservation issues for 2010. *Trends in Ecology and Evolution*, **25**, 1–7.

Sutherland, W.J., Bardsley, S., Bennun, L., Clout, M., Côté, I.M., Depledge, M.H. *et al.* (2010). A horizon scan of global conservation issues for 2011. *Trends in Ecology and Evolution*, **26**, 10–16.

Weiss, C.H. (1977) Research for policy sake: the enlightenment function of social research. *Policy Analysis*, **3**, 521–545.

Received 1 June 2010; accepted 11 November 2010
Handling Editor: Jana M McPherson