

Ans Vercammen*, Mark Burgman

Centre for Environmental Policy, Imperial College London, Weeks Hall, 16-18 Prince's Gardens, South Kensington SW7 1NE, U.K.

*email ans.vercammen15@imperial.ac.uk

RUNNING HEAD: Collective intelligence

KEYWORDS: decision science , teams , group cognition , swarm intelligence , crowdsourcing , conservation psychology

ARTICLE IMPACT STATEMENT: Applying principles of collective intelligence could improve conservation and environmental decision making.

Abstract

Environmental decisions are often deferred to groups of experts, committees, or panels to develop climate policy, plan protected areas, or negotiate trade-offs for biodiversity conservation. There is, however, surprisingly little empirical research on the performance of group decision making related to the environment. We examined examples from a range of different disciplines, demonstrating the emergence of collective intelligence in the elicitation of quantitative estimates, crowdsourcing applications, and small-group problem solving. We explored the extent to which similar tools are used in environmental decision making. This

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/cobi.13335](https://doi.org/10.1111/cobi.13335).

revealed important gaps (e.g., a lack of integration of fundamental research in decision-making practice, absence of systematic evaluation frameworks) that obstruct mainstreaming of collective intelligence. By making judicious use of interdisciplinary learning opportunities, collective intelligence can be harnessed effectively to improve decision making in conservation and environmental management. To elicit reliable quantitative estimates an understanding of cognitive psychology and to optimize crowdsourcing AI tools may need to be incorporated. The business literature offers insights into the importance of soft skills and diversity in team effectiveness. Environmental problems set a challenging and rich testing ground for collective-intelligence tools and frameworks. We argue this creates an opportunity for significant advancement in decision-making research and practice.

Introduction

Judgment and decision making are integral to conservation and environmental management, but the processes are complex and error-prone, in part due to the inherent uncertainty of social-ecological systems (Schick et al. 2017) and because decision makers have cognitive limitations. Under certain conditions, groups can deliver more accurate factual estimates and creative solutions to problems, referred to as collective intelligence (CI) (Malone & Bernstein, 2015). Democratic elections, jury decisions, and recruitment selection panels are founded on the idea that collectives have access to greater problem-solving resources (Landemore 2012), especially if processes for eliciting and aggregating information, beliefs, and preferences are effective (Hastie & Kameda 2005). Disciplines, including organizational psychology and economics, have long studied group dynamics, but scientific interest in CI has flourished only recently. Enhanced group judgements are reported in medical diagnostics (Mayo & Woolley 2016), taxonomic classification (Prestopnik & Crowston 2012), and meteorological forecasts (Hueffer et al. 2013). There has

been little recognition of the opportunities for collective intelligence in conservation or environmental management.

Collective intelligence applies to “groups of individuals acting together in ways that seem intelligent” (Malone & Bernstein 2015) and covers a range of practices and tools. We evaluated its potential to enhance conservation and environmental decision making. We considered the strengths and weaknesses of 3 specific applications (Fig. 1) and aimed to set a research agenda. We examined wisdom-of-the-crowd effects (specifically how quantitative estimates are improved through structured elicitation) and investigated how crowdsourcing is used to gather distributed and diverse inputs to solve complex problems. We also considered CI as an emergent property in teams making collaborative decisions and reviewed the conditions under which this may occur. If judiciously applied, CI could be deployed more effectively in environmental decision making .

Improving quantitative judgements

Decision makers often have inadequate information, but accurate estimates can be achieved by aggregating individual judgements, most straightforwardly by simple averaging (Armstrong 2001). Wise crowd judgement relies on a large set of diverse and independent opinions in which random errors cancel out to reveal underlying information (Galton, 1907; Surowiecki 2005).

Yet many real-world decision-making contexts are messy, and the simple rules of crowd wisdom may not apply. People use shared resources to inform their opinions, making use of multiple cues that are often spatially and temporally correlated and vary in reliability. Large groups are subject to biasing effects of correlated information (Kao & Couzin 2014). Decision making may also go awry when social interactions among participants, such as dominance, trigger individuals to revise their judgments, ultimately leading to convergence

without accuracy improvement (Lorenz et al. 2011). The impact is more notable in centralized social networks (Becker et al. 2017). Aggregated judgment accuracy improves when individual estimates are negatively correlated, suggesting that divergent opinions enhance collective judgment (Davis-Stober et al. 2014).

Individual contributions may be weighted based on criteria, such as stated confidence or level of expertise (e.g. Prelec et al. 2017). Unfortunately, there is little relationship between people's confidence and their accuracy (Burgman et al. 2011). Those considered experts by others typically provide confidence bounds that are too narrow (Soll & Klayman 2004), leading to poor judgement on all but the simplest questions (Griffin & Tversky 1992). The best predictor of accuracy is prior performance on questions of a similar kind, irrespective of experience, qualification, or training (McBride et al. 2012). Implementing such testing procedures is not always feasible due to resource or time constraints or the difficulty of obtaining and validating test questions.

While aggregation can minimize random error, systematic bias may arise from entrenched, value-based positions or anchoring on previous judgements (e.g., Bosetti et al. 2017). Such cognitive biases are pervasive and occur most often when people use mental shortcuts (heuristics) rather than engaging in in-depth processing (Tversky & Kahneman 1974). When applied ineffectively, heuristics distort judgements; this affects group judgements too because standard aggregation techniques do not correct bias (e.g., Simmons et al. 2011).

Success of mitigation approaches vary (Montibeller & von Winterfeldt 2015). Education in logic and probability theory sometimes enhances judgments (Larrick 2004). Even a single training session based on innovative “serious games” persistently reduces judgement bias (Morewedge et al. 2015). However, motivational incentives have little effect (Camerer & Hogarth 1999). Changing the problem presentation can generate more accurate

judgements, e.g. by restating probabilistic questions as natural-frequency problems (Gigerenzer & Hoffrage 1995). Overprecision can be reduced by asking for bounds before eliciting a best estimate and requiring respondents to estimate how confident they are (Speirs-Bridge et al. 2009). Deliberate practice and unambiguous and immediate feedback are also helpful (Fischhoff 1982). Socially mediated biases, such as group-think and dominance effects, can be managed through well-structured interaction (e.g. Schultze et al. 2012), and feedback about others' estimates improves individual judgments (Wintle et al. 2012).

Structured elicitation methods have been developed that offer multiple strategies to reduce bias and improve accuracy. In conventional Delphi protocols, experts interact through a facilitator who provides feedback about others' estimates. The aim is to reach consensus rapidly (Rowe & Wright 2001), but accuracy is not guaranteed (Murphy et al. 1998). Improving on this, the IDEA protocol uses guided social interactions to avoid the biasing elements of group deliberation and behavioral aggregation (Hanea et al. 2017). Participants provide individual estimates before receiving anonymized information about peer judgments. Ensuing group discussion introduces new information and reconciles differences in understanding. In a second round, individual, anonymous estimates are averaged. The method generates relatively accurate judgements that usually improve with performance-based weighting (Hanea et al. 2018).

Appropriately managed, aggregate estimates from many contributors (experts or not) typically surpass those of conventionally knowledgeable individuals. Most real-life problem settings, although complex and multifaceted, require some element of judgement and are informed by factual estimates. This suggests great potential in harnessing this form of CI to provide more accurate and reliable decision making.

Distributed processing

Complex problem solving requires more than collating and estimating facts; it assumes active coordination of cognitive resources and mental activity. Even cognitively simple animals (e.g., ants) act collectively to undertake complex tasks and solve problems that are intractable for individuals (Krause et al. 2010). Social learning (i.e., learning from observation of others' behavior) probably plays an important role in this phenomenon in animal collectives (Kao et al. 2014) as well as some human decision-making groups (e.g. Kurvers et al. 2014). In "swarm intelligence," a population of unintelligent or uninformed agents, each following simple rules, interacting locally, may produce intelligent global behaviour without the need for centralised control.

This principle has been used to amplify the intelligence of human groups by connecting networked individuals through an online interface that is moderated by artificial intelligence (AI) algorithms (Rosenberg 2016). Artificial swarm intelligence (ASI) enables groups of people to answer questions, make predictions, express opinions, and reach decisions as a unified emergent intelligence by tracking group members as they signal their intent toward choice alternatives. The group's decision is dynamic, representing real-time negotiation among group members collectively exploring the decision space and converging upon the most agreeable answer. Artificial swarm intelligence outperforms medical experts and machine-learning algorithms and makes relatively accurate predictions for financial markets and the outcomes of sporting events (Rosenberg et al. 2017, 2018).

Despite the success of AI-supported systems, the dominant form of crowdsourcing still relies on some central control. Originally, crowdsourcing enabled tasks once performed by a single expert or a small group to be executed through a wide, undefined network of individuals, connected via the web (Howe 2006). A vast range of applications, from simple, repetitive tasks to multifactorial problem solving and design innovation have since been explored. Typical crowdsourcing marketplaces (e.g., Amazon's Mechanical Turk [AMT])

connect requesters with thousands of potential workers to complete so-called human intelligence tasks rapidly and in parallel. The Lego Company, for instance, launched a public call to design new products, rewarding successful creators with a 1% royalty on net revenue. Many applications (e.g., Google Earth) rely on intrinsic public-good motivation to mobilise distributed knowledge. Its success extends to scientific advancement (e.g., a new protein structure discovered through an online game; Cooper et al. 2010).

Although most crowdsourcing is collaborative, distributed knowledge can also be aggregated efficiently in competitive settings. In prediction markets, participants trade the probabilities of the outcomes of events, receiving pay outs when events occur. The market price reflects a consensus forecast of the underlying event probability, which is typically more accurate than probabilities gathered through conventional polls and surveys (Paton et al. 2009). Prediction markets effectively forecast, for example, election results (Rothschild 2009), company valuations (Berg et al. 2009), and spread of infectious disease (Polgreen et al. 2007).

Despite its uptake in commercial and public sectors, there exists only a rudimentary understanding of the intellectual gains crowdsourcing may achieve and under which conditions this happens (Zhao & Zhu 2014). Some evidence suggests that a traditional microtasking approach has potential for significant intelligence amplification. Kosinski et al (2012) crowdsourced the completion of a nonverbal intelligence test by AMT workers and aggregated responses by majority vote. They found that collective intelligence increases with the size of the crowd, although gains are marginal in groups with >6 members, a finding replicated by Verecammen et al. (2019).

Yet, crowdsourcing platforms do not automatically enhance collective intelligence (Guth & Brabham 2017). Crowdsourcing can produce problematic or ineffectual solutions (Greengard 2011) and may not be egalitarian (Brabham 2012), or it may produce unbalanced

views that do not reflect the majority's voice (e.g., Wikipedia entries [Lee & Seo 2016]).

Because contributors' efforts cannot be observed directly and individual accuracy is difficult to monitor due to volume of data, potential free riding, malicious, or manipulative activities need to be curtailed, for example, through motivational tools, such as reputation metrics (Allahbakhsh et al. 2013). Crowdsourcing complex or creative products may require active promotion of collective learning and tailored feedback (e.g. Cullina et al. 2015). For instance, output quality on a creative task improves when workers are informed of other contributors' efforts and their rationales (Xiao 2014).

To maximise the utility of crowdsourcing as a support tool further research is required, focusing on the management of contributors' behaviour, quality-control measures, incentive schemes, and systematic evaluation of crowdsourcing performance.

Small-group collaboration

The small, collaborative consensus-seeking group is the most popular forum for decision making. Groups often outperform the average and the best-regarded individual within a group (Hill 1982). However, "group think" may compromise decisions by introducing dominance effects and correlated judgements. Cases, such as the Bay of Pigs invasion (Janis 1982), highlight that perceived expertise, social status, cognitive and motivational biases, differences in personality and thinking styles, and social processes all potentially influence a group's ability to act intelligently. So, can one measure and predict a group's collective intelligence? The psychometric concept of intelligence is based on the finding that a single underlying dimension, *g*, (general intelligence), explains 40-50% of individual variability in test performance across a range of cognitive domains (e.g., numeracy, working memory, visuospatial skills). General intelligence can be measured reliably, varies among individuals, but is relatively constant over time within an individual

(Spearman 1904). It also correlates well with scholastic achievement, job performance, and other indicators of success, consistent with the concept of intelligence as an “adaptive capacity” (Nisbett et al. 2012).

To examine whether a group equivalent of the individual g exists, Woolley et al (2010) assessed small, randomly assembled groups on tasks relevant to group performance (e.g., creative thinking, decision making, rational judgments under uncertainty). A single underlying dimension explained 43% of group-performance variability, which the authors took as evidence of a collective intelligence factor (c). When the same groups were challenged on a strategy game or an architectural design challenge, c predicted group achievement, whereas individual intelligence did not, suggesting the problem-solving capacity of groups does not depend directly on intellectual ability of individual group members.

Group performance appears to depend more on soft skills than individual cognitive acumen. The single best predictor of collective intelligence is average social perceptiveness, or group members’ ability to correctly identify and appropriately respond to ones’ own mental states and those of others, a skill linked to empathy (Woolley et al. 2010). Even a few group members with low social perception can adversely affect collective intelligence (Engel et al. 2014).

Diversity of identity (e.g., demographic differences) and disposition (differences in problem-solving approaches and heuristics) drive CI in small groups (Hong & Page 2004). An intermediate level of diversity in thinking styles appears to be particularly advantageous (Aggarwal et al. 2015). Too much diversity within groups hampers communication through lack of trust and respect (Jackson et al. 2003). Within-group communication patterns are also relevant; high-performing online collaborations typically exhibit high volumes of interaction and more equal conversational turn taking (Engel et al. 2014; Aggarwal et al. 2015). Groups

that synchronize their activities to allow for more immediate feedback also tend to outperform those that do not use these strategies (Kim et al. 2017).

Some qualitative aspects of group communication also improve group CI. To solve difficult problems, groups must develop a shared mental model (e.g., Maciejovsky et al. 2013), and the more diverse the shared information, the better the group does (Riedl & Woolley 2016). This emerges in so-called hidden profile tasks, where some relevant information is known to most group members, whereas other facts are available only to individuals. A bias toward discussing shared information and avoiding unique or private information contributes to pooling distributed facts (Stasser & Titus 1985), which may have important consequences (Lu et al. 2012). This suggests sharing of rare information should be incentivised so that collaborative groups benefit from all available knowledge (Tausczik & Boons 2018).

Although some argue that coordination cost associated with group problem solving may nullify the potential intellectual gain (Bates & Gupta 2017), well-managed teams have substantial CI. To enable reliable engineering of CI-supportive conditions, more systematic research is required into effects of compositional features (e.g., individual cognitive ability, empathic skill) and operational characteristics (e.g., opportunity for participation, group communication patterns) in teams working on real-world problems.

Collective intelligence in conservation

Expert judgements are indispensable when data are scarce and decisions about socioecological systems are complex and pressing (Martin et al. 2012). The use of structured elicitation methods may increase the rigour with which decisions are made, aid the management of uncertainty, and mitigate prevalent and persistent biases that undermine judgement. Despite successful applications in, for example, threatened species assessments

(McBride et al. 2012), prioritizing management strategies (e.g. Carwardine et al. 2019), risk assessments (Smith et al. 2015), and estimating population trends (Adams- Hosking et al. 2016), most researchers do not use a structured approach (Drescher & Edwards 2019). Expert-elicitation protocols require testing in a wider range of environmental decision making (Hemming et al. 2017), including in the harnessing of collective local ecological knowledge in groups without a strong numerical background or with different knowledge systems (Mantyka-Pringle et al. 2017). More research is also required to establish who should be consulted, the required number of experts, methods for combining judgments, techniques for training and feedback, and tools for independent verification of expert judgments (Martin et al. 2012). Evidence of how quickly expert judgments can be derailed by individual cognitive limitations highlights that experimental studies on human cognition in applied settings and with actual decision makers are a priority. Other fields, including education (e.g. Fay & Montague 2015) and medicine (e.g. Reilly et al. 2013), have integrated behavioural science into decision making and have devised and tested quality-control procedures and instructional materials. Methodological and conceptual advances could be made relatively quickly if practitioners and researchers adapt existing resources to the conservation context.

Crowdsourcing applications in conservation are advancing rapidly. For instance, the immensely popular Zooniverse (<https://www.zooniverse.org/>) hosts dozens of citizen science projects that combine the efforts of human volunteers completing research microtasks on anything from counting penguins to classifying galaxies. iNaturalist (www.inaturalist.org) supports over 9000 different projects and has gathered millions spatially explicit observations and species identifications submitted by the general public. These applications are forerunners of an increasingly participatory science model (Tinati et al. 2015).

Beyond collection and aggregation of data, crowdsourcing provides opportunity for innovation. For example, Climate CoLab, an online problem-solving platform, connects over

90,000 people to design and evaluate plans for addressing climate change. Projects may be combined into integrated plans and are finally evaluated by experts for cash prizes. It has so far generated proposals on new technologies, community projects, marketing strategies, businesses, and policies, attracting a global stream of new and returning visitors who generate innovative and high-quality solutions (Duhaime et al. 2015). Climate CoLab has also provided rich content to evaluate the use of these sociocomputational systems to solve other difficult problems (Introne et al. 2013).

Overall, crowdsourcing applications show great promise to scale up initiatives where resources and data are limited. They may also have additional benefits in education and raising awareness, supporting adaptive management, revealing low-frequency events, and improving scientific methods. Further improvements can be expected through integration with machine learning and AI-supported interaction.

However, to reach their full potential, environmental crowdsourcing initiatives must coordinate efforts, make data freely available, improve quality control, and explain how activities are linked to specific scientific objectives (Cox et al. 2015). Innovation-focused crowdsourcing tools may improve their impact by tapping into existing social networks and reinforcing a sense of community. To boost wide uptake and increase impact, contributors should receive personalised feedback and local, accessible, and directly relevant advice (Piccolo et al. 2018).

The main challenge across all forms of crowdsourcing is the lack of systematic evaluation. A better understanding of the mechanisms that support success could also inform how problems can best be structured for crowdsourcing. Existing research suggests platform management, problem decomposition strategies, preferred characteristics of contributors, quality control, incentive setting, management of participant motivations, and clarity on intellectual property rights are all parameters that require further testing (Ghezzi et al. 2017).

Evaluation of these elements should be guided by specific frameworks (Tredick et al. 2017) so that ultimately standards for best practice could evolve into a refined, generalizable, and effective model for problem solving (Biggar 2010).

The empirical study of teamwork has largely been restricted to the realm of business studies and organizational psychology, and there has been little diffusion of this work into conservation. Organizations typically recruit employees based on individual technical skills and knowledge, possibly negating soft skills and group compositional features. This may come at the expense of collective performance, with some initial evidence from the conservation sector indicating potential impacts on judgement accuracy (e.g., Hemming et al. 2018) and decision quality (e.g., Buckingham 2010).

Metrics have been developed that enable organizations to assess the potential for CI in teams (e.g. theory of mind [Bosco et al. 2016], group cognition [Woolley et al. 2010]). To have a real impact in terms of enabling capacity building for CI, these tools must be made widely available, validated against meaningful performance criteria, and tested specifically in environmental decision-making settings. Data from other fields suggest that to enable collective intelligence, organizations will have to consider restructuring communications, studying the formation of formal and informal networks and the roles team members play in each, strategically managing incentives, and promoting an organizational culture of psychological safety in which team members feel confident in expressing personal opinions and alternative views (Edmondson 1999). This shift in conception of what makes an effective conservation professional must be embedded in the curriculum for environmental and conservation courses and become part of organizational recruitment. We acknowledge that this will challenge institutional biases, bureaucratic processes, and accepted norms of power and decision making.

Conclusions

Conservation and environmental decision making is complex and dynamic, providing a rich testing ground for existing and new CI tools. Some techniques (e.g., structured expert elicitation, crowd computing) may improve input into the decision-making process by enhancing data-processing capacity and judgement accuracy. Other tools (e.g., crowd-based design and innovation contests) may add value by enabling more innovative or representative solutions. Recognition of the emergent properties of groups, rather than a focus on individual capacity could improve capacity for problem solving. We challenge researchers and decision makers to draw on lessons learned in other disciplines and to further develop, implement, and evaluate the utility of CI tools for decision making.

Literature Cited

- Adams-Hosking C, et al. 2016. Use of expert knowledge to elicit population trends for the koala (*Phascolarctos cinereus*). *Diversity and Distributions* **22**:249-262.
- Aggarwal I, Woolley AW, Chabris CF, Malone TW. 2015. Cognitive diversity, collective intelligence and learning in teams. Page 232 in TW Malone, and MS Bernstein, editors. *Handbook of collective intelligence*. MIT Press, Cambridge, Massachusetts.
- Allahbakhsh M, Benatallah B, Ignjatovic A, Motahari-Nezhad HR, Bertino E, Dustdar S. 2013. Quality Control in Crowdsourcing Systems: Issues and Directions. *IEEE Internet Computing* **17**:76-81.
- Armstrong JS. 2001. Combining forecasts. Pages 417-439 in Armstrong JS, editor. *Principles of forecasting: a handbook for researchers and practitioners*. Kluwer Academic, Norwell, Massachusetts.
- Bates TC, Gupta S. 2017. Smart groups of smart people: Evidence for IQ as the origin of collective intelligence in the performance of human groups. *Intelligence* **60**:46-56.

Becker J, Brackbill D, Centola D. 2017. Network dynamics of social influence in the wisdom of crowds. *Proceedings of the National Academy of Sciences* **114**:e5070.

Berg J, Neumann GR, Rietz TA. 2009. Searching for Google's Value: Using Prediction Markets to Forecast Market Capitalization Prior to an Initial Public Offering. *Management Science* **55**:348-361.

Biggar J. 2010. Crowdsourcing for the environment: The case of brighter planet. *PLATFORM: Journal of Media and Communication* **2**:15.

Bosco FM, Gabbatore I, Tirassa M, Testa S. 2016. Psychometric Properties of the Theory of Mind Assessment Scale in a Sample of Adolescents and Adults. *Frontiers in Psychology* **7**:566.

Bosetti V, Weber E, Berger L, Budescu DV, Liu N, Tavoni M. 2017. COP21 climate negotiators' responses to climate model forecasts. *Nature Climate Change* **7**:185.

Brabham DC. 2012. The myth of amateur crowds. *Information, Communication & Society* **15**:394-410.

Buckingham S. 2010. Call in the women. *Nature* **468**: 502.

Burgman MA, McBride M, Ashton R, Speirs-Bridge A, Flander L, Wintle B, Fidler F, Rumpff L, Twardy C. 2011. Expert Status and Performance. *PLOS ONE* **6**:e22998.

Camerer CF, Hogarth RM. 1999. The Effects of Financial Incentives in Experiments: A Review and Capital-Labor-Production Framework. *Journal of Risk and Uncertainty* **19**:7-42.

Carwardine J, Martin TG, Firn J, Reyes RP, Nicol S, Reeson A, Grantham HS, Stratford D, Kehoe L, Chadès I. 2019. Priority Threat Management for biodiversity conservation: A handbook. *Journal of Applied Ecology* **56**:481-490.

- Cooper S, Khatib F, Treuille A, Barbero J, Lee J, Beenen M, Leaver-Fay A, Baker D, Popović Z, players F. 2010. Predicting protein structures with a multiplayer online game. *Nature* **466**:756.
- Cox J, Oh EY, Simmons B, Lintott C, Masters K, Greenhill A, Graham G, Holmes K. 2015. Defining and Measuring Success in Online Citizen Science: A Case Study of Zooniverse Projects. *Computing in Science & Engineering* **17**:28-41.
- Cullina E, Conboy K, Morgan L. 2015. Measuring the crowd: a preliminary taxonomy of crowdsourcing metrics. Pages 1-10. Proceedings of the 11th International Symposium on Open Collaboration. ACM, New York.
- Davis-Stoher CP, Budescu DV, Dana J, Broomell SB. 2014. When is a crowd wise? *Decision* **1**:79-101.
- Drescher M, Edwards RC. 2019. A systematic review of transparency in the methods of expert knowledge use. *Journal of Applied Ecology* **56**: 436-449.
- Duhaime E, Olson GM, Malone TW. 2015. Broad participation in online problem solving can influence participants and lead to better solutions: evidence from the MIT Climate CoLab. MIT CCI working paper 2015-001. Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Edmondson A. 1999. Psychological safety and learning behavior in work teams. *Administrative Science Quarterly* **44**:350-383.
- Engel D, Woolley AW, Jing LX, Chabris CF, Malone TW. 2014. Reading the mind in the eyes or reading between the lines? Theory of mind predicts collective intelligence equally well online and face-to-face. *PLOS ONE* **9**:e115212.

- Fay R, Montague NR. 2015. Witnessing your own cognitive bias. A compendium of classroom exercises. *Issues in Accounting Education* **30**:21.
- Fischhoff B. 1982. Debiasing in Kahneman D, Slovic P, and Tversky A, editors. *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge University Press, Cambridge, United Kingdom.
- Galton F. 1907. Vox Populi. *Nature* **75**:450.
- Ghezzi A, Gabelloni D, Martini A, Natalicchio A. 2017. Crowdsourcing: A Review and Suggestions for Future Research. *International Journal of Management Reviews* **20**:343-363.
- Gigerenzer G, Hoffrage U. 1995. How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review* **102**:684-704.
- Greengard S. 2011. Following the crowd. Page 3. *Communications of the ACM*. ACM, New York, NY.
- Griffin D, Tversky A. 1992. The weighing of evidence and the determinants of confidence. *Cognitive Psychology* **24**:411-435.
- Guth KL, Brabham DC. 2017. Finding the diamond in the rough: Exploring communication and platform in crowdsourcing performance. *Communication Monographs* **84**:510-533.
- Hanea AM, McBride MF, Burgman MA, Wintle BC. 2018. The Value of Performance Weights and Discussion in Aggregated Expert Judgments. *Risk Analysis* **38**:1781-1794
- Hanea AM, McBride MF, Burgman MA, Wintle BC, Fidler F, Flander L, Twardy CR, Manning B, Mascaro S. 2017. Investigate Discuss Estimate Aggregate for structured expert judgement. *International Journal of Forecasting* **33**:267-279.

- Hastie R, Kameda T. 2005. The robust beauty of majority rules in group decisions. *Psychological Review* **112**:494-508.
- Hemming V, Burgman MA, Hanea AM, McBride MF, Wintle BC. 2017. A practical guide to structured expert elicitation using the IDEA protocol. *Methods in Ecology and Evolution* **9**:169-180.
- Hemming V, Walshe TV, Hanea AM, Fidler F, Burgman MA. 2018. Eliciting improved quantitative judgements using the IDEA protocol: A case study in natural resource management. *PLOS ONE* **13**:e0198468.
- Hill GW. 1982. Group versus individual performance: Are $N + 1$ heads better than one? *Psychological Bulletin* **91**:517-539.
- Hong L, Page SE. 2004. Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences of the United States of America* **101**:16385.
- Howe J. 2006. The rise of crowdsourcing. *Wired Magazine* **14**: 1-4.
- Hueffer K, Fonseca MA, Leiserowitz A, Taylor KM. 2013. The wisdom of crowds: predicting a weather and climate-related event. *Judgement and Decision Making* **8**:14.
- Introne J, Laubacher R, Olson G, Malone TW. 2013. Solving wicked social problems with socio-computational systems. Working paper. Massachusetts Institute of Technology, Cambridge, Massachusetts.
- Janis IL 1982. *Groupthink*. Houghton-Mifflin, Boston.

Kao AB, Couzin ID. 2014. Decision accuracy in complex environments is often maximized by small group sizes. *Proceedings of the Royal Society B: Biological Sciences* **281**: e20133305.

Kao AB, Miller N, Torney C, Hartnett A, Couzin ID. 2014. Collective Learning and Optimal Consensus Decisions in Social Animal Groups. *PLOS Computational Biology* **10**:e1003762.

Kim YJ, Engel D, Woolley AW, Lin JY-T, McArthur N, Malone TW. 2017. What Makes a Strong Team?: Using Collective Intelligence to Predict Team Performance in League of Legends. *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*. ACM, New York.

Kosinski M, Bachrach Y, Kasneci G, Van Gael J, Graepel T. 2012. Crowd IQ: Measuring the intelligence of crowdsourcing platforms. *Proceedings of the 4th Annual ACM Web Science Conference*. ACM, New York.

Krause J, Ruxton GD, Krause S. 2010. Swarm intelligence in animals and humans. *Trends in Ecology and Evolution* **25**:28-34.

Kurvers RHJM, Wolf M, Krause J. 2014. Humans use social information to adjust their quorum thresholds adaptively in a simulated predator detection experiment. *Behavioral Ecology and Sociobiology* **68**:449-456.

Landemore HE. 2012. Whe the many are smarter than the few and why it matters. *Journal of Public Deliberation* **8**.

Larrick RP. 2004. Debiasing in Koehler DJ, and Harvey N, editors. *Blackwell Handbook of Judgment and Decision Making*. John Wiley & Sons, .

- Lee J, Seo D. 2016. Crowdsourcing not all sourced by the crowd: An observation on the behavior of Wikipedia participants. *Technovation* **55-56**:14-21.
- Lorenz J, Rauhut H, Schweitzer F, Helbing D. 2011. How social influence can undermine the wisdom of crowd effect. *Proceedings of the National Academy of Sciences* **108**:e9020.
- Lu L, Yuan YC, McLeod PL. 2012. Twenty-five years of hidden profiles in group decision making: a meta-analysis. *Personality and Social Psychology Review* **16**:54-75.
- Maciejovsky B, Sutter M, Budescu DV, Bernau P. 2013. Teams Make You Smarter: How Exposure to Teams Improves Individual Decisions in Probability and Reasoning Tasks. *Management Science* **59**:1255-1270.
- Malone TW, Bernstein MS 2015. *Handbook of Collective Intelligence*. MIT Press, Cambridge, MA.
- Mantyka-Pringle CS, et al. 2017. Bridging science and traditional knowledge to assess cumulative impacts of stressors on ecosystem health. *Environment International* **102**:125-137.
- Martin TG, Burgman MA, Fidler F, Kuhnert PM, Low-Choy S, McBride M, Mengersen K. 2012. Eliciting Expert Knowledge in Conservation Science. *Conservation Biology* **26**:29-38.
- Mayo AT, Woolley AW. 2016. Teamwork in health care: Maximising collective intelligence via inclusive collaboration and open communication. *AMA Journal of Ethics* **18**:8.
- McBride MF, Fidler F, Burgman MA. 2012a. Evaluating the accuracy and calibration of expert predictions under uncertainty: predicting the outcomes of ecological research. *Diversity and Distributions* **18**:782-794.

- McBride MF, et al. 2012b. Structured elicitation of expert judgments for threatened species assessment: a case study on a continental scale using email. *Methods in Ecology and Evolution* **3**:906-920.
- Montibeller G, von Winterfeldt D. 2015. Cognitive and Motivational Biases in Decision and Risk Analysis. *Risk Analysis* **35**:1230-1251.
- Morewedge CK, Yoon H, Scopelliti I, Symborski CW, Korris JH, Kassam KS. 2015. Debiasing Decisions: Improved Decision Making With a Single Training Intervention. *Policy Insights from the Behavioral and Brain Sciences* **2**:129-140.
- Murphy MK, Black NA, Lamping DL, McKee CM, Sanderson CF, Askham J, Marteau T. 1998. Consensus development methods, and their use in clinical guideline development. *Health Technology Assessment* **2**: 1-88.
- Nisbett RE, Aronson J, Blair C, Dickens W, Flynn J, Halpern DF, Turkheimer E. 2012. Intelligence: new findings and theoretical developments. *American Psychologist* **67**:130-159.
- Paton D, Siegel Donald S, Vaughan Williams L. 2009. The Growth of Gambling and Prediction Markets: Economic and Financial Implications. *Economica* **76**:219-224.
- Piccolo LSG, De Liddo A, Burel G, Fernandez M, Alani H. 2018. Collective intelligence for promoting changes in behaviour: a case study on energy conservation. *AI & Society* **33**:15-25.
- Polgreen PM, Nelson FD, Neumann GR, Weinstein RA. 2007. Use of Prediction Markets to Forecast Infectious Disease Activity. *Clinical Infectious Diseases* **44**:272-279.
- Prelec D, Seung HS, McCoy J. 2017. A solution to the single-question crowd wisdom problem. *Nature* **541**:532.

- Prestopnik N, Crowston K. 2012. Purposeful gaming and socio-computational systems: a citizen science design case. Pages 75-84. Proceedings of the 17th ACM international conference on Supporting group work. ACM, New York.
- Reilly JB, Ogdie AR, Von Feldt JM, Myers JS. 2013. Teaching about how doctors think: a longitudinal curriculum in cognitive bias and diagnostic error for residents. *BMJ Quality & Safety* **22**:1044.
- Riedl C, Woolley AW. 2016. Teams vs. Crowds: A Field Test of the Relative Contribution of Incentives, Member Ability, and Emergent Collaboration to Crowd-Based Problem Solving Performance. *Academy of Management Discoveries* **3**:382-403.
- Rosenberg L. 2016. Artificial Swarm Intelligence, a Human-in-the-loop approach to A.I. Association for the Advancement of A.I. Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence. ACM, New York.
- Rosenberg L, Pescetelli N. 2017. Amplifying prediction accuracy using swarm A.I. Intelligent Systems. Proceedings of the Intelligent Systems Conference. XXX, London, UK.
- Rosenberg L, Pescetelli N, Willcox G. 2017. Artificial swarm intelligence amplifies accuracy when predicting financial markets. Proceedings of the 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference. IEEE, Piscataway, New Jersey.
- Rosenberg L, Willcox G, Halabi S, Lungren M, Baltaxe D, Lyons M. 2018. Artificial Swarm Intelligence employed to amplify diagnostic accuracy in radiology. Proceedings of the 9th annual information technology, electronics, and mobile communication conference. IEEE, Piscataway, New Jersey.
- Rothschild D. 2009. Forecasting Elections: Comparing Prediction Markets, Polls, and Their Biases. *The Public Opinion Quarterly* **73**:895-916.

- Rowe G, Wright G. 2001. Expert Opinions in Forecasting: The Role of the Delphi Technique. Pages 125-144 in Armstrong JS, editor. Principles of Forecasting: A Handbook for Researchers and Practitioners. Springer US, Boston.
- Schick A, Hobson Peter R, Ibisch Pierre L. 2017. Conservation and sustainable development in a VUCA world: the need for a systemic and ecosystem- based approach. Ecosystem Health and Sustainability **3**:e01267.
- Schultze T, Mojzisch A, Schulz-Hardt S. 2012. Why groups perform better than individuals at quantitative judgment tasks: Group-to-individual transfer as an alternative to differential weighting. Organizational Behavior and Human Decision Processes **118**:24-36.
- Simmons JP, Nelson LD, Galak J, Frederick S. 2011. Intuitive Biases in Choice versus Estimation: Implications for the Wisdom of Crowds. Journal of Consumer Research **38**:1-15.
- Smith M, Wallace K, Lewis L, Wagner C. 2015. A structured elicitation method to identify key direct risk factors for the management of natural resources. Heliyon **1**:e00043.
- Soll JB, Klayman J. 2004. Overconfidence in interval estimates. Journal of Experimental Psychology: Learning, Memory and Cognition **30**:299-314.
- Spearman C. 1904. "General Intelligence," Objectively Determined and Measured. The American Journal of Psychology **15**:201-292.
- Speirs- Bridge A, Fidler F, McBride M, Flander L, Cumming G, Burgman M. 2009. Reducing Overconfidence in the Interval Judgments of Experts. Risk Analysis **30**:512-523.
- Stasser G, Titus W. 1985. Pooling of unshared information in group decision making: Biased information sampling during discussion. Journal of Personality and Social Psychology **48**:1467-1478.

Surowiecki J 2005. The wisdom of crowds. Doubleday, New York.

Tausczik Y, Boons M 2018. Distributed knowledge in crowds: crowd performance on hidden profile tasks. Proceedings of the Twelfth International AAAI Conference on Web and Social Media. AAAI, Menlo Park, California.

Tinati R, Kleek MV, Simperl E, Luczak-Roesch M, Simpson R, Shadbolt N. 2015. Designing for Citizen Data Analysis: A Cross-Sectional Case Study of a Multi-Domain Citizen Science Platform. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. ACM, New York.

Tredick CA, Lewison RL, Deutschman DH, Hunt TANN, Gordon KL, Von Hendy P. 2017. A Rubric to Evaluate Citizen-Science Programs for Long-Term Ecological Monitoring. BioScience **67**:834-844.

Tversky A, Kahneman D. 1974. Judgment under Uncertainty: Heuristics and Biases. Science **185**:1124-1131.

Vercammen A, Ji Y, Burgman MA. 2019. The collective intelligence of random small crowds: A partial replication of Kosinski et al. (2012). Judgment and Decision Making **14**: 91-98.

Wintle Bonnie C, Fidler F, Vesik Peter A, Joslin LM, Armsworth P. 2012. Improving visual estimation through active feedback. Methods in Ecology and Evolution **4**:53-62.

Woolley AW, Chabris CF, Pentland A, Hashmi N, Malone TW. 2010. Evidence for a Collective Intelligence Factor in the Performance of Human Groups. Science **330**:686.

Xiao L. 2014. Effects of rationale awareness in online ideation crowdsourcing tasks. Journal of the Association for Information Science and Technology **65**:1707-1720.

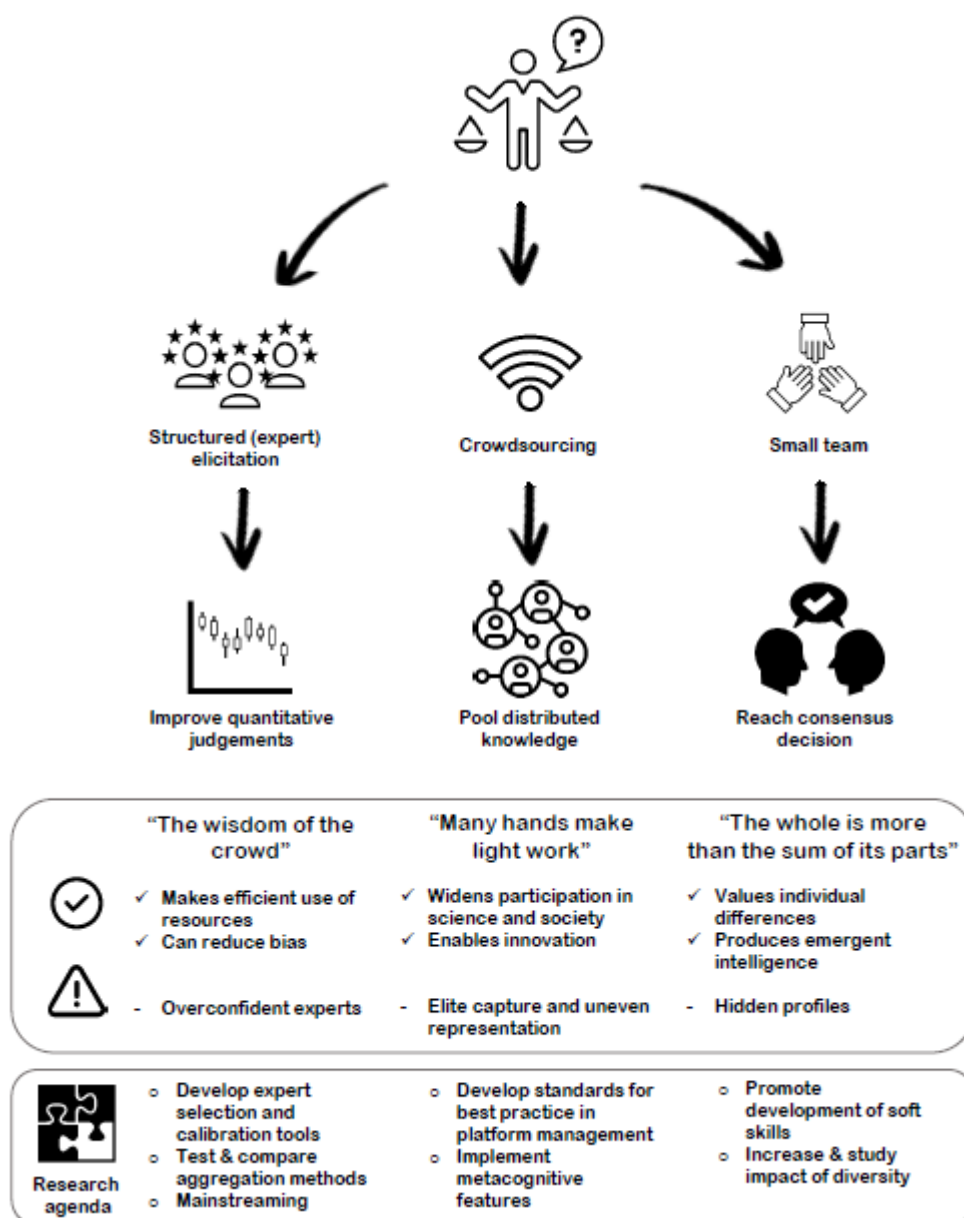


Figure 1. Three forms of collective intelligence and how they may assist in improving decision-support and decision-making processes in conservation.



Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:

Vercammen, A;Burgman, M

Title:

Untapped potential of collective intelligence in conservation and environmental decision making.

Date:

2019-12

Citation:

Vercammen, A. & Burgman, M. (2019). Untapped potential of collective intelligence in conservation and environmental decision making.. *Conserv Biol*, 33 (6), pp.1247-1255.
<https://doi.org/10.1111/cobi.13335>.

Persistent Link:

<http://hdl.handle.net/11343/285714>