

Reply to comment from W.P. Aspinall on "Social studies of volcanology: knowledge generation and expert advice on active volcanoes" by Amy Donovan, Clive Oppenheimer and Michael Bravo [Bull Volcanol (2012) 74:677-689]

Amy Donovan, Clive Oppenheimer, Michael Bravo

► **To cite this version:**

Amy Donovan, Clive Oppenheimer, Michael Bravo. Reply to comment from W.P. Aspinall on "Social studies of volcanology: knowledge generation and expert advice on active volcanoes" by Amy Donovan, Clive Oppenheimer and Michael Bravo [Bull Volcanol (2012) 74:677-689]. Bulletin of Volcanology, Springer Verlag, 2012, 74 (6), pp.1571-1574. 10.1007/s00445-012-0626-9 . insu-00723552

HAL Id: insu-00723552

<https://hal-insu.archives-ouvertes.fr/insu-00723552>

Submitted on 28 Aug 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Reply to comment from W.P. Aspinall on “Social studies of volcanology: knowledge generation and expert advice on active volcanoes” by Amy Donovan, Clive Oppenheimer and Michael Bravo [Bull Volcanol (2012) 74:677-689]

Amy Donovan^{1,2}, Clive Oppenheimer^{1,3,4} and Michael Bravo²

(1) Department of Geography, University of Cambridge, Downing Place, Cambridge, CB2 3EN, UK

(2) Scott Polar Research Institute, University of Cambridge, Lensfield Road, Cambridge, CB2 1ER, UK

(3) Le Studium, Institute for Advanced Studies, Orléans and Tours, Orléans, France (4) Institut des Sciences de la Terre d’Orléans, University of Orléans, 1a rue de la Férollerie, 45071 Orléans Cedex 2, France

While it was not our intention to discuss the expert elicitation procedure in detail in our paper (Donovan et al. 2012), the *Comment* proffered by Aspinall (2012) raises important questions for volcanologists. We are broadly in agreement with Aspinall (2012) that the Cooke Method of expert elicitation provides a means of *increasing* objectivity and quantifying at least some of the uncertainty surrounding risk assessment on active volcanoes. It does also provide a means for accountability and the tracking of scientific assessments, which, as Aspinall notes, is particularly important in litigious times. Risk assessment forces an opening-up of our understanding of scientific reasoning, its social context and its associated uncertainties – not least for the sake of transparency. However, we maintain that the elicitation method is epistemologically, psychologically and probabilistically subjective. This does not detract from its valuable contribution to scientific debate, quantitative risk assessment and the communication of the uncertainty concerning volcanic risk. Indeed, a degree of subjectivity increases the types of knowledge and intuitive expertise that can be applied to the problems of volcanic risk assessment alongside other more “objective” evidence.

Kuhn (1977), referencing Kant (2003, original 1781), argued against the use of “objective” in opposition to “subjective”: they are not strictly opposed and the history of philosophy contains multiple readings of each. Rather, the terms should be regarded as separate; they are neither antonyms nor part of a spectrum. There are different types of knowledge involved in expert advice, and distinguishing between them can be challenging. In the following, we explore the terminology and highlight aspects of the social and scientific context of the elicitation method¹ that may be helpful in its application in volcanic risk assessment.

Epistemology and probability

The question of whether or not human beings are ever capable of objective judgement is one that pervades the history and philosophy of science. It is closely tied to the nature of knowledge, knowing and the knowable – and by inference the nature of uncertainty, ambiguity and ignorance. This epistemology is further complicated by probability, because it is a human construct (e.g. de Finetti 1974) that seeks to apply knowledge to make inferences, either using frequencies or “degrees of belief”. In elicitation, the experts’ knowledge base is the assumed primary source of information for the assessment. The nature of scientific knowledge was of great concern to the logical positivist school of Mach, Russell, Wittgenstein and others, who argued that such knowledge can only be generated by empirical science, and must be testable. According to the great objectivist, Popper (1959, p.22),– “the objectivity of scientific statements lies in the fact that they can be inter-subjectively tested”.

This can be true of statements of knowledge (as make up much of the expert's knowledge base), but knowledge is not the same as belief. In light of this, Gillies (2000) introduced a notion of "intersubjective probability", where a probability is generated as the degree of belief of a group. This may increase the objectivity of a subjective probabilistic method, but raises issues about the nature of the questions and their testability.²

We note that two epistemologically distinct varieties of question have been applied in elicitation on Montserrat. There are questions requiring assessment of physical quantities such as dome collapse volumes, and questions that require the assessment of *probabilities* of particular events. The seed questions, however, only assess quantities (as these have "known" answers). It is not possible for the seed questions to assess the experts' ability to produce probabilities, because probabilities other than 0 or 1 cannot be verified or falsified (unless the system under discussion is limited in some way and a definite frequency is observed). A skill score can be used to demonstrate a poor calibration, but requires a large dataset of prior elicitation for each expert. A single probabilistic assessment that gives a high probability of an event that then does not happen is not "wrong" (unless $P = 1$). This is also where the use of skill scores based on past results to calibrate experts breaks down under the conditions of high uncertainty (and very varied questions) that prevail in volcanic crises. At best, experts can only produce an accurate representation of *their own belief*. Thus, while some correction can be made either for probabilistic reasoning (imperfectly, using a skill score) or for volcanological expertise (using seed questions), neither method can *completely* encapsulate the expert's aptitude for both volcanological issues and probabilistic inference, or their ability to estimate their uncertainty. It may, however, enable a more refined assessment than one without weightings (Aspinall 2010). The key, according to Gillies (2000), is that all the experts are agreed on the final value: the final value gets its authority from a knowledgeable social source.

It is generally accepted in statistics that there is a distinction between frequency-based and belief-based probabilistic methods. A common nomenclature is to refer to the former as "objective" and the latter as "subjective" – though philosopher of science Ian Hacking (2001, p.131) notes that "these are terrible terms, loaded with ideology". Frequentist methods are based on the frequency of particular phenomena, such as tossing a coin, where experiments and data can be used to define this frequency. Subjective probabilities are based on degrees of belief. Many Bayesian methods are regarded as subjective because they generally involve the selection of a prior distribution. Similarly, expert judgements arising from the elicitation process may be based on knowledge of relevant information, but are ultimately expressed as degrees of belief in the likelihood of particular phenomena that have not yet taken place. Thus the discussion is as much about deduction and induction as it is about objectivity and subjectivity.³

As Hacking (2006) points out, both types of probabilistic method seek to evade the problem of induction in science. Elicitation complicates this picture: in the volcanological application, the experts are being asked to make inferences based on their knowledge of the specific volcano and also of other volcanoes – some inductive, some deductive. Experiences and knowledge gained at one volcano may play a part in assessments at a different volcano, generally because of a lack of other knowledge when a volcano awakens. There are thus individual choices about which knowledge applies where. The consideration of the probability, in the next 6 months, say, of a dome collapse of a given volume towards a specified direction, requires the combined interpretation of multiple observables, models and the uncertainty on each – not only in terms of error but also in terms of relevance. Each expert

has a different knowledge base and will attribute importance to the various types of information available to them in different ways. There is therefore an element of what Kuhn refers to as “theory choice” (Kuhn 1977), especially relevant in a multidisciplinary context like volcanic risk assessment. Experts may also be conscious of political, economic and social factors that are linked to the risk assessment and may invite a level of precaution. Additionally, experts may be aware (or not) of quantities that we simply cannot measure and have to make educated guesses about, and related to this is the aleatory uncertainty (also known as the objective uncertainty, in yet another use of the term!) in the natural system.

With such a variety of sources of uncertainty, it is not surprising that its quantification is challenging. Different groups of experts have yielded different distributions during the elicitation procedures on Montserrat, as is acknowledged in the reports (e.g. SAC Second Addendum to the September 2002 Report, MVO January 2003, Appendix 5). This implies that the reproducibility referred to by Aspinall is dependent on the use of the same group of experts and is hence subjective in the normal sense (“pertaining to the subject”, Oxford English Dictionary). Nevertheless, the use of a broad range of experts remains advantageous in increasing the volume of the evidence base being used for the assessment.

Psychology and cognition

A substantial body of relevant work on expert elicitation can be found in the social psychological literature (e.g. O’Hagan et al. 2006; Kynn 2008 and references therein). It focusses on the ability of individuals – experts included – to make probabilistic inferences. It has also revealed a range of biases and heuristics concerning the ways in which inferences are made based on experience, knowledge and the ability to use that knowledge. The terminology of “biases” does not imply prejudiced judgement but rather that the human brain is conditioned to make judgements in particular ways. This can result in subject-specific and not readily quantifiable aspects to the elicitation of probabilities (O’Hagan et al. 2006). In a sense, the experts act as “black boxes” (Latour 1987): while the input and output are known, the path between them is complex and hidden. The seed questions seek to calibrate the black box by its outputs, but cannot open it.

Beyond lies another psychological issue: numbers provide a semblance of knowledge that may not be “real”. This is crucial when asking experts to assess their own uncertainty. The complexity of the natural system has to be considered by an expert, who is also complex, and then from this multiplied complexity comes a single distribution. It is in this context that sociologists have sought to provide frameworks that move beyond the quantification of risk assessments and involve both quantitative methods – such as the Cooke Method – and qualitative methods. This is important in the understanding of indeterminate (ambiguous) uncertainty and “unknown unknowns” (Wynne 1992; Stirling 2007; Spiegelhalter and Riesch 2011): these are not easily quantifiable, and it could be argued that the provision of quantities without adequate narrative explanation and disclaimers is misleading. Hence the communication of elicitation results to the authorities has to be framed within the scientific debates that elicitation seeks (partially) to quantify.

Conclusions

This argument is not intended to discredit the elicitation procedure in any way. Indeed, we have argued that its subjectivity neither detracts from its usefulness nor renders it incapable of

generating a level of “objectivity”. In this respect, there is some gain to be had in combining the use of statistical methods with the use of social scientific methods (and as Foucault argued, statistics – literally, “pertaining to the state” – is the purest form of social science and has its origins in the quest to understand the social). Social science depends on subjective methods because it involves the study of subjects by subjects. Thus, the experts have to consider their own subjectivity in the application of their research to real-world problems (known as “reflexivity”). In a similar way, the use of expert elicitation should be guided by an awareness of its context and its limitations – as it has been on Montserrat – and accompanied by an explanation of the challenges and uncertainties. It does, as Aspinall argues, provide a means of documentation of the scientific reasoning process, and ensures that a diverse range of models, measurements and interpretations can be included in the evidence base. It also reduces the impact of unhelpful personality factors, such as overconfidence, on the results whilst enabling the incorporation of “tacit” and local knowledges, and it provides a formalised means of providing expert advice. Finally, the use of probabilistic methods enables the acknowledgement of uncertainties surrounding the evidence, its interpretation and potential implications. This is a significant advantage over deterministic methods (see Marzocchi and Zechar 2011, for discussions of this issue in the seismological context).

In summary, the expert elicitation procedure, as applied on Montserrat and elsewhere, is an important tool in volcanic risk assessment. It is mathematically elegant, and socially embedded. On Montserrat, it has involved extensive discussions both prior to the elicitation, and after the results have been calculated to ensure they are acceptable to the group. This ensures that there is an auditing process and that the reasoning behind the results is sound: in other words, there is a level of “peer review” involved. The key factors, then, which render elicitation a subjective method are: individual expressions of belief cannot be readily verified or falsified unless $P = 0$ or 1 (or there is a series of values for each expert – and each type of question – that can be scored); different groups of experts produce different results; the uncertainty on the results (and uncertainties!) can never be completely measured; the method is belief-based rather than frequency-based; and the calibration/informativeness values derived from the seed questions cannot fully remove all subject-specific aspects of judgement. Nevertheless, subjective ideas, theories and interpretations may be an integral part of the evidence base for volcanic risk assessments, and a key benefit of elicitation is that the evidence base is broad. The involvement of social scientific expertise might help to refine the method and its framing and provide a broader conceptualisation of risk and its implications for society. Given the complexity of inter-subjectivity between experts, the range of methods, models, experience and data, and the uncertainty associated with each, providing a philosophical account of *objectivity* remains a challenge. Perhaps it is more helpful to focus on the language of impartiality, consistency, and transparency in providing expert advice in volcanic crises.

Acknowledgements

The authors thank Professor David Spiegelhalter and members of the risk and uncertainty reading group in the Centre for Research in the Arts, Social Sciences and Humanities, University of Cambridge, for helpful discussions over the past several years.

References

Aspinall WP (2006) Structured elicitation of expert judgement for probabilistic hazard and

risk assessment in volcanic eruptions. In: Mader HM, Coles SG, Connor C, Connor L (eds) *Statistics in volcanology*. Geological Society of London, London, pp 15–30

Aspinall WP (2010) A route to more tractable expert advice. *Nature* 463:294–295

Aspinall WP (2012) Comment on Donovan et al. 2012 “Social studies of volcanology: expert advice and knowledge generation on active volcanoes”. *Bull Volcanol. this issue*

De Finetti B (1974) *Theory of probability*. Wiley, London

Donovan A, Oppenheimer C, Bravo M (2012) Social studies of volcanology: knowledge generation and expert advice on active volcanoes. *Bull Volcanol* 74(3):677–689

Gillies D (2000) *Philosophical theories of probability*. Routledge, London

Hacking I (2001) *An introduction to probability and inductive logic*. Cambridge University Press, Cambridge

Hacking I (2006) (Second edition; first published 1975). *The emergence of probability*. Cambridge University Press, Cambridge

Kant I (2003) (Original 1781) *Critique of pure reason* (trans: Meikljohn JMD). Dover, New York

Kuhn T (1977) Objectivity, value judgement, and theory choice. In: Kuhn T (ed) *The essential tension: studies in scientific tradition and change*. University of Chicago Press, Chicago, pp 320–329

Kynn M (2008) The ‘heuristics and biases’ bias in expert elicitation. *J R Stat Soc: Ser A (Stat Soc)* 171(1):239–264

Latour B (1987) *Science in action: how to follow scientists and engineers through society*. Harvard University Press, Cambridge

Marzocchi W, Zechar JD (2011) Earthquake forecasting and earthquake prediction: different approaches for obtaining the best model. *Seismol Res Lett* 82(3):442–448

O’Hagan A, Buck CE, Daneshkhah A, Eiser JR, Garthwaite P (2006) *Uncertain judgements: eliciting experts’ probabilities*. Wiley, London

Popper K (1959) (German original, 1935). *The logic of scientific discovery*. Routledge, Abingdon

Popper K (1983) *Realism and the aim of science*. Routledge, London

Russell B (1912) (reissued 2001). *The problems of philosophy*. Oxford University Press, Oxford

Spiegelhalter DJ, Riesch H (2011) Don't know, can't know: embracing deeper uncertainties when analysing risks. *Philos Trans R Soc A: Math, Phys Eng Sci* 369(1956):4730–4750

Stirling A (2007) Risk, precaution and science: towards a more constructive policy debate. *EMBO Rep* 8(4):309–315

Wynne B (1992) Uncertainty and environmental learning: reconceiving science and policy in the preventive paradigm. *Glob Environ Chang* 2(2):111–127

Footnotes

¹ Expert elicitation according to the Cooke Method involves the weighting of expert opinions on the basis of a set of “seed questions” (e.g. Aspinall 2006). The seed questions are devised by the facilitator and have known true values. The experts are calibrated by both statistical likelihood (their accuracy), and informativeness (how certain they are about their answer). Their responses to the actual elicitation questions are then weighted by their scores.

² Popper himself went further and rejected probabilistic methods that are not based on frequencies (or on what he termed “propensity theory”), thereby denying any inductive/epistemic/subjective basis for probability. He emphasised the fallibility and uncertainty of science, and the challenge of assessing a hypothesis based on uncertain evidence using probabilistic methods (e.g. Popper 1983), referring to this as “knowledge out of ignorance”.

³ Bertrand Russell gives an example of the problem of induction: “The man who has fed the chicken every day throughout its life at last wrings its neck instead, showing that more refined views as to the uniformity of nature would have been useful to the chicken.” Russell (1912), p.35.