International Journal of Disaster Risk Reduction 10 (2014) 75–101

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: www.elsevier.com/locate/ijdrr

Uncertainty and decision making: Volcanic crisis scenarios

Emma E.H. Doyle [a,*], John McClure [b], Douglas Paton [c], David M. Johnston [a,d]

[a] Joint Centre for Disaster Research, Massey University, PO Box 756, Wellington 6140, New Zealand
[b] School of Psychology, Victoria University of Wellington, Kelburn Parade, PO Box 600, Wellington 6012, New Zealand
[c] School of Psychology, University of Tasmania, Newnham Campus, Locked Bag 1342, Launceston, Tasmania 7250, Australia
[d] GNS Science, PO Box 30 368, Lower Hutt 5010, New Zealand

ARTICLE INFO

Article history:
Received 9 April 2014
Received in revised form
17 July 2014
Accepted 17 July 2014
Available online 25 July 2014

Keywords:
Decision making
Uncertainty
Probabilities
Forecasts
Evacuations
Profession

ABSTRACT

The impact of uncertainty on Disaster Risk Reduction decision-making has become a pressing issue for debate over recent years. How do key officials interpret and accommodate uncertainty in science advice, forecasts and warnings into their decision making? Volcanic eruptions present a particularly uncertain hazard environment, and to accommodate this scientists utilize probabilistic techniques to inform decision-making. However, the interpretation of probabilities is influenced by their framing. We investigate how verbal or numerical probabilities affect decisions to evacuate a hypothetical town, and reasons given for that decision, based upon a volcanic eruption forecast. We find fewer evacuations for verbal terms than for equivalent numerical terms, and that the former is viewed as more ambiguous. This difference is greater for scientists, which we suggest is due to their greater familiarity with numerical probabilities and a belief that they are more certain. We also find that many participants have a poor understanding of the relationship between probability and time window stated, resulting in an incorrect assessment of overall likelihood and more evacuations for the lower likelihood version of two scenarios. Further, we find that career sector (scientist or non-scientist) influences evacuation decisions, with scientists tending to reduce the uncertainty by focusing on the quality and volume of information provided, while non-scientists tended to either acknowledge or suppress the uncertainty, focusing on actions to take. These findings demonstrate the importance of identifying communication strategies that mitigate different perceptions of forecasts, to both enhance end-user decision making and to prevent premature, delayed, or unnecessary actions.

1. Introduction

Volcanic crises can create conditions of extreme uncertainty for scientists and key officials managing the crisis. A volcano showing signs of unrest may exhibit changes in geophysical signals (such as seismic and geodetic changes), geochemical signals, cause felt earthquakes and result in visual signs of impending activity (such as increased degassing or bulging flanks). While indicative of changes occurring in the volcanic edifice, they do not always lead to eruptions and may be unreliable as indicators of when an eruption may occur [62,66,72]. They thus present a challenging environment for effective response, emergency management planning, and decision management.

Once an eruptive phase has occurred, volcanoes can also go through cycles of quiescence, followed by periods of unrest, periods of activity and then periods of quiescence

* Corresponding author. Tel.: +64 4 801 5799x62458; fax: +64 4 801 4822.
E-mail addresses: e.e.hudson-doyle@massey.ac.nz (E.E.H. Doyle), john.mcclure@vuw.ac.nz (J. McClure), Douglas.Paton@utas.edu.au (D. Paton), david.johnston@gns.cri.nz (D.M. Johnston).

http://dx.doi.org/10.1016/j.ijdrr.2014.07.006
2212-4209/

again. Since 1995, Soufrière Hills Volcano, Montserrat, has gone through at least 3 distinct eruptive periods [23]. Eighteen years after it re-awoke, Soufrière Hills continues to display unrest and activity, presenting challenges for long term management and communication as communities learn to live with the volcano [35,36,67]. The epistemic uncertainty (the unknowns) and the aleatoric uncertainty (the stochastic variability) of the volcanic physical process thus contributes to considerable uncertainty in the crisis management process itself. In addition, an event such as this requires inter-disciplinary interaction, and if relationships and procedures are not well established and practiced before, considerable uncertainty can arise in its management due to problems with inter-agency communications, collaborations, and the understanding of each others' roles, responsibilities, and inter-dependencies [76,77,19], particularly when under high pressure, short time situations characterized by high ensuing risk and stress.

This variation in unrest periods, and the potential for eruption, non-eruption, or continued eruptive cycles, creates extreme challenges for those involved in the response as decision makers balance the issue of life safety and community continuity through the crisis [78,76]. Issues have arisen due to conflicting scientific advice either from internal and external agencies, or due to the presence of a wide range of scientific advisory bodies and individuals (e.g., [51,102,28,12,84,72,101,67]). Based upon experiences from Guadeloupe in 1976 and St. Vincent in 1979, Fiske [28] highlighted that successful volcanic crisis management is not just dependent on improved monitoring techniques but upon the communication between scientists, journalists and the public, and in particular, the need for experienced chief scientists, who "while not suppressing scientific disagreements, would attempt to coordinate the activities of the scientists involved into a single group effort, to increase communication between the scientists, and to help ensure that a single and complete stream of information is made available to civil authorities and journalists" (ibid, p. 176). Over 20 years later, during New Zealand's (NZ) Civil Defence exercise Ruaumoko, which practiced the nationwide response to an eruption in the Auckland Volcanic Field, one of the highlighted recommendations was again the demonstrated importance of having official scientific advice provided by "one trusted source" [65]. This was in the form of the Auckland Volcanic Science Advisory Group, which gathered together the social, geological, economic, geophysical and monitoring groups and communicated this pooled expertise to the decision makers while also responding to direct requests for information from the emergency managers. This process was identified as being of a considerable benefit to the response process as it helped to prevent conflicting or confusing messages [65]. Numerous other volcanic and geophysical events have identified the need for official scientific advice to be the trusted source of advice, that delivers appropriate, accurate advice that meeds the needs of (diverse) decision makers, emergency managers, and the public. This makes it impossible to consider a one-size fits all approach, and makes it important that recipent organizations not only establish a single source of advice, but build strong relationships and trust across agencies via exercises, workshops and meetings prior to an event, to help build individual and team mental models of each other's roles, responsibilities, and their information needs as well as to develop an understanding of other political and legal issues that may play a role in the implementation of advice (see [77,19]). As discussed by Jordan [45, p. 6], in light of the L'Aquila earthquake and trial, it is vital that the different roles of the science advisors and the civil decision makers are understood and kept distinct, as "confusing these roles can lead to trouble". In addition, as stated by McGuire et al. [67, p. 75], it is vital that the "mechanisms that underpin effective communication during a volcanic crisis are in place long before a volcano shows signs of unrest" so that seamless communication between main stakeholder groups can occur in-event. Thus all efforts need to focus on the "building of trust between stakeholders, the maintenance of good working relationships, and the safeguarding of an open and continuous information flow between all key players" (p. 75).

Currently in NZ, advice is communicated by a number of Scientific Advisory Groups across the volcanic regions, established to bring expertise from various scientific agencies together [65,98,19], and who often sit within wider volcanic advisory groups established by the regional authorities. For example, the Central Plateau Volcanic Advisory Group was established by the Horizons Regional Council to ensure that all responding organizations are "working together to increase community resilience to volcanic hazards within the Central Plateau" [14]. This group encompasses a Science Focus Group, a Planning Focus Group, and a Communications Focus Group, all guided by a framework strategy and Contingency Plan, and who meet every six months to report on work programmes, outcomes and future plans, and to help build relationships and inter-agency coordination.

However, while the goal is to have the experts within a SAG familiar with each other and other responding agencies via exercises, training, workshops, planning, and other relationship building activities, the need to call on other experts (e.g., the Ministry of Health for ash fall advice or psychosocial issues), changes of personnel within an agency over time, and changes in agency structure can create a situation where in fact these groups may be relatively unfamiliar with each other and their respective roles and expertise; particularly as the size of an event grows and extends beyond the regional to the national or international level, with the impact and management consequences of these differences being magnified by the high risk, high stress environment in which they would have to interact. Developing more comprehensive inter-agency training and embedding inter-agency operations into the organizational culture is a crucial first step to creating the kind of organizational learning required for "superordinate" organizational response to rapidly escalating, complex volcanic crises [29]. In the absence of such capabilities, the Scientific Advisory Group (SAG), which may thus include individuals familiar or unfamiliar with each other, thus has an important role in soliciting, collecting and pooling together a wide range of expertise into a single source of advice. However, little research has been conducted to guide them in how best to communicate any non-consensus advice and how to communicate the uncertainty related to this non-consensus to officials and the public. In addition, the implementation of this advice is affected by the variability in the legal framework across countries [31]. To date, many volcanologists have thus been cautious to communicate uncertainty or disagreement to public officials, for fear of a loss of credibility or competence [86]. This is not just limited to volcanology. For example, in the field of meteorology, similar concerns have been raised regarding the explicit communication of uncertainty for weather forecasts ([73], as cited in [70]).

Added to the uncertainty implicit in managing the volcanic event itself, uncertainty thus emerges in relation to activities such as deciding on and advising of the need for evacuation in the context of concerns about making an "economically disastrous, unnecessary evacuation" (Tazieff 1983, as cited in [111, p. 88]; see also [28]). As discussed by Kilburn [51] the trust between scientists and the public can be damaged by "accusations of 'crying wolf'". Historic cases identified by Tazieff [102] as "erroneous volcanological diagnoses" due to wrong interpretations of actual facts, or use of deliberately false data, have led to significant socio-economic consequences for the affected communities, as well as damaging trust and relationships. To this end, in the 70s, both Tazieff [102] and Kilburn [51] identified the need for practising volcanologists to abide by a "deontological code" or a "Hippocratic-type" code of practice to fulfil moral obligations to society. Since then, the International Association for Volcanology and the Earth's Interior (IAVCEI) Subcommittee for Crisis Protocols (1999) have established a set of best practice guidelines for conduct of scientists during volcanic crises, that incorporates a guiding principle, and outlines areas that can cause problems. They highlight that uncertainty should be acknowledged in communications between scientists and public officials, and remind scientists that "non-scientists can deal much better with worrisome information and uncertainty than many scientists believe" (p. 331). These issues do not just apply to volcanic eruptions, but also to tsunami warnings, earthquakes and associated aftershock sequences, weather warnings, and wildfires. As stated by Jordan et al. [44], the next vital step in operational forecasting is to apply the principles of effective public communication already established by social science research to the communication of hazard information.

In this manuscript, we next review some of the psychology literatures regarding decision making and uncertainty and its implication for volcanology. We then present a study investigating the effect of using either numerical or verbal probabilistic forecasts on decisions to evacuate for a hypothetical crisis scenario, and explore via a thematic analysis the reasons provided by non-scientists and scientists for those decisions. The implications of our research findings for the communication of uncertainty and probabilistic forecasts in volcanology are then discussed.

1.1. Managing uncertainty and decision making

The communication of advice about complex processes whose activity (timing, intensity, duration etc.) is characterized by considerable uncertainty over time has been highlighted as a global issue in disaster research. In particular, the Risk Interpretation and Action working group of the Integrated Research on Disaster Risk (IRDR) framework has identified the need to understand "the ways in which people interpret risks and how they respond based on these interpretations" which are "shaped by their own experience, personal feelings and values, cultural beliefs and interpersonal and societal dynamics", particularly when making decisions under uncertainty [22]. Fearnley [112] highlights that there are many approaches across disciplines for negotiating uncertainty and risk, and managing 'non-quantitative' uncertainty (that is the 'un-measurable uncertainty'), that would be beneficial to the management of volcanic hazards, particularly when assigning an alert level. These include more holistic approaches to the co-production of knowledge across the different stakeholders involved (including the public) and that combines uncertainty, ignorance, ambiguity and risk.

Science advice about natural hazards is often subject to many levels of uncertainty, due to the natural stochastic uncertainty (the variability of the system), and the epistemic uncertainty (lack of knowledge) [82,1]. In addition, uncertainty and disagreement between scientists regarding how diverse interpretive judgments influence their decisions about what advice to communicate can arise due to (a) incomplete information, (b) inadequate understanding, and (c) undifferentiated alternatives [59, p. 151]. Uncertainty in the context of action can affect both the scientific advisors in their decisions regarding communication, as well as the decisions made by the emergency managers that depend on that advice. Moss and Schneider [71] proposed seven recommendations for the IPCC regarding the handling and communication of uncertainties, which include "Document key causes of uncertainty" and "determine the appropriate level of precision of your conclusions after considering the nature of the uncertainties and the state of science". Wiedemann et al. [110] advise that experts should help lay people make informed judgements by communicating the full range of uncertainties, and stating clearly what is known. However, they question whether advice communicated in this manner actually helps non-experts reach the right conclusions, given differences in how information is processed by the respective groups. These approaches do not account for relative differences in how scientists and non-scientists or citizens process information when confronting, complex, uncertain and threatening events [24,96,11,97]. Nor do they consider how non-scientist, emergency managers are being influenced by their awareness of the social, economic and political dimensions of their decision making (e.g., the implicit need to manage response decisions with economic or political views that are imposed upon emergency managers). As stated in the IAVCEI Subcommittee for Crisis Protocols [40], volcanologists must "understand that a decision-making official who is presented with an uncertain scientific forecast, clear socioeconomic issues, and political pressure might not always follow scientific advice" (p. 330).

Epstein [24] outlines theory for two parallel processing systems that drive decision making (see also [96,11,97]).

One is the *analytic* approach which involves more deliberate (and hence makes for longer decision times) computational cognitive processes, and is a learnt process that consciously and deliberately applies rules and procedures to the analysis of data (e.g., formal logic, utility maximization that have been developed and applied to the problems being tackled). The analytic system uses algorithms, normative rules and logic, and does not operate automatically. It is also oriented towards delayed action, and decisions made often require 'justification via logic and evidence' [24]. This approach is used predominantly by scientists and risk management professionals to quantify their analyses and express the outcomes as probabilities and develops as a result of extensive training in analysing data surrounding volcanic processes. In contrast, the *affective* approach, which is often termed the *experiential system*, involves rapid, unconscious affective processes and is oriented toward immediate action. It is a holistic approach where behaviour is "mediated by 'vibes' from past experiences" [24], and is an evolutionary adaptation that automatically converts uncertain and adverse aspects of experience into affective responses (e.g., fear, dread, and anxiety). It thus results in people interpreting risk as an affective or emotional state or feeling [60,97]. It is thus influenced by factors in the decision context that affect emotions (such as the cost of incorrect decisions such as ordering evacuation when it is unnecessary, or failing to do so when it is necessary).

There has been recent debate regarding the dual processing theories of higher cognition [25,26,50,75,104], criticizing the evidence available for two systems [49,50], and proposing whether a unified theoretical approach that explains both intuitive and deliberative judgments as rule based is more appropriate [53]. Evans and Stanovich [25] suggest however an approach where "rapid autonomous processes (Type 1) are assumed to yield default responses unless intervened on by distinctive higher order reasoning processes (Type 2)", the latter of which heavily loads working memory and utilizes hypothetical thinking. Thompson [104] argues that in fact type 2 processes might also be triggered automatically but that, "unlike Type 1 processes, their completion requires working memory resources, and that the outcome of those processes is more flexible than that of Type 1 processes".

To discuss this debate further would be beyond the scope of this manuscript, thus we refer herein to the *affective* (similar to Type 1) and *analytical* (similar to Type 2) processing systems presented by Epstein [24], and consider that while both systems influence people's judgements, scientists are likely to develop a more analytical process, and while non-scientists or emergency managers may adopt either processes depending on the governing operating procedures, prior experience, and time pressures, they will generally tend more towards affective processes. The adoption of the affective and analytical processing systems is thus not an either-or situation, but rather a more complex balancing act influenced by the degree of uncertainty or threat in the decision context, and the relative experiences of both scientists and emergency managers. While ordinary citizens use some analytical processing, they are more likely to base their decisions on the more readily available affective processing. Even if non-scientists adopt an analytical process for their decision, if the outputs from the two processing systems disagree, the affective system usually prevails [60]. Analytic processing of scientific information may lead people to recognize hazards as a significant threat, but their affective processing of the output often reduces the likelihood of their acting on this knowledge [108].

Considering decision making under uncertainty for the analytic system, Lipshitz and Strauss [59, p. 150] state that uncertainty in the context of action is "a sense of doubt that blocks or delays action", and classify it according to the issue (i.e., what the decision maker is uncertain about) and the source (i.e., what causes this uncertainty). Uncertainty in the context of the issue can relate to the outcome, the situation itself, and the alternative actions available, and uncertainty such as this can also arise due to "undifferentiated alternatives". Many formal and behavioural decision theories identify the R.Q.P. heuristic for coping with uncertainty in decision-making (see [59]), which represents the *Reduction* of uncertainty by information searching, the *Quantifying* of the magnitude of uncertainty that cannot be reduced, and the *Plugging* of the result into a formal decision-making scheme that incorporates uncertainty.

From an analysis of written accounts of decisions made under high levels of uncertainty, Lipshitz and Strauss [59] found that decision makers reduce, acknowledge or suppress the uncertainty (see Table 1). From this, they proposed their RAWFS (Reduce, Assumption-based reasoning, Weighing pros and cons, Forestalling, Suppressing) heuristic, stating that "how decision makers cope, or ought to cope, with uncertainty is principally determined by the nature or quality of the uncertainty" [59, p. 160]. This model assumes however a 'rational' or analytical decision maker, and does not incorporate the uncertainties introduced by individual interpretative processes, biases and interactions, or the role of more implicit or experiential modes of thinking [24], except for in the suppressing stage when behaviours such as "relying on intuition" are considered (Table 1). Nor does it accommodate the different interpretive bases of the different stakeholders in the decision process. The latter becomes important in the context of one of the key elements of the reduction phase of their RAWFS heuristic: the soliciting of advice and opinions of experts; demonstrating how science advice is not just about providing information for situation assessment, but also about providing advice to help decision makers understand, acknowledge and reduce the uncertainties in the source and the complex physical systems.

1.2. A survey to investigate uncertainty and probabilities

To address the many risks and uncertainties involved in volcanic eruptions, due to their complex nature, and to help facilitate decision making for end users and stakeholders, it has become increasingly popular for scientists to use probability statements in their communications (see full review in Doyle et al. [20]). The IAVCEI Subcomittee for Crisis Protocols [40, p. 330] recommend the use of "probabilities to calibrate qualitative assessments of risk".

Table 1
Methods used to reduce, acknowledge and suppress uncertainty while making a decision, as identified by Lipshitz and Strauss [59, p. 153–154], see also [58].

Reduce	• Collecting additional information, • deferring decisions until additional information becomes available, • soliciting advice and following SOPs, • filling in gaps in factual knowledge through assumption based reasoning.
Acknowledge	• Taking the uncertainty into account in the selection of an action by incorporating slack into the decisions/actions, • improving readiness by generating new alternatives to pre-empt a specific potentially negative outcome, • weighing up the pros and cons of an approach.
Suppress	• Ignoring the uncertainty, • relying on intuition, • Rationalizing and removing the doubts that block action.

Solana et al. [101] identified however that civil authorities prefer to receive a deterministic forecast (eruption expected within a time window) over a probabilistic statement (*X*% probability within a time window), attributing this preference to a "reluctance to take responsibility for interpreting inherently uncertain information during a crisis" (p. 312). The communication of probabilistic statements such as these, whether in numeric or linguistic format, is complicated by the fact that they can be misinterpreted due to the framing, directionality and probabilistic format adopted [9,38,47,48,57,103]. Verbal phrases are seen as more 'accessible' to the public, and are generally better understood than their numerical counterparts [107,83,101]. However, such verbal phrases (e.g., *likely, unlikely*) are subject to the 'translation issue', whereby people's interpretation of the numerical equivalent of these terms varies widely [103,17,89,57]. This can result in poor and/or inconsistent decisions based not on the data itself, but on the misinterpretation of the probabilistic data.

To address the need to convey quantitative probabilistic risk assessments, while still conveying information in an 'accessible' format that can inform (diverse) decision needs, reports should adopt a dual approach with sophisticated parts using a numeric format, and more general parts using verbal phrases and narratives [82]. 'Translation tables' (Table 2) are then used to convert from verbal to numerical terms. It is important to highlight the use of tables such as this, as the framework may not match people's intuitive use of the language [82], and they may find it hard to suppress the meanings they normally associate with these terms [48].

Our previous research has investigated the perception of numerical and probability forecasts for volcanic eruptions with scientific and response agencies who would usually respond to a natural hazard event (see Section 2), including

Table 2
IPCC Qualitative Descriptors used for the Third Assessment Report Climate Change 2001, as given in Patt and Schrag [83] from Houghton et al [39, p. 28]. A similar table has been developed based on this by the WMO [33], but with additional categories directly either side of medium likelihood.

Probability range	Descriptive term
$<1\%$	Extremely unlikely
1–10%	Very unlikely
10–33%	Unlikely
33–66%	Medium likelihood
66–90%	Likely
90–99%	Very likely
$>99\%$	Virtually certain

1. The use of translation tables in volcanology [21], finding that *scientists* and *non-scientists* differ in translations of verbal likelihood phrases into numerical equivalents, and that inconsistency with the IPCC translation scales (e.g., Table 2) supports the development of a translation table built from the target community (non-scientists) for hazard communications, to match their intuitive use of these terms.
2. Whether scientists and key decision makers differ in their interpretations of event likelihoods within forecast time windows (e.g., "*there is a 68–88% chance of an explosive eruption in the next 10 years*"; [21]), finding that for long time windows (10 years) participants skew their judgment of an event likelihood towards the end of the time window, rating the likelihood of an event occurring 'this year' as less than year 10. Using the term "*within*" instead of "*in*" was found to mitigate some of this skew.
3. That this skew in perception also occurs for short-term time windows (under one week; [20]) that are of most relevance for emergency warnings, and that the use of phrasing "*within*" instead of "*in*" did not mitigate this skew for short forecasts. They also found a skew in perceptions for probability-free statements such as "*threat of an eruption within 2–3 days*".

Doyle et al. [20] suggest that these skew in time perceptions for forecasts reflects participants overlaying information in the statement with their understanding of volcanoes, in a manner similar to the *base rate* effect [48,17,82,57]. However, both of these studies [21,20], considered only the participants' perceptions of the likelihood of an eruption. Herein, we investigate how these perceptions influence participants' action choices, and the reasons for decisions made. We investigate how using a numerical or verbal probability format in a forecast influences the choice to issue an evacuation, and how those decisions and reasons vary between scientists and non-scientists, and between different likelihood scenarios.

2. Method

Following Doyle et al. [21,20], participants were directed to the online survey described briefly above through a snowball approach via email contact with individuals in organizations from the natural hazard, physical and social science community in New Zealand, and from emergency management and civil defence authorities. These organizations included GNS Science (Geological and Nuclear Science), NIWA (National Institute for Water and Atmospheric Research), NZ universities, MCDEM (Ministry of Civil Defence and Emergency Management), regional and local CDEM groups, emergency and protective services, and lifeline companies. In addition, the survey was delivered internationally to capture global perspectives. Additional recruitment was conducted though advertisements in bulletins and on-line notice boards, such as the MCDEM e-bulletin, the international 'Volcano Listerv' (run by Arizona State University), the bulletin board of the Comprehensive Emergency Management Research Network (CEMR), and in the Oceania newsletter of the International Association of Emergency Managers. In total, 179 participants completed the survey; 76 were located in NZ, 36 in North America, 15 in Europe, 7 in Australia or Oceania/ Pacific, 2 in South America, 2 in Central America, 1 in Asia, 1 in the Caribbean, and 6 undeclared.

The online survey was administered through the Qualtrics Survey Research Suite software (Qualtrics Labs Inc., Version 2.03s©2011) which randomly assigned each participant to either a 'numerical' or 'verbal' experimental group for the survey. The survey was anonymous, and participants identified their primary employment sector from the options listed in Table 3, from which they were grouped into *scientists* ($n=92$) and *non-scientists* ($n=85$) with 2 undeclared.

For the two questions discussed herein, participants were presented a hypothetical volcanic scenario depicting an unrest episode affecting a town. The scenario provided limited information about the town, to enhance the ambiguity of the decision scenario. Participants were then told that it takes at least 48 h to execute an evacuation of the "capital city", and given the volcanologists' forecast of eruption likelihood within a time window. Participants were then asked whether they would evacuate the town, when they would evacuate, and to comment on their decision. The format of this hypothetical decision-making scenario was based upon a standard approach used in psychology to explore the effect of specific wordings, formats and probabilities on action choices (see e.g., [46,47]), and this approach was adopted to facilitate an investigation of the effect of using either verbal or numerical probabilities on decisions to evacuate.

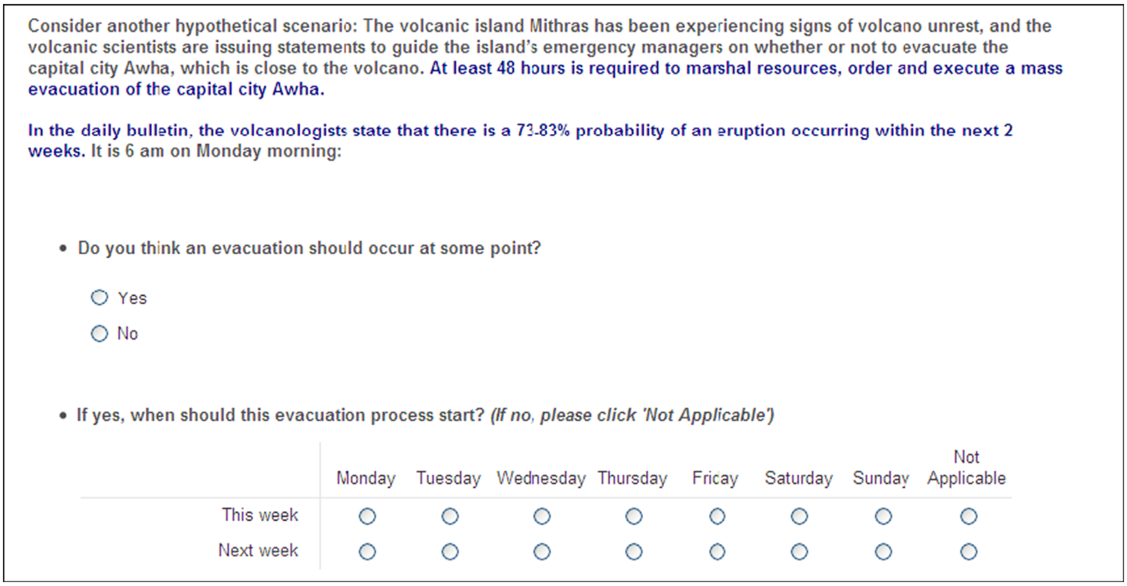
The probabilistic statement of event likelihood presented to participants included a time window which for experiment 1 stated "*the volcanologists state that there is a 73–83% probability of an eruption occurring within the next 2 weeks*", and for experiment 2 stated "… *there is a 45–55% probability of an eruption occurring within the next 3 days*", as depicted in Figs. 1 and 2 respectively. For each experiment, one group received the 'numerical' probability statements described ($n=73$ for experiment 1, $n=69$ for experiment 2), while the other received the equivalent 'verbal' probability statements using the translation outlined in the IPCC translation Table [83]; resulting in the statements "… *a likely chance of an eruption occurring within the next 2 weeks*" and "… *a medium likelihood chance of an eruption occurring within the next 3 days*" ($n=71$ for both experiments).

Both a within- and between- subject designs [3] were used to analyse the responses. All participants were presented with both experiments, and a randomised coin-toss was used to assign a participant to either the numerical or the verbal condition for each experiment. Between-subject analyses were thus conducted between the responses to either the numerical or verbal version, and between the scientists' and non-scientists' responses, for each experiment. Experiment 2 represents a higher likelihood scenario (45–55%, 3 days) compared to experiment 1 (73–83%, 2 weeks), and thus an additional within-subject analysis was conducted across the two experiments. For both experiments, investigations initially focused on a quantitative analysis of the action choices (evacuation and timing) made by each group presented with either the numerical or verbal probability statement, as well as the differences between the judgments of scientists and non-scientists within each group. For these comparisons, we use the χ^2 test ([106, p. 44–45]).

We then analysed the comments shared by participants in response to their decisions using thematic analysis to identify patterns of meaning [6,91]. This enabled us to identify key themes in the understanding of the science in the scenario and evacuation decision considerations. To

Table 3
Categories provided for participants to select their primary employment sector [21,20].

Category	Grouped as
Scientific or technical (agency, university or research institute)	Scientists
Central/national government, civil defence, emergency management (Ministry, agency etc.). Local/regional government, civil defence, emergency management (Council, agency, etc.). Public safety, emergency services (police, fire, ambulance, rescue, response, etc.). Lifelines (infrastructure, water, telecommunications, electricity, transportation, gas, etc.). Other.	Non-scientists


Consider another hypothetical scenario: The volcanic island Mithras has been experiencing signs of volcano unrest, and the volcanic scientists are issuing statements to guide the island's emergency managers on whether or not to evacuate the capital city Awha, which is close to the volcano. At least 48 hours is required to marshal resources, order and execute a mass evacuation of the capital city Awha.

In the daily bulletin, the volcanologists state that there is a 73-83% probability of an eruption occurring within the next 2 weeks. It is 6 am on Monday morning:

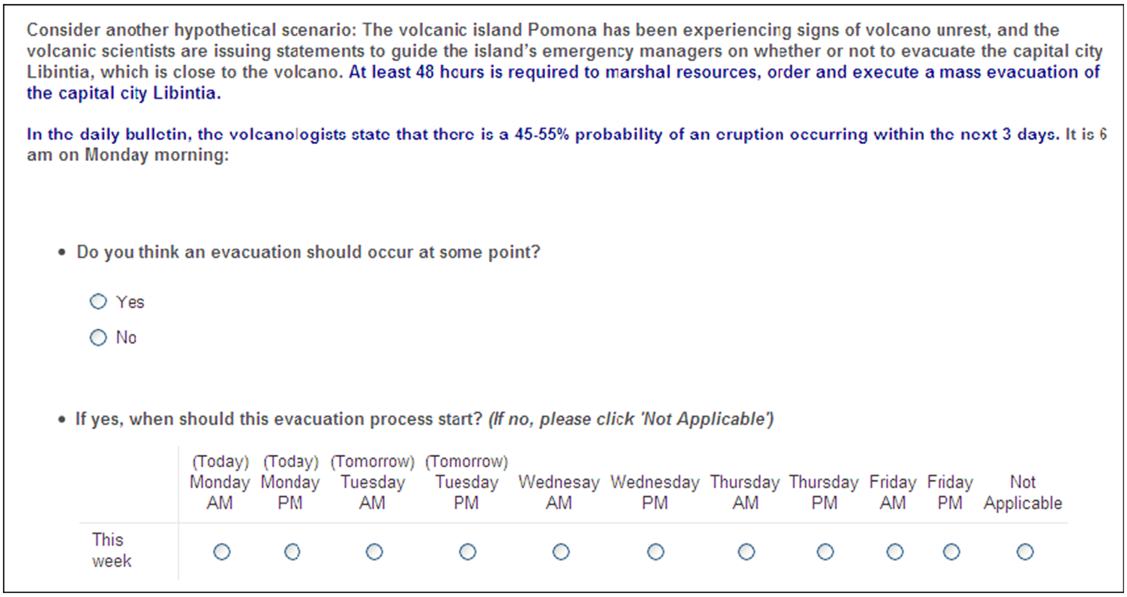
- Do you think an evacuation should occur at some point?

○ Yes
○ No

- If yes, when should this evacuation process start? *(If no, please click 'Not Applicable')*

	Monday	Tuesday	Wednesday	Thursday	Fricay	Saturday	Sunday	Not Applicable
This week	○	○	○	○	○	○	○	○
Next week	○	○	○	○	○	○	○	○

Fig. 1. Screen shot of experiment 1, showing the probability statement presented to the numerical experiment group. The verbal group received the equivalent probability statement using the translation outlined in the IPCC translation table [83], and resulting in the statement "… *a likely chance of an eruption occurring within the next 2 weeks*".


Consider another hypothetical scenario: The volcanic island Pomona has been experiencing signs of volcano unrest, and the volcanic scientists are issuing statements to guide the island's emergency managers on whether or not to evacuate the capital city Libintia, which is close to the volcano. At least 48 hours is required to marshal resources, order and execute a mass evacuation of the capital city Libintia.

In the daily bulletin, the volcanologists state that there is a 45-55% probability of an eruption occurring within the next 3 days. It is 6 am on Monday morning:

- Do you think an evacuation should occur at some point?

○ Yes
○ No

- If yes, when should this evacuation process start? *(If no, please click 'Not Applicable')*

	(Today) Monday AM	(Today) Monday PM	(Tomorrow) Tuesday AM	(Tomorrow) Tuesday PM	Wednesay AM	Wednesday PM	Thursday AM	Thursday PM	Friday AM	Friday PM	Not Applicable
This week	○	○	○	○	○	○	○	○	○	○	○

Fig. 2. Screen shot of experiment 2, showing the probability statement presented to numerical experiment group. As of experiment 1, the verbal group received the equivalent probability statement using the translation outlined in the IPCC translation table [83], and resulting in the statement "… *a medium likelihood chance of an eruption occurring within the next 3 days*".

conduct this thematic analysis, we followed the approach of Braun and Clarke [6] and followed a five step process involving (a) familiarization with the data (through reading and re-reading the free text responses), (b) generating initial codes from interesting features within the data in a systematic fashion, (c) searching for themes by collating and clustering these codes into potential themes, (d) reviewing the themes and developing a thematic map, and (e) defining and naming the themes. Throughout this process, we utilized a theoretical thematic analysis approach [6], where our identification of themes was driven by our research questions regarding decision making, uncertainty, and inter-agency communications and evacuation issues. This is in contrast to an inductive analysis where themes are identified in a 'bottom up' way without being influenced by theoretical and epistemological commitments. Throughout this process, the codes were continuously reviewed and evaluated in a recursive manner to assess whether the existing codes accommodated the breadth of participants' responses (as of [52]).

Throughout our results discussion, we may refer to the number of participants who raised a particular comment or theme. However, to analyse the prevalence of a theme across experiment groups, we adopt the approach of Dalla [16] and Kim [52] (as cited in [7]) whereby we report frequency counts of theme occurrences to indicate the strength and consistency of a theme between groups, rather than the number of different speakers who articulated that theme. We thus use only descriptive data (numbers and percentages) to report the comparative analysis between themes in each group, to illustrate the strength and consistency of a theme within each group. We do not conduct a statistical analysis on the occurrence of the themes within each group, as we have ‘no way of interpreting what is not reported in qualitative data’ [7, p. 261]. Thus we consider the meaning, importance, value and validity of a particular theme is not defined by whether it is significant or not in statistical analysis of its frequency [7] and that such a statistical analysis would be inappropriate and misleading.

3. Experiment 1: verbal and numerical probabilities

We consider first the decision making scenario where the numerical group read the probability statement “… *the volcanologists state that there is a 73–83% probability of an eruption occurring within the next 2 weeks...*”, whereas the verbal group read the probability statement “…*the volcanologists state that there is a likely chance of an eruption occurring within the next 2 weeks*” (see Fig. 1). Fig. 3 illustrates the evacuation decision or not for the scientists and non-scientists, for both the numerical and verbal groups.

For the group presented with the numerical statement, 90.4% of participants who answered ($n=73$) chose an evacuation when presented with the numerical statement, with 83.6% in the first week and 46.6% *immediately* on Monday (today). In comparison, 71.8% of the group presented with the verbal statement ($n=71$) chose an evacuation when presented with the verbal statement, with 64.8% evacuating in the first week and 45.1% *immediately* on Monday (today). There is a significant relationship between the format of the probability statement and the evacuation decision, with participants choosing an evacuation more often when presented with the numerical statement ‘73–83%’ compared to the verbal statement ‘likely’ ($\chi^2(1, N=144)=7.914$, $p=0.005$, medium effect size $V=0.234$).

Considering the behaviour of participants who identified as scientists and non-scientists separately, for the numerical group, 95.1% of scientists ($n=41$) and 84.4% ($n=32$) of non-scientists choose an evacuation. However for the verbal group, only 69.7% of scientists ($n=33$) and 73.7% of non-scientists ($n=38$) choose an evacuation. There is a significant relationship between likelihood format and the decision to evacuate for the scientists, ($\chi^2(1, N=74)=8.699$, $p=0.003$, moderate effect size $V=0.343$), but not for the non-scientists ($\chi^2(1, N=70)=1.179$, $p=0.278$). The scientists are more likely to choose an evacuation for the numerical statement than the verbal, and are thus effectively more ‘sensitive’ to the different probability format than the non-scientists. This may be due to a preference for, or greater familiarity with, numerical probability statements that makes them appear more precise or certain to scientists, while their verbal equivalents seems more vague and less likely.

Interestingly, there is no significant relationship between the numerical and verbal groups and the timing of the evacuation decision *immediately* on Monday (today) for all respondents ($\chi^2(1, N=131)=0.204$, $p=0.652$), for scientists ($\chi^2(1, N=69)=0.305$, $p=0.581$), or for non-scientists ($\chi^2(1, N=61)=1.346$, $p=0.246$). This suggests that, for this scenario, while probability format affects the evacuation decision, it is not affecting the urgency of the resultant evacuation. There are however some respondents who do choose to delay their evacuation (see Fig. 3), and it is important to understand both their perspectives and those who chose not to evacuate.

3.1. Reasons for evacuation decisions

A thematic content analysis was conducted on the free text responses provided by participants to experiment 1, to help us understand the reasons behind their evacuation decisions and timing, as discussed in our methodology. A number of key themes were identified (see Table 4), discussed in turn in the following sections.

3.1.1. Understanding the probability and time forecasts

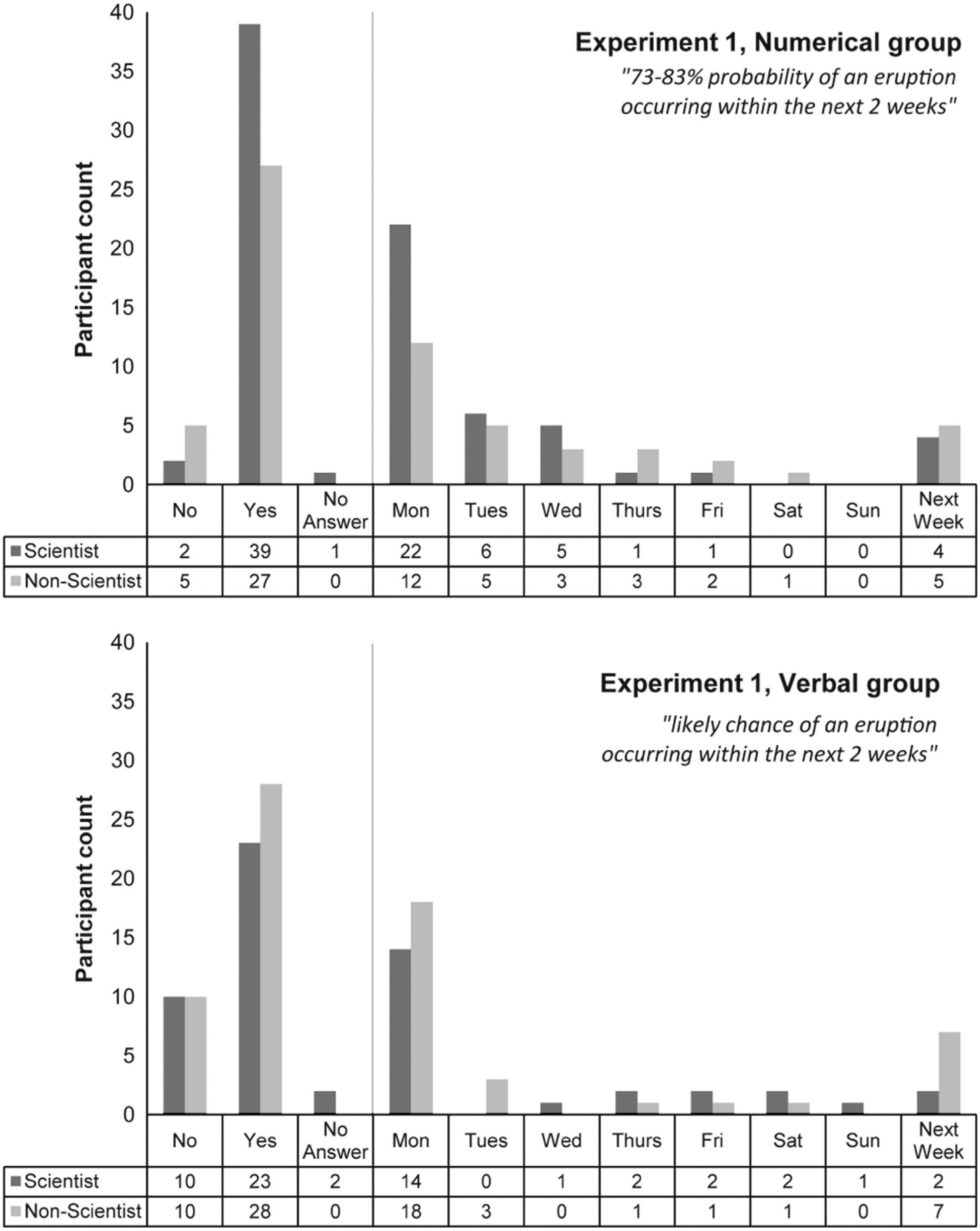
Considering first the group who read the numerical version of the probability statement (78–83%, 2 weeks), four participants stated that the event could occur anytime from now and that an evacuation should occur as soon as possible, therefore recognizing the real risk of an event today.

- *I understand the statement as implying that the eruption could occur anytime within the next two weeks, therefore if an evacuation should take place, it should be BEFORE the actual eruption, not waiting until it is too late.* [scientist, numerical]

In the comment above, the decision of when to evacuate may have involved working back from the end of the time window and incorporating the time required for an evacuation, and comments by a number of participants appeared to reflect this approach. Two participants also noted that the likelihood of an eruption being equal throughout the week.

- *The probability of an eruption occurring is constant, i.e., the likelihood of an eruption occurring on the first Monday is the same as it happening on the last Sunday, so evacuate the town as soon as possible (assuming the town is in the path of likely hazards)* [scientist, numerical]

This likelihood distribution is reasonable based on the information provided, particularly when an evacuation decision needs to be made; however, strictly speaking, the probability described may represent an increase, decrease or plateau of activity throughout the time



	No	Yes	No Answer	Mon	Tues	Wed	Thurs	Fri	Sat	Sun	Next Week
Scientist	2	39	1	22	6	5	1	1	0	0	4
Non-Scientist	5	27	0	12	5	3	3	2	1	0	5

	No	Yes	No Answer	Mon	Tues	Wed	Thurs	Fri	Sat	Sun	Next Week
Scientist	10	23	2	14	0	1	2	2	2	1	2
Non-Scientist	10	28	0	18	3	0	1	1	1	0	7

Fig. 3. The evacuation decision or not, and the time chosen to evacuate, for experiment 1 (Fig. 1) and numerical probability experiment group (73–83%) and the verbal probability experiment group (likely). Results are shown for participants who identified as scientists (dark grey) and non-scientists (light grey).

specified. As stated by Tilling [105], a period of unrest could culminate in a major eruption or dormancy after a short duration of unrest, after a long lull in unrest, or continue as consistent ongoing unrest over long periods of time, or ongoing irregular periods of unrest.

The participants who read the verbal probability statement (likely, 2 weeks) raised more comments about the probability translations and time forecasts than the numerical group, with 10 occurrences of this theme in the verbal group responses compared to 6 occurrences in the numerical group (as illustrated in Table 5). As above, participants noted that the eruption could "*happen any time from now on*", and "*anytime within those 2 weeks*", or assumed "*the estimated probability applies equally for the entire two week*

Table 4
The key themes and codes identified from the free text responses to experiments 1 and 2.

Key theme	Code from text
Evacuation thresholds	Evacuation threshold – probability is high enough for evacuation; evacuation threshold – probability needs to be higher for evacuation, probability too low
Evacuation technicalities	Complexities of evacuation issues and processes; evacuation process – education awareness of public; evacuation time scheduling; humanitarian vulnerable pets special needs; lives, safety, order, daylight etc; mandatory evacuation, or not mandatory; seat of government; self or voluntary evacuation (in evacuation technicalities); staged evacuation, prioritized evacuation; time restraints, short time
No time to lose	ASAP, why wait, no time to lose, better safe than sorry, before % increases; earlier evacuation better; evacuation time equals time until probability increases; evacuation can be cancelled; evacuation planning, prep, marshalling resources; precautionary principle, proactive, prudent, err on side of caution; too late
Let's wait and see	Depends on volcano behaviour; evolving processes n probabilities; monitoring of situation; wait delay; waiting may increase certainty; no action – indecisive decision makers
Managing uncertainty (MU)	
MU - Cost benefit	Cost – benefit; risk, risk tolerance, threat
MU - Assumptions made	Future outcome ok; information good, reliable; no prior concern; processes (emergency management and evacuation); time and forecasts; what information means; impacts; ignored some stats
MU - Incomplete information, information needed	Volcano and eruption type, hazards, location of town, risk to city, impact, intensity; better information, general, more information, more evidence, population demographics etc. history so far; general science dialogue monitoring, data; risk, general risk, loss of life; probability, time forecasts, temporal constraints
MU - Level of certainty or confidence	Certain, certainty; discussion of confidence; uncertain, uncertainty, no certainty, vague
Time forecasts	Could occur anytime – no definite time, as some point in near future; could occur now, today, within next 48 h, any time from now on – focus is on NOW; decision depends on timescale; difficulty understanding probability for a particular day; don't believe it is possible to give stats over this short time frame; don't know when within timescale; equal probability throughout the week; greater probability within 48 h; it may not erupt at all, if at all etc; likelihood increasing at rapid pace; long time scale means ramping up; not enough % variability through time; probability increases after 48 h; short window implies accelerating unrest; should be a shorter time interval; views end of time window as event occurrence; will occur tomorrow
Probability translations	Verbal to numerical, numerical to verbal, etc.
Science advice, understanding, role	Eruption intensity, magnitude; information quality; scales; scientists make decision, should not make decision, science conflict, scientists reluctant to give certain advice; trends, increasing or decreasing probability; volcano type, different type of hazards
Public action	Empowerment of the public; preparing the public; public evacuation plans; public unlikely to evacuate, won't evacuate, non-compliance, non-compliance in the future, non-compliance management; self or voluntary evacuation (in public action)
Public perception	Justify; public criticism trust, credibility; public observations of activity and reassurance, behaviour
Future issues – negatives and positives	False alarm, evacuation fatigue, cry wolf; positive – education, experience, practice; recovery and when to return

period". However, additional comments were also made about the fact that the eruption "*may not occur*":

- *Although there is a likely chance of an eruption occurring in the next 2 weeks, it may not erupt at all.* [non-scientist, verbal]
- *We don't know within that 2 week period when [or if] the explosion will take place,* [no career given, verbal]

This possibility of a non-event was rarely discussed for the numerical statement, which suggests that the verbal probability statement is more ambiguous than the numerical probability statement. This is in line with the quantitative analysis, where 90.4% of people choose an evacuation for the numerical statement "73–83%", compared to 71.8% for the verbal statement "likely".

- *High probability of an eruption and people should be moved to save lives* [non-scientist, numerical]

Interestingly, for the verbal probability statement scenario, several chose to translate the verbal probability term into a numerical equivalent, including "*I consider "likely" to be around 75% chance*", "*Likely sounds like 50/50 chance*", "*Likely = 60–80%*", "*interpreting likely as a … probability > 10–30%*". For the numerical probability statement scenario (73–83%), a number of participants commented about this being a "3 in 4 chance", translating the probability into a 'natural frequency format' which is suggested by Gigerenzer and Edwards [32] to enhance Bayesian reasoning. Both these translations demonstrate the wide range of equivalents to these verbal terms, and the relative familiarity that some participants may have with probabilities. It also illustrates how participants attempted to reduce the ambiguity of the verbal probability term by first converting it into a numerical counterpart prior to making a decision. This demonstrates the key limitation of verbal statements: while they are more 'accessible' to readers (Section 1.2), any translation made by participants increases the cognitive load while also increasing the potential for translation mistakes, biases, and miscommunications.

3.1.2. *Using and understanding the science advice*

For both the numerical and verbal groups, several participants raised comments about the science described in the statement, with similar issues raised by both groups. The majority of these highlighted that the evacuation

Table 5
The total number of occurrences (*N*) of each theme provided in the free text responses for the numerical and verbal condition of each experiment, as well as the number of those occurrences which came from those who chose to evacuate (Yes) and who chose not to evacuate (No). As described in our methodology, these numbers do not represent the number of participants that articulated that theme, but the number of times a theme occurred, and for this table and subsequent tables they are shown to indicate prevalence and strength of a theme.

Themes	Experiment 1						Experiment 2					
	Numerical			Verbal			Numerical			Verbal		
	N	Yes	No	*N*	Yes	No	*N*	Yes	No	*N*	Yes	No
Evacuation thresholds	15	13	2	7	3	4	10	3	7	11	4	7
Evacuation technicalities	29	24	5	25	20	5	14	11	3	23	20	3
No time to lose	32	26	6	26	23	3	16	15	1	21	18	3
Let's wait and see	11	9	2	10	5	5	8	5	3	12	10	2
Managing uncertainty (MU)	25	22	3	34	13	21	17	10	7	15	8	7
MU – cost benefit	3	3	0	3	0	3	4	4	0	7	4	3
MU – assumptions Made	4	4	0	2	2	0	2	2	0	1	1	0
MU – incomplete information, information needed	14	12	2	19	7	12	10	3	7	6	3	3
MU – level of certainty or confidence	4	3	1	10	4	6	1	1	0	1	0	1
Time forecasts	6	6	0	10	9	1	5	5	0	1	1	0
Probability translations	4	4	0	4	2	2	2	1	1	3	3	0
Science advice, understanding, role	9	9	0	8	4	4	7	3	4	3	2	1
Public action	9	8	1	6	6	0	3	2	1	8	4	4
Public perception	2	1	1	5	2	3	2	0	2	4	2	2
Future issues – negatives, positives	4	3	1	1	1	0	2	2	0	2	0	2

decision depends on the expected eruption intensity, the volcano type, and the different type of hazards, and that they made assumptions about these impacts to make their decision. For example

- *Would depend on the anticipated magnitude or intensity of the likely eruption. If catastrophic eruption is the typical activity, start preparing for evacuation immediately; begin the process of evacuation of non-essential residents as soon as is practical. If a lower intensity eruption is more likely, then begin preparation and logistics for removal and track the intensity of the activity. If intensity is on an upward trend, then begin evacuation of non-essential residents as soon as it can be done in a orderly way.* [scientist, numerical]

Several participants also considered the potential for the situation to change during the specified time forecast, referring to the 'trend' of activities, and 'increasing' or 'decreasing' probabilities. They then chose to either delay their evacuation based on this evolving situation, or to modify the evacuation process when more information became available

- *Preparations should be made for evacuation and the volcano should be closely monitored. If activity increases then evacuations should be made* [scientist, verbal]
- *Evacuation can be ramped down if activity in subsequent days shows waning unrest, but marshalling plans and transportation resources need to be staged appropriately to respond to rapid escalation of unrest.* [non-scientist, verbal]

These responses indicate that people are making some decisions about when action might be required. That is, they appear to be forming decision thresholds.

3.1.3. Evacuation decision thresholds

In the example comments provided above, we can identify a number of decision 'thresholds' where the evacuation decision depended on the current or evolving situation. This theme was raised more often in the numerical group than the verbal group, with this theme occurring 15 times in the responses to the numerical statement and 7 times in response to the verbal statement (Table 5). This raises the possibility that people can move up and down a decision process more readily with numerical information than with verbal information.

For the numerical statement, 13 of the occurrences of this theme referred to the probability being "high enough for evacuation", and also originated from those who chose to evacuate (see Table 5)

- *A 78% chance of an eruption is too high to not take action. I would evacuate asap* [non-scientist, numerical]
- *If the probability of eruption is that high, then I wouldn't muck around – dependent on whether it's possible to get a better reading within a 48 h window. If this is possible then I'd delay until the probability is higher.* [non-scientist, numerical]

However for the verbal probability 5 of the 7 occurrences of this theme referred to the probability "needing to be higher for evacuation".

- *I think that the probability of an eruption needs to be higher than 'likely' to require evacuation.* [scientist, verbal]
- *I would not recommend evacuation based on the daily bulletin alone with a statement such as "likely chance". It would get my attention to start the planning process only*. [non-scientist, verbal]

In the design of this experiment, the verbal term "likely" was chosen from the IPCC tables to be the verbal equivalent of the "73–83%" probability (likely = 66–90%, Table 2). However, the evacuation decision, and the comments raised for each statement above, suggests that the participants do not see the terms "73–83%" and "likely" as equivalent, with one scientist stating it "*sounds like 50/50 chance*". They are assuming that the verbal statement represents a lower 'possibility' scenario than the corresponding numerical statement, preferring to wait for a higher likelihood in the former case before initiating an evacuation. Inconsistencies with the IPCC translation table were also identified by Doyle et al. [21] who found that 'likely' is translated to 68 ± 14% and "very likely" is 83 ± 10%, with wide ranges of 28% and 20% respectively due to the varying interpretations or the 'meaning' of these words. The judgments found here, with lower evacuations for the "likely" scenario than its numerical equivalent (73–83%), suggest that the choice to use either verbal or numerical terms in a forecast not only affects translations and likelihood perceptions, but has clear impacts on evacuation decisions, timing, and reasoning.

3.1.4. Decision urgency: "No time to lose", "Let's wait and see"

When commenting on why they chose an evacuation, a number of participants stated that there was "*no time to lose*", that an evacuation should occur "*ASAP*", and that "*it is better to be proactive*" or to take a "*precautionary*" approach.

- *73–83% would indicate to me that an eruption is imminent and evacuations should begin immediately and probably should've started much earlier when the probability was around 30%* [scientist, numerical]

As would be expected based on the findings so far, there was a greater occurrence of the theme "no time to lose" in responses to the numerical statement (32 occurrences) than in responses to the verbal statement (26 occurrences), reflecting the urgency of the situation. Within this theme the participants who received the verbal statement also raised more comments that related to evacuation planning, preparation and marshalling of resources, with 14 occurrences of the theme 'evacuation technicalities' overlapping with the 'no time to lose' theme for the verbal group (compared to 7 overlapping occurrences for the numerical group).

- *The process should start immediately so that the planning and resourcing is done before the eruption starts as it is going to take 48 h to complete the evacuation. If the preplanning is done and those evacuated who wish to move of their own free will then it makes it easier to evacuate those who are reluctant to go but will when the eruption is imminent.* [non-scientist, verbal]

Unsurprisingly, the majority of those comments relating to evacuation technicalities came from participants who identified themselves as *non-scientists*, whose responsibility it would be to conduct these activities during an actual event (see Table 6). In addition, it is unclear whether these participants are interpreting these probabilities as an indication of *when to plan*, or *when an eruption will occur*.

Interestingly, some respondents in both the numerical and the verbal probability group raised comments that were classified under the theme "*let's wait and see*"; with 11 occurrences of this theme in the numerical group's comments, and 10 occurrences for the verbal group (Table 5). They preferred to monitor the situation, with the eventual evacuation decision stated as depending upon the volcano behaviour, evolving processes and probabilities and that waiting may "*increase certainty*".

- *A few more days monitoring data should improve confidence of advice re probability. But can't wait too long. Review decision when to start process in 2 days time unless prior increase in parameters/certainty.* [non-scientist, numerical]
  Preparations should be made for evacuation and the volcano should be closely monitored. If activity increases then evacuations should be made. [scientist, verbal]

The decision to either "wait and see", or act as there is "no time to lose", is not just dependent upon the scenario and forecast format, and participants' decision concerns, but also their prior perceptions of the potential outcomes of a period of volcanic unrest and their perceptions of event likelihood throughout the time window [20]. Thus many participants may assume the situation must only escalate ("no time to lose") and fail to recognize the multiple potential outcomes [105].

3.1.5. Managing uncertainty

In the comments cited above, several approaches can be identified where participants reduced the uncertainty of the scenario to help them decide whether and when to evacuate. This includes making assumptions about the situation, the hazards or the advice, and requesting or waiting for more monitoring data, before making a decision to evacuate. Those that delayed their decision (for more information) are still making a decision: the decision not to act, not to issue an evacuation. However, it is unclear how this relates to how they manage uncertainty versus being paralyzed by it. It also raises the interesting question as to how the data they have available now (probability forecast) may frame their interpretation of any future data. For example, the "boiled frog" scenario[1] could reduce their sensitivity to incremental differences in "likelihood" over time. Thus once the process commences, the process itself may be contributing to how people frame new data.

[1] The "boiled frog" scenario is referred to often in disaster and business communities to describe "creeping" disasters and crises, it describes a frog that when placed in a pot of cold water which is being gradually heated, will fail to recognize the increasing danger and thus get boiled alive. However, if this proverbial frog is put directly into hot water, it will recognize the dangerous situation and jump straight out of the pot. This metaphor thus describes a number of problems that can arise with creeping disasters: failing to recognize the accumulation of many small changes which can amount to a major crisis; normalization bias; and personnel fatigue and performance issues for long duration events.

Table 6
The total number (N_t) of occurrences of all identified themes from the free text responses for each experiment group (e.g., scientists in the numerical condition) in experiment 1, using the themes as described in Table 4. The number (N), and percentage (%) of occurrences of each particular theme relative to N_t, is also shown for each experiment group. Note that not all participants indicated whether they were non-scientists or scientists, and so for some themes there is a difference in the total number for an experimental group, as illustrated in Table 5.

Experiment 1	**Numerical**				**Verbal**			
N_t (number of occurrences of all themes in each experiment group):	**Science 93**		**Non-science 78**		**Science 80**		**Non-science 91**	
Number (*N*) and % of each theme occurrence relative to all theme occurrences								
Evacuation thresholds	**8**	9%	**7**	9%	**4**	5%	**3**	3%
Evacuation technicalities	**13**	14%	**16**	21%	**7**	9%	**19**	21%
No time to lose	**15**	16%	**17**	22%	**9**	11%	**17**	19%
Let's wait and see	**6**	6%	**5**	6%	**5**	6%	**5**	5%
Managing uncertainty (MU)	**17**	18%	**8**	10%	**19**	24%	**16**	18%
MU – cost benefit	**2**	2%	**1**	1%	**1**	1%	**2**	2%
MU – assumptions made	**4**	4%	**0**	0%	**2**	3%	**0**	0%
MU – incomplete information, information needed	**9**	10%	**5**	6%	**11**	14%	**9**	10%
MU – level of certainty or confidence	**2**	2%	**2**	3%	**5**	6%	**5**	5%
Time forecasts	**4**	4%	**2**	3%	**4**	5%	**3**	3%
Probability translations	**3**	3%	**1**	1%	**4**	5%	**1**	1%
Science advice, understanding, role	**5**	5%	**4**	5%	**6**	8%	**2**	2%
Public action	**4**	4%	**5**	6%	**1**	1%	**5**	5%
Public perception	**0**	0%	**2**	3%	**2**	3%	**3**	3%
Future issues – negatives, positives	**1**	1%	**3**	4%	**0**	0%	**1**	1%

All scenarios in this questionnaire provided limited information about the town, to enhance the uncertainty of the scenario. For example, participants were not told how far away the town was from the volcano, or what type of volcano it was. Through this design, we aimed to identify what they judged to be the main factors affecting their decisions, and how they managed the uncertainty. As noted in Section 1, Lipshitz and Strauss [59] showed that decision makers reduced, acknowledged, or suppressed the uncertainty in order to make a decision (see Table 1). In our themes defined above, the theme *let's wait and see* and related comments can be considered to fall into the "reduced" uncertainty category of Lipshitz and Strauss [59] as individuals defer (or avoid) their decision until additional information becomes available, or until after they can reinterpret incremental data. Considering just those comments that display approaches to *managing uncertainty*, the majority refer to having incomplete information, and that they would solicit more advice and information or monitoring data to make their decision, or make a change to the type of evacuation (again in the 'reduced' category of Lipshitz & Strauss).

- *There would need to be a lot more information from survey points etc. prior to a full evacuation, however judging from this information, I would be focusing on an evacuation and securing the area by early Friday morning.* [non-scientist, verbal].

This theme *incomplete information, information needed* occurred more in response to the verbal statement (19 occurrences) than the numerical statement (14 occurrences) (Table 5). Considering both statements, scientists commented more about this lack of information than non-scientists, where the theme occurred 20 times in the scientists' comments and 14 times in the non-scientists'. These scientists stated a reluctance to make a decision with this incomplete information, with their focus being on the lack of information about impacts and the volcano type and hazards.

- *Evacuation or not depends entirely on likelihood, type, severity of impacts on Awha. We're not given this information. For example if Awha is upwind of the volcano, beyond ballistic impact zones and away from lava, lahar or pyro*[clastic] *flow paths, then evacuation is unlikely to be needed. Volcano status has to be continuously monitored and EM's kept updated of change.* [scientist, numerical]

The theme *assumptions made* (as part of the 'reducing' approach to managing uncertainty of Lipshitz and Strauss [59]; Table 1) only occurred 4 times in the responses to the numerical statement, and 2 times in the responses to the numerical statement (Table 5), and those assumptions raised were primarily about the impacts as well as evacuation processes and the evolution of activity throughout the week. However, it is likely that many other assumptions were made, but not referred to.

- *My assumption is that any scale for the eruption would affect Awha and be a threat to its inhabitants. The other assumption is that an evacuation plan is ready and the authorities operational.* [scientist, numerical].
- *I assume that if a 2-week window is given, activity level is ramping up but not yet close to a critical level. But I*

would wait only a few days though, because rate could accelerate rapidly, and 48 h could be too long if precursors accelerate [scientist, verbal].

As would be expected based on our analysis so far, the theme *level of certainty or confidence* occurred more in the responses to the verbal probability statement than for the numerical statement (Numerical: 10, Verbal: 4; Table 5):

- *Without knowing what level of confidence 'likely chance' represents in this scenario it's not possible to recommend an evacuation date.* [scientist, verbal].

Lipshitz and Strauss [59] also highlighted that individuals 'acknowledged' uncertainty as an approach to decisions under uncertainty, and from the themes analysed so far, the use of *evacuation thresholds* can be seen as a way of acknowledging uncertain information. A number of participants specifically weighed up the pros and cons of different approaches as a way of acknowledging uncertainty, using a *cost-benefit approach* for *managing uncertainty*, with 6 occurrences of this theme in the responses to both the numerical and verbal conditions. For example:

- *Necessary to determine the cost, in lives, of the evacuation against the cost, in lives, of an eruption. Only then are you able to determine which is the best option that minimizes loss of life. Both scenarios have the potential to cause the loss of life, and numerous variables must be considered.* [scientist, verbal].
- *In terms of managing risk, the likelihood (high) combined with the consequence (unstated) results in the level of risk. This, combined with proximity (within two weeks) must be balanced with the veracity of the information and the opportunity cost of evacuating.* [non-scientist, verbal].

Very few comments mentioned tactics that might fall within Lipshitz & Strauss's approach of 'suppressing' the uncertainty through ignoring it, relying on intuition, or rationalizing away doubts, an approach that reflects more *affective* type of processing (see Section 1.1). However, our analysis is based only on those people that chose to comment and share their thoughts, and this strategy may have been adopted but not reflected in the comments, especially if, due to the intuitive nature of the decision, participants did not feel the need to comment.

3.1.6. Public action, public perception and future issues

A number of respondents raised issues relating to *public action*, in terms of preparing the public for an evacuation, that a self or voluntary evacuation may occur, or that the public would be non-compliant to an evacuation order

- *If it is likely to occur in the next 2 weeks, residents should be evacuated now; as certainly as some portion of the population will not comply to evacuation orders, once the eruption begins, efforts would have to be focused on a rapid evacuation of that smaller population who did not evacuate* [non-scientist, verbal].

These three themes occurred 15 times in response to the numerical statement, and 12 times in response to the verbal statement (see Table 5); and for both statements the majority of these occurrences originated from non-scientists reflecting their response concerns and roles, with 19 occurrences for the non-scientists' responses and 8 for the scientists' responses (see Table 6). The issue of compliance to warnings is closely related to *public perception* issues raised by a number of participants:

- *The population can see and feel the volcanic activity. They need assurance that authorities are also aware and are responding and advice of what they should or should not do. This includes preparing the public for the possibility of evacuation.* [non-scientist, verbal].

There is some discussion in the disaster and emergency management community that a fear of non-compliance to future warnings may prevent emergency managers from issuing an evacuation order in response to an event that may not occur. This fear of being labelled 'cry-wolf' and the damage that it may have on future warning compliance has anecdotally been said to affect key decision makers [34,27,94,111,93], even though there is little empirical evidence for this behaviour [61]. This behaviour could reflect a move towards an *affective* style of processing rather than an *analytical* style (see Section 1.1) by the decision makers, whereby their (lack of) previous (analytical) experience (of volcanic uncertainty), acts to heighten levels of anxiety and a lack of perceived control (i.e., it activates the 'dread' aspect of risk) to constrain their decision making from a logical reason based approach towards one mediated by past experience [24], with social and political concerns (e.g., the cost of an evacuation) coming to dominate how they frame their interpretation. This elevation of anxiety may also affect their belief or trust that the community will respond, resulting in a fear of being labelled 'cry wolf', further impacting their decision making. This issue was indeed raised as a decision consideration, and was specifically mentioned by 3 participants:

- *You need to pay credence to expert advice. My opinion better to err on the side of caution than to risk lives. There is the risk factor then of 'the boy who cried wolf' if the eruption does not occur. This has to be managed as best it can.* [non-scientist, numerical].
- *To play it perfectly safe, we should evacuate now. However, if we do and nothing happens, there is a high probability that the next time the people will be less likely to heed our warning.* [non-scientist, numerical].
- *It really depends on the volcano – evacuations can be done too early losing the trust of those living alongside the volcano; or too late, resulting in mass casualties and deaths.* [non-scientist, verbal].

Related to this, participants commented on a number of other *future issues*, including the view that any evacuation is an "*opportunity to validate plans and provide experience for responders, receivers, and the general public*" [non-scientist, numerical]. In addition, some participants considered that an earlier evacuation "*may provide for a better qualitative result for taking care of complex human needs and long term recovery during all phases of an orderly and safe evacuation*" [non-scientist, verbal].

3.1.7. Evacuation technicalities

The other dominant themes identified in the free response comments were those that referred specifically to *evacuation technicalities*, ranging from the actual evacuation process, to the evacuation time scheduling, and a consideration of complexities such as humanitarian needs, pets, conditions (daylight), and what type of evacuation should be issued (mandatory or voluntary), as well as whether it should be staged or prioritized. This theme was very prevalent in all responses, and occurred 29 times in the responses to the numerical statement, and 25 times in the responses to the verbal statement (Table 5). As expected, the *non-scientists* raise these issues more than the *scientists*, particularly with the verbal statement where the theme occurred 19 times in the non-scientists' responses and 7 times in the scientists' responses (see Table 6).

- *In the absence of any further forecasts I would begin a prioritized limited evacuation (e.g., nearest areas, congested/remote areas, mobility-impaired populations) and make preparations for an accelerated "mass" evacuation should indications warrant.* [non-scientist, numerical].
- *There may be a staged evacuation of vulnerable communities immediately with other people encouraged to self evacuate and a forced evacuation pending dependant on further information being available.* [scientist, numerical].

3.2. People who chose not to evacuate

For the **numerical** statement, 9.6% of participants chose not to evacuate (4.9% of scientists, 15.6% of non-scientists). For the verbal statement, 28.2% chose not to evacuate (30.3% of scientists, 26.3% of non-scientists). Considering first the numerical statement, very few reasons or explanations were given for why people chose not to evacuate. For those that did choose to comment, reasons included probability *evacuation thresholds* and that "*there is not a high enough likelihood to evacuate the population – there is still a 17% likelihood that there won't be an eruption*" [non-scientist, numerical], and that there "*is a high degree of uncertainty – would not alarm me*" [scientist, numerical]. Others chose to *let's wait and see*, citing fear of falsely alarming the public (*public perception*) and that "*if we do* [evacuate] *and nothing happens, there is a high probability that the next time the people will be less likely to heed our warning*" [non-scientist, numerical], or that there is not enough accuracy to "*shut down the city*" [non-scientist, numerical]. Others who chose not to evacuate explained that they would get the evacuation planning and preparation underway now (*evacuation technicalities*), noting that they would have their "*plans dusted off*" [non-scientist, numerical], and would "*increase public awareness and preposition some resources to expedite the evacuation process if needed*" [non-scientist, numerical].

However, for the **verbal** statement, there were many occurrences of themes that related to non-evacuation for those that chose not to evacuate (see Tables 4–6), reflecting the higher number of people who chose not to evacuate. The majority of these comments referencing non-evacuation fell in the *incomplete information* theme (12 occurrences, see Tables 4 and 5), where respondents' approach to *managing uncertainty* was to reduce it by waiting for further information about the monitoring data and the hazards, and updates in the probability forecasts

- *More definitive explanations would be required. The city should be on 'high alert' with residents advised that a rapid evacuation may be necessary at any time.* [non-scientist, verbal].

Others stated issues around probability *evacuation thresholds* and the level of confidence, stating that "*a decision for evacuation can be made when there is greater certainty*" [non-scientist, verbal]. The limitations of the science advice and probability forecasts were also raised:

- *Without knowing what level of confidence 'likely chance' represents in this scenario it's not possible to recommend an evacuation date. e.g., if 'likely chance' is only the second step on a ten step warning scale then the answer is a safe no. If, however, 'likely chance' is the highest level of threat on the scale then I would say yes and order the evacuation to begin immediately, especially if there was no further chance of evidence to indicate the eruption was more or less imminent.* [scientist, verbal].

A number of respondents suggested that *evacuation technicalities* get under way in preparation for any future response or that a limited prioritized eruption should be started, with an increase in dialogue with the scientists:

- *A determination of "likely" chance of eruption within the next two weeks should result in increased emergency operations – i.e. increased situational awareness, increased monitoring, and more frequent updates/briefings; it should trigger evacuation preparedness activities and contingency operations, not necessarily the evacuation itself. The second week, if conditions remain the same, I would begin evacuation of special needs populations – i.e. those that cannot evacuate themselves.* [non-scientist, verbal].

As with the numerical statement, the issue of *public perception* was raised. However, in terms of needing enough information or having high enough levels of risk to "*justify an evacuation plan of this magnitude* [scientist, verbal], with some participants also choosing to 'acknowledge' the uncertainty through a *cost-benefit* approach".

In sum, these comments indicate that respondents who chose (or reacted to uncertainty) not to evacuate did not completely discount the risk, as the majority decided to monitor the situation and delay an evacuation decision until more advice was available, and to ready resources and the public for such evacuation if needed. The majority of these latter comments came from *non-scientists*, reflecting the responsibilities that those individuals would have during an eruption. As expected, more respondents in the verbal group (likely) adopted this latter *wait and see* approach, reflecting the lower perceived likelihood for that statement compared to the numerical equivalent (73–83%). For the numerical statement, the people who chose not to evacuate cited the possibility of a non-event and the damage to public perception of the authorities, with some still readying resources.

Interestingly, when we compare all comments raised by the *scientists* and *non-scientists* across both conditions, the scientists tended to query the scientific advice and express dissatisfaction about the quality and lack of information as reasons for not evacuating, whereas the non-scientists tended more towards expressing a desire to wait for more information while however preparing for evacuation. This can be seen in the prevalence of themes illustrated in Table 6, where the themes 'evacuation technicalities' and 'let's wait and see' occur more often for non-scientists than the scientists; and themes discussing the 'science advice' and the 'quality of the information' relating to ways of 'managing uncertainty' occur more often for the scientists than the non-scientists.

More of the non-scientists chose an evacuation even with the limited scientific information available than the scientists, commenting on the process and technicalities and the *actions* they would take, thereby *acknowledging* or *suppressing* the uncertainty [59]. The scientists however focused more on the *information* and the quality of the information, choosing to *reduce* the uncertainty in the source (information) before proceeding.

4. Experiment 2: a higher likelihood scenario

We now investigate whether the evacuation decision is affected by the probability format when a higher likelihood scenario is considered. For this, the numerical group were presented with "… *a 45–55% probability of an eruption occurring within the next 3 days*", and the verbal group with "… *a medium likelihood chance of an eruption occurring within the next 3 days*" (see Fig. 2). Results are presented in Fig. 4, and we find that for the numerical group, 72.9% of respondents choose an evacuation ($n=70$, 48.6% on Monday morning, omitting non-responders throughout), while for the verbal group, 71.8% choose an evacuation ($n=71$, 43.7% on Monday morning). We find no significant relationship between the probability format and the evacuation decision (see Table 7). Considering only scientists, 69.7% evacuate in the numerical group ($n=33$) and 71.1% in the verbal group ($n=38$); while for the non-scientists, 75% evacuate in the numerical group ($n=36$) and 72.7% in the verbal group ($n=33$). Again, no significant difference is found between responses to the numerical and verbal phrasing for each of these career groups (see Table 7). These results demonstrate that the different evacuation decisions observed between participants presented with verbal and numerical probabilities for the 2 week scenario (experiment 1: Figs. 1 and 3), is not observed for the shorter time frame scenario (experiment 2: Figs. 2 and 4). It appears that the shorter scenario is reducing any effects due to the different probability format.

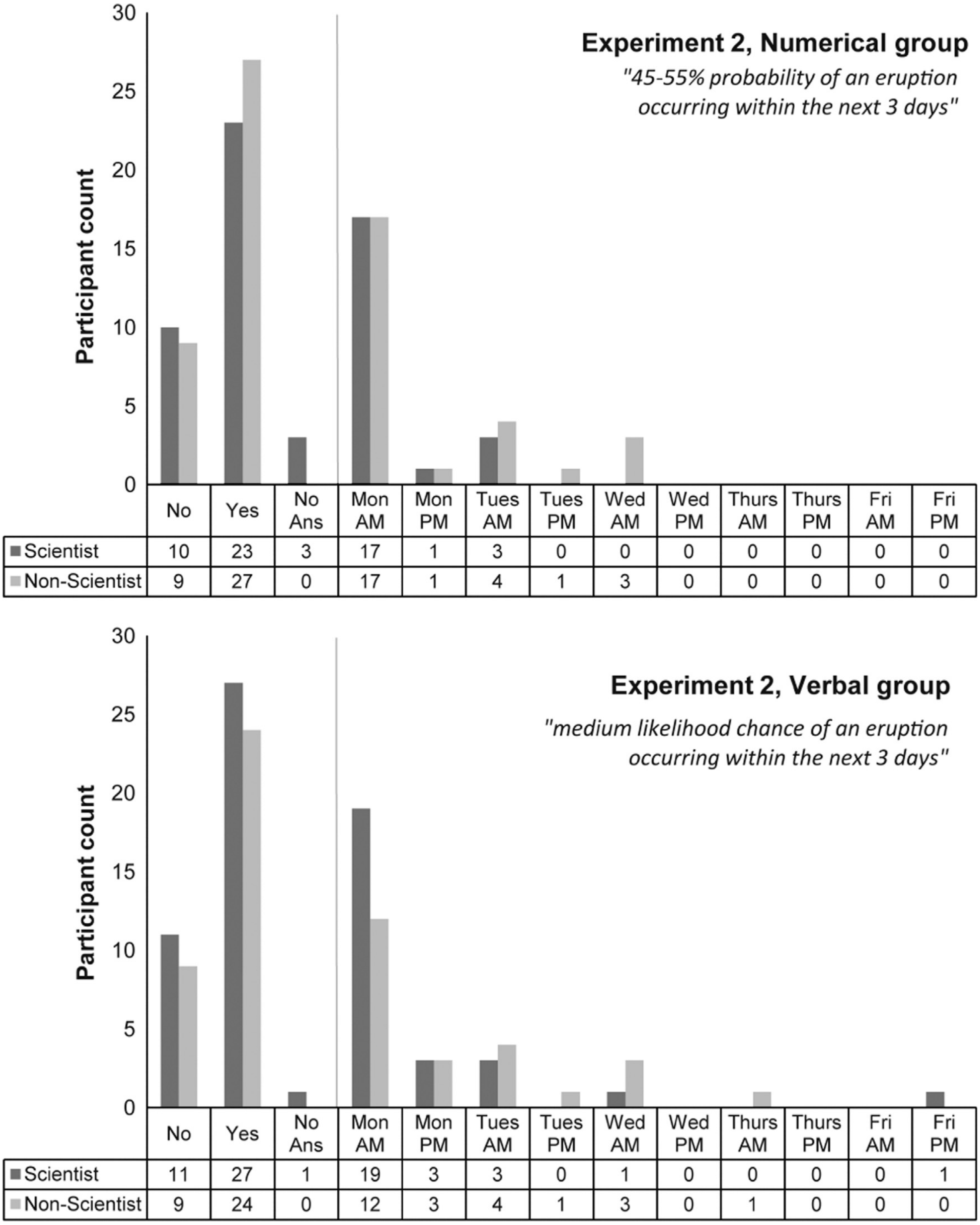
It is interesting to also directly compare the decision behaviour across the groups presented with the numerical statements in experiments 1 and 2, as experiment 2 (45–55%, 3 days) represents a higher likelihood, or more certain, scenario than experiment 1 (73–83%, 2 weeks). To explain, if we were to hypothetically consider the likelihood throughout the time windows to be independent on each day, and of equal likelihood distribution (i.e. neither decreasing or increasing throughout the stated time), then the probability on Monday for experiment 1's 2 week statement is 9 to 12%, while the probability on Monday for experiment 2's 3 day statement is 18–23%. Admittedly we cannot infer these values directly, as the scenarios provide no information about distribution throughout the week (it could increase, decrease or plateau) or any conditional probability between days. However, based on the information provided in each statement, it is reasonable to assume that the second shorter scenario corresponds to a higher likelihood on the Monday.

Interestingly, however, when we directly compare the decision behaviour between experiments for all participants, more participants choose an evacuation for experiment 1 (90.4%) than experiment 2 (72.5%). This significant difference ($\chi^2(1, N=143)=7.402$, $p=0.007$, medium effect size $V=0.228$) is in the opposite direction to what the logic described above would predict, whereby experiment 2 should have a higher number of evacuations. It seems that participants fail to recognize the influence of the stated time window on the overall likelihood. Rather than evacuate in greater numbers for the higher likelihood situation (experiment 2, 45–55%, 3 days), more participants evacuate for the situation with the higher probability numbers (experiment 1, 73–83%, 2 weeks) failing to recognize that it represents a lower likelihood on any given day within the time window. It appears that even though some participants commented that the shorter time window "*increased the risk*" (see qualitative analysis below), many participants base their decision on the value of probability communicated (73–83% versus 45–55%) and do not account for the different time windows (2 weeks versus 3 days).

For the statements that utilize the verbal probabilities, logic would again indicate that experiment 2 (medium likelihood, 3 days) represents a higher likelihood, or more certain scenario, than experiment 1 (likely, 2 weeks). However, the same proportions of respondents chose an evacuation for both scenarios (71.8% in each case). These data show the same issues as for the numerical statement, but to a lesser degree due to the uncertainty and variance in perception inherent in these verbal terms.

4.1. Reasons given for evacuation decision

The reasons given for evacuation are not dissimilar to those provided in response to experiment 1's long time window scenario (see Section 3.1), and so we only briefly



	No	Yes	No Ans	Mon AM	Mon PM	Tues AM	Tues PM	Wed AM	Wed PM	Thurs AM	Thurs PM	Fri AM	Fri PM
Scientist	10	23	3	17	1	3	0	0	0	0	0	0	0
Non-Scientist	9	27	0	17	1	4	1	3	0	0	0	0	0

	No	Yes	No Ans	Mon AM	Mon PM	Tues AM	Tues PM	Wed AM	Wed PM	Thurs AM	Thurs PM	Fri AM	Fri PM
Scientist	11	27	1	19	3	3	0	1	0	0	0	0	1
Non-Scientist	9	24	0	12	3	4	1	3	0	1	0	0	0

Fig. 4. The evacuation decision or not, and the time chosen to evacuate, for experiment 2 (Fig. 2) and numerical probability experiment group (45–55%) and the verbal probability experiment group (medium likelihood). Results are shown for participants who identified as scientists (dark grey) and non-scientists (light grey).

describe them here. However, there are interesting differences in the distribution and prevalence of the theme occurrences in the free text responses to the two scenarios (Tables 5, 6 and 8), as follows.

Considering first the group that received the **verbal** statement (medium likelihood) a greater proportion of the comments mentioned a *cost-benefit* approach to *managing uncertainty* for this short time window statement (7 theme occurrences) compared to the previous long time window scenario (3 theme occurrences); *evacuation thresholds* were also discussed more often (Exp 1: 7 occurrences, Exp 2: 11 occurrences), as well as issues relating to *public*

Table 7
The total number (N) of participants in each experiment and group and the percentage that chose to evacuate (% Yes), where non-responses were omitted from the analysis. Results are shown for each experiment group (presented with either the numerical or verbal probabilities), and for all participants, as well as for just those that identified as scientists and non-scientists. The percentage that chose to evacuate on Monday (experiment 1), and on Monday AM (experiment 2), is also shown (% Mon). In addition, χ^2 test results comparing the number who chose to evacuate for the numerical vs. verbal group are also shown, and for those found to be significant the effect size (V) is given.

	Numerical		% Mon	Verbal		% Mon	χ^2 Test results and effect size
	N	% Yes		N	% Yes		
Experiment 1: '73–83%' or 'likely' … within the next 2 weeks							
All	73	90.4	46.6	71	71.8	45.1	$\chi^2(1, N=144)=7.914$, $p=0.005$, medium effect size $V=0.234$
Scientists	41	95.1	53.7	33	69.7	42.4	$\chi^2(1, N=74)=8.699$, $p=0.003$, moderate effect size $V=0.343$
Non-scientists	32	84.4	37.5	38	73.7	47.4	$\chi^2(1, N=70)=1.179$, $p=0.278$
Experiment 2: '45–55%' or 'medium likelihood' … within the next 3 days							
All	70[a]	72.9	48.6	71	71.8	43.7	$\chi^2(1, N=141)=0.19$, $p=0.892$
Scientists	33	69.7	51.5	38	71.1	50	$\chi^2(1, N=71)=0.016$, $p=0.901$
Non-scientists	36	75	47.2	33	72.7	36.4	$\chi^2(1, N=69)=0.046$, $p=0.830$

[a] One participant in this group chose not to declare their career sector.

Table 8
The total number (N_t) of occurrences of all identified themes from the free text responses for each experiment group (e.g. scientists in the numerical condition) in experiment 2, using the themes as described in Table 4. The number (N) and percentage (%) of occurrences of each particular theme, relative to N_t, is also shown for each experiment group. Note that not all participants indicated whether they were non-scientists or scientists, and so for some themes there is a difference in the total number for an experimental group as illustrated in Table 5.

Experiment 2	**Numerical**				**Verbal**			
N_t (number of occurrences of all themes in each experiment group):	**Science** **49**		**Non-science** **56**		**Science** **66**		**Non-science** **52**	
Number (N) and % of each theme occurrence relative to all theme occurrences								
Evacuation thresholds	**3**	6%	**7**	13%	**4**	6%	**7**	13%
Evacuation technicalities	**4**	8%	**11**	20%	**14**	21%	**9**	17%
No time to lose	**9**	18%	**7**	13%	**11**	17%	**10**	19%
Let's wait and see	**1**	2%	**7**	13%	**8**	12%	**4**	8%
Managing uncertainty (MU)	**10**	20%	**9**	16%	**10**	15%	**5**	10%
MU – cost benefit	**1**	2%	**3**	5%	**6**	9%	**1**	2%
MU – assumptions made	**1**	2%	**1**	2%	**1**	2%	**0**	0%
MU – incomplete information, information needed	**8**	16%	**4**	7%	**3**	5%	**3**	6%
MU – level of certainty or confidence	**0**	0%	**1**	2%	**0**	0%	**1**	2%
Time forecasts	**2**	4%	**0**	0%	**0**	0%	**1**	2%
Probability translations	**1**	2%	**1**	2%	**3**	5%	**0**	0%
Science advice, understanding, role	**4**	8%	**3**	5%	**2**	3%	**1**	2%
Public action	**2**	4%	**1**	2%	**3**	5%	**5**	10%
Public perception	**1**	2%	**1**	2%	**1**	2%	**3**	6%
Future issues – negatives, positives	**2**	4%	**0**	0%	**0**	0%	**2**	4%

action (Exp 1: 6 occurrences, Exp 2: 8 occurrences). A comparison of the prevalence of themes for each experiment, as indicated in Tables 6 and 8, illustrates that more of the *non-scientists* commented on both these issues (see Table 9 for notable comments). However, fewer of all participants discussed methods for *managing uncertainty* in terms of *assumptions made* and issues relating to *incomplete information* compared to the previous longer time window verbal scenario. There was also less discussion of the *science advice* and any limitations around that advice, or the understanding and translations of the *probability forecast*s or time windows. Comments relating to *no time to lose* and evacuating *ASAP* or initiating evacuation processes or adopting the precautionary principle were discussed about as often in response to this verbal scenario as the long time window verbal scenario in experiment 1 (see Tables 6 and 8).

The only major difference observed for the **numerical** statement compared to the longer time window used in experiment 1, was that there were fewer comments about there being *no time to lose*, and few comments about it being better to evacuate earlier, to adopt the precautionary principle, or to prepare resources for evacuation (see the prevalence of themes indicated in Tables 6 vs. 8, with notable comments in Table 10). This is surprising, as strictly speaking, this scenario has a higher likelihood than the scenario in experiment 1. However, it is consistent with our quantitative analysis, where fewer participants evacuate for experiment 2 (72.5% evacuations for "45–55%, 3 days" versus 90.4% evacuations for "73–83%, 2 weeks").

Table 9
Notable comments raised in response to the verbal statement of experiment 2 (medium likelihood, 3 days).

Experiment 2, Verbal probability statement: non-scientists

- Need something more than a medium likelihood. Concern with evacuating the population and the "medium likelihood" persists for weeks or months, how do you allow people to return when the risk remains the same as when they were initially evacuated.
- Again, marshal the resources and be ready to issue the order, but I would not tell the population to leave based on a medium likelihood. I seriously doubt that even if an evacuation order was issued that many people would leave, based on that probability.
- Even a 50% chance of an eruption is too high not to evacuate immediately.
- Need more specifics than "medium likelihood". If you were to evacuate based on that information the "medium likelihood" may persist for several weeks or months and if that were to happen you would have a large percentage of the population returning to their homes due to evacuation fatigue, and it would be much more difficult to get them to evacuate a second time.
- 3 Days isn't very long. better to err on the side of caution, better to be criticized for something you did rather than something you didn't do...
- I would begin phased evacuation preparedness activities and operations – specifically – evacuation of non-ambulatory medical special needs populations, requesting hospitals transfer patients out of the area, cancel elective surgeries, prepare to increase urge capacity, cancel essential personnel vacations, etc.
- There is not enough information to justify an evacuation.
- I would understand the statement to say execute the plan to marshal the resources. At some point within the next 48 h the decision to move with or without the resources in place maybe necessary. I would recommend executing the plan. At a minimum, the most time intensive elements.
- Planning and resourcing needs to be done asap so that the evacuation can be done as soon as possible when the eruption is imminent

Experiment 2, Verbal probability statement: scientists

- Again the balance between loss of life and Livelihood(evacuation and eruption) is important.
- Begin evacuation immediately. The financial loss of an evacuation for only a few days is not worth the risk of full destruction of your capital city. PR-wise, overabundance of caution is also less damaging than failure to protect your people (including our government people) from catastrophe. If you evacuate your government but not your citizens, there would be huge loss of faith and trust in the government.
- Not clear likelihood of eruption high enough to warrant evacuation – and I'm glad it's only hypothetical
- Any evacuation may be done in terms of identifying those most vulnerable and providing information to the rest of the community to enable them to make their own decision as to whether to self evacuate.
- If the primary goal is to save lives then why wait? Beginning the evacuation process could be as simple as mobilizing transport, coordinating relief centres and identifying those with special needs (e.g. transportation, livestock, grain storage, etc.). his should begin immediately.
- This is answer involves taking a precautionary principle approach on the assumptions that 'medium likelihood' is a higher than even chance probability and that this warning represents a new increase in the level of warning. If the scenario said this was the 10th day in a row the scientists had given this same bulletin then I might answer differently. In this scenario the trend is as or more important than the probability alone.
- Same reasoning as previous answer, more urgent given the time frame is shorter, even though the probability is lower (my perception of lower at least).
- In simple terms I believe a medium likelihood to be a 50/50 probability of an eruption and therefore preparations for an evacuation to be a prudent measure
- What does medium likelihood mean, that is the key, isn't it? I would interpret it to be $>50\%$ which means it is prudent to begin evacuation.
- I feel the mobilization process should begin now and let the volcanic activity develop to the point of needing evacuation. The proper authorities will be ready and the evacuation should run more quickly.

They fail to recognize it is a higher likelihood, choose to evacuate less, comment less about evacuating straight away or *ASAP,* and generally not feeling the need to comment on why they chose an evacuation or to justify their decision.

Considering only experiment 2, for the verbal statement there were slightly more comments relating to *incomplete information* and ways of *managing uncertainty* compared to the equivalent numerical statement (see prevalence of themes indicated in Tables 6 and 8), with participants stating that they would reduce this uncertainty by waiting for more advice as well as acknowledging the uncertainty through a *cost-benefit* approach. For these participants, this may reflect their view that the verbal statement represents a lower likelihood and greater uncertainty. This reflects the greater ambiguity in verbal statements and the greater complexities in deciding on evacuations when that phrasing is used.

4.2. People who chose not to evacuate

For the verbal statement, the people who chose to not evacuate commented mostly on *evacuation thresholds,* noting that the likelihood was not high enough for evacuation, *cost-benefit* issues between the cost of evacuation and the damage to the *public perception* and any future compliance due to evacuation fatigue. Some also noted that *evacuation preparations* should be made considering *evacuation technicalities* such as encouraging voluntary evacuation rather than issuing any mandatory evacuation order. For the numerical statement, people who chose not to evacuate commented mostly on *evacuation thresholds* and noted that the stated probability is "*too low to make a decision evacuate*", and discussed issues relating to the *lack of information* and the related uncertainty preventing them from making a decision.

5. Discussion

Many factors affect the understanding of probability forecasts and influence decision making [21,20]. These include framing, past experience, time and risk pressures, base rate, background knowledge, outcome severity, the level of uncertainty, competing time frames, and political and economic factors (see Fig. 5). The results here show that evacuation decisions are further influenced by (1) perceived higher levels of certainty for numerical probabilities over verbal likelihoods, particularly for scientists;

Table 10
Notable comments raised in response to the numerical statement of experiment 2 (45–55%, 3 days).

Experiment 2, Numerical probability statement: non-scientists
• This is dependent upon local risk tolerance and the quality of predictions provided by the volcanologists – assuming high reliability for the prediction and low risk tolerance the better option is to at least begin evacuation of vulnerable populations early. • You can't risk people's lives – order the evacuation straight away – give yourselves plenty of time to get out • A 50:50 chance of significant death cannot be dismissed comfortably. • It seems like it's worth waiting for better information. • Tricky one – depends on whether a higher probability warning can be given with a 48 h window. But hard to justify an evacuation on a 50% probability. • There is a less than 55% chance of an eruption. At this point I would encourage people to have an evacuation plan, and if they have somewhere else they can go, it would be advisable. But Not a full evacuation. • The evacuation process should start when the volcanologists raise the probability somewhat. 45–55% isn't enough to make me want to displace the residents and the seat of government. Another 15–20% probability, and I'd recommend initiating an evacuation. • Evacuation is almost never an all-or-nothing proposition. I would begin an immediate prioritized evacuation of most vulnerable (including nearby and isolated) and mobility-impaired populations while preparing for accelerated evacuation of the remaining population should conditions deteriorate. • Requires an instant decision and public safety would take precedence over the disruption • Need something more than a medium likelihood. Concern with evacuating the population and the "medium likelihood" persists for weeks or months, how do you allow people to return when the risk remains the same as when they were initially evacuated.
Experiment 2, Numerical probability statement: scientists
• Extent of risk combined with likelihood make the situation too hazardous to ignore. • Presumably, you have some information about the intensity of the eruptions from Pomona – if the eruptions are always extremely violent, then you begin the evacuation on an up trending intensity of precursory evidence. If the eruptions are of modest intensity, then prepare for an evacuation but await progress toward a higher probability. • Again, what is the expected impact of the eruption? If "there is a 45–55% probability of a large number of casualties within the next 3 days, if no evacuation is undertaken", then my answer is "Yes". The evacuation should start Monday AM. • If it's ignimbrite eruption yes evacuation on Monday morning; if it's fire fountain, no evacuation. I don't think a 50% chance is a good probability to be given by scientist.. Throwing a coin is better! • 1/2 within 3 days is high. I think we need a probability trend within the next 7 and 15 days to take a more reasonable decision. • I believe the probability is too low to warrant an evacuation just yet, since evacuating and then not having an eruption would lower the credibility of the volcanologists in the future, when the probability of an eruption could be much higher. • May still raise awareness and preparedness so that time for evacuation is reduced from 48 h to less. • 50–50 chance of it happening, and any prep done for evacuation is probably better than none. chance is that you'll save lives. • The 3-day time window suggests rapidly accelerating unrest, so no time to lose, even with a 50/50 probability • I'm not enough of statistician to know how to phrase the problem that's bothering me, but in essence I feel that "X% in *T* time" does not contain enough information about how *X* will be reassessed as time passes for me to know how to interpret it • Review probability daily and if an increase then move to evacuate. Suggest people voluntarily evacuate.

(2) a poor understanding of the relationship between the probability value and the time window stated within a forecast; and (3) career sector, and thus background training and world view, with *non-scientists* focusing more on potential actions that should be taken and *scientists* focusing more on the quality and volume of the information available in their reasons for evacuation or not. Each of these is discussed in turn in the following sections.

5.1. Higher levels of certainty for numerical probabilities

Evacuations occur less frequently when a verbal likelihood is communicated instead of the numerical equivalent. To chose these equivalent translation terms we used the IPCC translation table (Section 1.2 and Table 2), in line with our previous studies [21,20]. The decision results examined in experiment 1 suggest that participants do not see the terms "73–83%" and "likely" to be equivalent at all, with scientists in particular making different decisions in response to these different terms. This is consistent with the results of Doyle et al. [21], who found that for this study group 'likely' is translated to $68 \pm 14\%$ and "very likely" is $83 \pm 10\%$, values which overlap with the IPCC translation table boundaries. If we modify the translation used herein to the alternative range suggested by Doyle et al. [21], we find that the equivalent of the verbal statement 'likely' is $68 \pm 14\%$, which is very slightly lower than the equivalent IPCC numerical value used herein (73–83%), and this difference may explain some of the differences in evacuation rates. However, it is unlikely that this translation difference is the sole reason that participants issue evacuations less frequently for the verbal likelihood, and the free text responses suggest that the main reason for these lower rates of evacuation are that the verbal terms (*likely, medium likelihood*) are viewed as more vague, are more ambiguous and represent a less certain situation than their numerical equivalents (73–83%, 45–55%). When presented with the verbal term, more participants acknowledged the possibility of a non-event, or chose to wait for the situation to evolve before evacuating by adopting the "wait and see" approach. They also discussed more of the complexities and problems that may be considered with any evacuation. In contrast, the group that received the numerical probability commented much more about the urgency of the situation, concentrating on evacuation processes and technicalities. This is in line with the finding of Budescu et al. [9] that the verbal probabilities communicated in the 2007 IPCC report may have

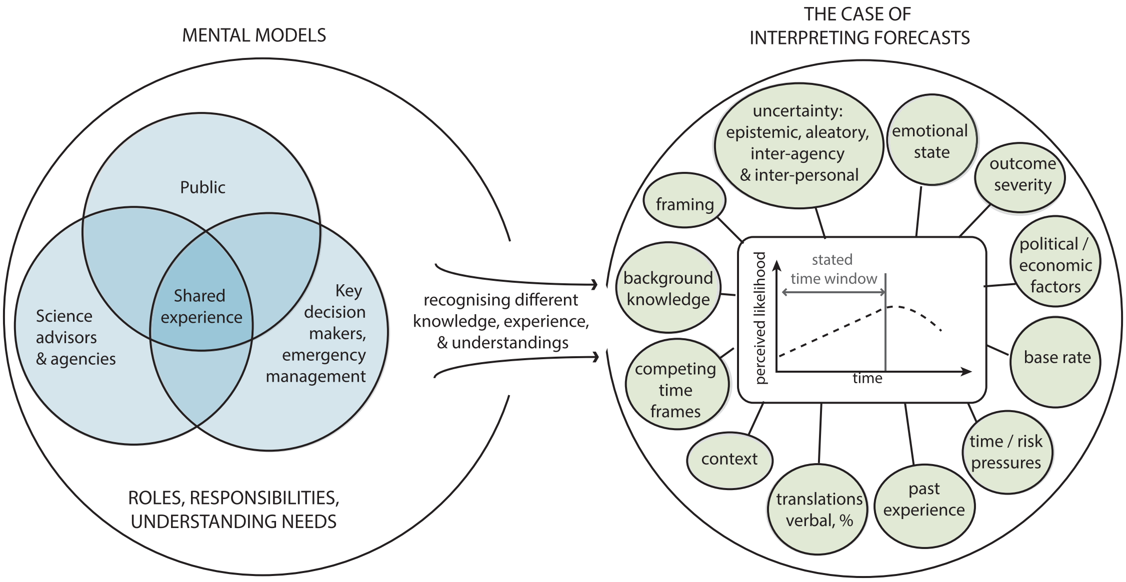


Fig. 5. The factors that affect the interpretations of probabilistic forecasts and the influences on resultant decision making, after the graphical abstract of Doyle et al. [20].

implied higher levels of imprecision than were actually present. The majority of these differences are due to the judgments of scientists, suggesting they are more sensitive to the format of the probability statements. These scientists are more likely than the non-scientists to choose an evacuation for the numerical statement rather than the verbal. This difference in behaviour between these two groups may be due to the scientists having a greater familiarity with numerical probability statements, eliciting a greater degree of confidence in numerical formats over verbal formats.

We also find that a number of participants chose to translate the verbal statement into a numerical percentage probability, or a natural frequency, before making a decision. This increases the cognitive load as well as the potential for translation mistakes, biases and miscommunication. It is clear that the choice to use either verbal or numerical terms in a forecast does not just affect translations and likelihood perceptions, but has clear and potentially larger impacts on evacuation behaviour and reasoning. This behaviour is not observed for the short time window examined in experiment 2, and as with the framing and job sector effects discussed in Doyle et al. [20], this is likely to be due to the shorter time window masking these differences.

When we consider the recommendation to utilize translation tables to standardize equivalent numerical and verbal probability terms [33,9,64,20], we must also be aware that failure of a scheme may not just be due to the different hazard under consideration (volcanology vs. climate change), the context, framing, or sample population, but also be due to sample fluctuations between study groups. Similar investigations to Doyle et al. [21] show a range of variations in translation values. For example, Budescu and Wallestein [113] identified that ‘likely’ is translated to 74% (st. dev: 15%), Reagan et al. [88] to the range of 65–85%, and Budescu et al. [9] that the central 90% of translations of ‘likely’ is in the range 58–75%. These differences in potential translations further highlight the need to adopt a standardised translation scheme, and to report numerical and verbal terms together [9]. Interestingly, the World Meteorological Office has developed a table for their use in weather forecasts based upon that of the IPCC [33], but with additional categories either side of the term *medium likelihood*. This is consistent with the findings of Doyle et al. [21] which illustrates a wide gap in translation values between the terms *medium likelihood* and *unlikely* or *likely*, and the potential need for an additional translation term between those categories.

The scheme that is adopted must still reflect, as close as possible, the intuitive translation values that people have for these terms and thus ideally should be developed from the target community's perspective (non-volcanologists), such that it more closely matches their intuitive use [20], helping to facilitate effective communication in high pressure, short time situations. It is also important to consider how these tables could be developed for different languages, and this should be investigated further in the context of work such as Karelitz and Budescu [48], who have explored the different translations of verbal probability phrases for six native languages, including English, Turkish, Spanish, German, Russian, and French; and who highlight that “communication of uncertainty between languages introduces another source of error that enhances miscommunication” (p. 35).

5.2. *The relationship between probabilities and time windows*

Our results indicate that many participants have a poor understanding of the relationship between the probability value and time window stated in a forecast. Doyle et al. [21,20] found that participants skew their perceptions of event likelihoods towards the end of stated time windows.

This is similar to the findings of Fischhoff [30] who identified that individuals may incorrectly interpret a forecast due to ambiguity regarding the event being predicted and a failure to understand the reference class (in this case the time window). From a comparison of evacuation choices in experiment 1 and 2, we find here that many participants are also not correctly utilizing all the information provided (probability value and time duration). Instead they appear to be basing their decision on the probability value alone. The inability to correctly combine available information within a forecast or probability statement is similar to findings by Doyle [18] who identified that individuals incorrectly judge long-term probabilities due to annual probabilities, and that individuals' assessment of compound probabilities are substantially in error compared to normative statistical rules, due to the effect of intuitive heuristics and biases; and to Wagenaar et al. [109] who demonstrated that misunderstandings can occur due to the translation from the surface structure of our decision problem (the story) to the deeper structure of the decision problem itself.

We found in Section 4 that, considering all participants, significantly fewer participants chose an evacuation for the higher likelihood scenario described in the **numerical** statement of experiment 2 (45–55%, 3 days) than for the lower likelihood scenario described in experiment 1 (73–83%, 2 weeks). This is thought to be due to them focusing on the lower probability number (45–55%) in experiment 2's statement, and a failure to recognize the influence of the shorter time window, which actually results in a much higher likelihood scenario.

The incorrect perception that experiment 2 represents a lower likelihood event is also reflected in the comments provided, where participants raise more issues relating to the management of uncertainty and there are fewer comments pertaining to the urgency of the situation or evacuation processes and technicalities. An incorrect understanding of the relationships between probabilities and time durations in a forecast could result in delays in evacuation decisions, which could be costly in terms of life safety, expense, and reputation. When we consider the career sector for these participants, we find that there is a significant difference between the number of **scientists** that chose to evacuate for the lower likelihood scenario (experiment 1, where 95.1% evacuate) than for the higher likelihood scenario (experiment 2, where 69.7% evacuate), considering the numerical statements in each case ($\chi^2(1, N=74)=8.699$, $p=0.003$, moderate effect size $V=0.343$). However, while more **non-scientists** also evacuate for experiment 1 (84.4%) compared to experiment 2 (75%), the difference in their behaviour is not statistically significant; and so it appears that the significant difference in evacuation behaviour seen for all participants between the two experiments is mainly due to the interpretations and behaviour of the **scientists**, who are focusing more on the value of the probability number than the time window.

The non-scientists thus appear to more accurately account for the influence of the time frame, and do not have a significant majority evacuating for what is actually the lower likelihood scenario (experiment 1). This may be because the experience and background of the participants considered here concentrates their thinking on the actions required, and thus the timing, scheduling and logistics required for that decision (discussed further in the following section), and thus they consider the stated time frame to be of equal priority in their decision process as the stated probability value. This is further supported by the fact that a number of non-scientists seemed to work backwards from the end of the time window to calculate the time an evacuation should occur, accounting for the time it would take to instigate any evacuation order. Their focus on time scheduling and sequencing in their decision does however have the potential to introduce more uncertainty into the decision process, as they assess how long they think these processes will occur and related consequences or impacts.

We suggest that how individuals combine numerical probability values and time frames stated within forecasts, and how that influences their perception of the likelihood of that forecast, should be investigated further. For example, would there still be the apparent greater focus on the probability number over the time window, if the order of time and probability were swapped in the forecast (such that it said “within the next 2 weeks, the probability is 75–83%” instead of “the probability is 73–83% within the next 2 weeks”)?

5.3. *Career sector, decision reasoning, uncertainty and trust*

Interestingly, as also discussed in Section 4.1, comparison of the percentage prevalence of themes identified from comments given by *scientists* and *non-scientists* across experiments 1 and 2 (and illustrated in Tables 6 and 8), indicate that the scientists tended to query the scientific advice and express dissatisfaction about the quality and lack of information as reasons for not evacuating (which reflects their core business of data collection), whereas the non-scientists tended more to express a desire to wait for more information (which reflects their role as recipients of information and as passive planners about what could happen) while still stating their intention to start the process of preparing for evacuation. Many more of the non-scientists chose an evacuation even with the limited information compared to the scientists, commenting on the process and technicalities and the *actions* they would take, thereby either *acknowledging* or *suppressing* the uncertainty in order to make their decisions [59]. The scientists however focused more on the *information* and the quality of that information, choosing to *reduce* the uncertainty in the source (information) before proceeding.

The different comments raised by scientists and non-scientists, as well as the significantly different evacuation behaviour and perceptions observed between these two groups, highlights the complexities that can arise when these two communities respond together to an event. Ideally, for an effective response to an event, all responding individuals and organizations should have a shared mental model about response and communication [69,5,37], and a mutual understanding of the needs, responsibilities, demands and roles of each party and their capacity to anticipate other parties' demands and decision needs [92,58,79,19]. The quality of these shared mental models can be improved through both training and effective team-based simulations [15,10,81,85,4], as well as from the analysis of past responses. While this process works well with objective information, the probabilistic data communicated from scientists to emergency decision makers is more open to interpretation [80]. In addition, it is unclear how uncertainty in the science and in the situation affects these mental models and communication pathways and the relationship between these two communities, as well as issues around trust, credibility and confidence in communications.

Throughout a volcanic crisis, it is vital that communities and stakeholders continue to trust the science providers and their message [35]. This trust can be challenged when high levels of uncertainty exist and recipients in situations of high pressure, low time, and limited resources, are totally reliant on scientific sources with whom they may rarely interact under normal circumstances. As stated by Siegrist et al. [95, p. 146], "One way people cope with […] lack of knowledge is to rely on social trust to reduce the complexity they are faced with". For the science providers, it is reassuring then to know that some researchers have shown that the communication of uncertainty can enhance the credibility and trustworthiness of the information provider (see reviews in [43,42,99,68,110]), making the provider seem more 'honest' [41], which should strengthen the relationships between key decision makers and scientists. Peterson and Tilling [84, p. 348] stated in their review of responses to volcanic eruptions that "people may have unrealistic expectations of scientists' forecasting ability, and if actual events do not coincide with cautiously expressed expectations of scientists, or if scientists decline to provide specific forecasts, people may question the ability, credibility, and motives of the scientists". Perhaps if the communications had included the range of uncertainties inherent in the forecasts, this could have been mitigated, and actually enhanced the relationships, understanding and trust.

However, Kuhn [54] cautions that if we do communicate a range of uncertainty estimates, decision makers may focus more on either side of a range of uncertainty estimates around an "anchor probability", resulting in the risk posed being attenuated or amplified (p. 43). This selective interpretation of uncertainty can be due to an alignment of pre-existing attitudes and beliefs (Lord et al. as cited in [54, p. 43]), and can be used as a reason to discount the seriousness of any threat or justify a particular political agenda. In addition, other studies have suggested that communicating uncertainty can decrease people's trust in, and credibility of, the provider [43,42,99,68,110], and that it can allow people to justify inaction or their own agenda, or to perceive the risk as being higher or lower than it is, depending on their attitudes.

What is clear is that the role of trust and mutual understanding between these two communities is particularly important for communications in the emergency management context where officials make decisions based upon scientific experts and advisors in situations characterized by high pressure, low time and limited resources. As stated by Miles and Frewer [68, p. 281], "if scientific uncertainty is to be increasingly communicated to the public, the effects on trust in both risk regulators and communicators must be empirically examined". This is particularly important for future response capability and long term responses, as any loss of trust can persist well into the future and affect the quality of relationships over long periods of time and for future events. This issue identifies a need to consider the long term implications of direct scientist-non-scientist relationships in the context of infrequent events.

6. Conclusions

The issuing of forecasts and warnings of natural hazard events, such as volcanic eruptions, earthquake aftershock sequences and extreme weather often involves the use of probabilistic terms, particularly when communicated by scientific advisory groups to key decision makers who differ in expertise and their function in the decision making process. We find here that it is not just whether uncertainty is communicated or not that can affect the decision making, but also the way we understand the actual terms we use to communicate uncertainties within our advice. If one adopts probabilities as a method to encompass some of these uncertainties, providers must be aware that there are differences in the perception of what they mean, how people make choices based on them, and the understanding of competing time frames as illustrated in Fig. 5. The severity of an outcome can influence the perception of a probability [8,82], and it is not unreasonable to assume that this 'severity' can range from physical through to political or economic impacts. Thus different audiences of scientific information may be influenced by the 'political severity' of decisions to differing degrees due to the relative importance of political or economic influence that they may place on the interpretation. This further affects their interpretations and perceptions of probabilities and event likelihoods. It is thus important to understand how scientists and other authorities impose meaning on, interpret and understand the distribution of the outcome likelihood within these time frames. This meaning is developed within different professional frames of reference and groups with different goals, and any meaning derived from the scientists' mental model may be inconsistent with a practitioner's mental model and goals (e.g., what may happen vs. do we need to evacuate or not; [80])

Our results show that people have a poor understanding of time windows stated with forecasts in particular, skewing their perceptions to the end of a time window and incorrectly understanding the relationship between time stated and probability, resulting in an under or overestimate of forecast likelihoods. In addition, they underestimate the scenario likelihood when verbal terms are used. These differences in perception significantly affect decisions, and if emergency managers receiving the scientific advice are affected by different perceptions of the wording of the probability itself, they may enact different actions based upon the same probabilistic advice which has been formatted in different ways (verbal vs. numerical, using the terms 'in' or 'within', focusing on short or long time windows, [20]; and see also [46]).

For most natural hazard crises, the successful management of a response is dependent upon not just the communication between scientists and emergency managers, but also the onward communication of this information to the communities through public education programs and warnings [102,51,28,12,13,2,55,101,67]. Thus, irrespective of the objective quality of the information from the scientific community, its ability to have the desired effect is influenced by how it is interpreted and filtered as it is transmitted through these various recipients. It is thus vital that any communication between scientific agencies and emergency managers maintains its true meaning in that communication prior to the next step, such that it is accurately communicated onward to the public through bulletins, warnings and advice. It is thus important to identify communication strategies that may help to mitigate these different perceptions and to help prevent onward miscommunications, or premature, delayed, or unnecessary action choices.

Based on our findings here and those of Doyle et al. [21,20], we suggest that when scientists have calculated or estimated likelihood forecasts involving time windows (which is itself a challenging task due to the fluctuating levels of unrest; [100,56,90,105]), that they make all possible efforts to communicate such forecasts, likelihoods, and probabilities over a range of relevant time windows, including a probability forecast for a shorter immediate time window in particular (such as the first 24 h). This will help to anchor this probabilistic information in the present, and prevent any underestimate of the probability of an event ‘today’, while also reducing the misunderstanding of the interaction between the stated probability and time window in a forecast. This is vital, as these short time frames are often when major decisions such as evacuations are made, and any interpretive bias due to a skew in time perception may lead key decision makers to overestimate any time available to enact their action choice.

Our results herein have considered a simple forecast evacuation decision scenario (“at least 48 h to evacuate”). However, as discussed by Marrero et al. [63], there can be a considerable time-lapse between the evacuation decision-making time and time in which an evacuation is complete, depending upon the socio-economic status of the resident population, the transportation network, evacuation plan, number of evacuees, compliance with evacuation orders, environmental schemes as well as many other complicating factors. Thus, these complexities in evacuation and the need for flexible evacuation planning further highlights the need to ensure that any potential interpretive bias due a skew in forecast time perception is mitigated for with the use of multiple time windows. As discussed earlier, it is important that scientific advice is timely, accurate and meets the needs of the diverse decision makers, emergency managers and public involved in a crisis. Thus, through pre-crisis exercises, scenarios and discussions, the ideal forecast time windows should be identified through an assessment of stakeholder decision making and action planning thresholds. As discussed by Morss et al. (in the context of meteorological forecasts; [70]), the effectiveness of information communication should be evaluated from different perspectives, it is more important that end-users understand the forecast well enough to inform their decisions, and less important that they understand the precise scientific details.

Budescu et al. [9], also suggest that a forecast or probability statement should state both the numerical and verbal equivalents of the probability, such that the message includes the more ‘accessible’ verbal term as well as a numerical term that is less subject to bias, and that a standardised translation scheme should be used to move from one to the other. From our results herein we further advocate this approach, as verbal probability terms alone seem to be viewed as representing a less certain, less confident, or lower likelihood scenario than their numerical equivalents; resulting in less decisive action. However, when incorporating these recommendations, and those outlined in Doyle et al. [20], care must be taken to ensure the message does not become too complex, as providing as much advice as possible may hinder the decision process, due to cognitive overload [87,74]. Finally, we highlight that these translation tables should not be mistaken for decision-making or warning tools, but rather translation guides to facilitate the move between verbal phrases and their equivalent numerical terms. Any formalised communication strategy should be accompanied by exercises, simulations, and education programs with both the decision-makers and the public to help facilitate a greater understanding of the complexities inherent in these uncertain forecasts.

Acknowledgements

Dr Emma E H Doyle is supported by a Foundation for Research Science & Technology NZ S&T Postdoctoral Fellowship MAUX0910 All research conducted within this study was done in accordance with Massey University's Code of Ethical Conduct for Research, Teaching and Evaluations Involving Human Participants. We thank José Marrero and two other anonymous reviewers for comments that enhanced this manuscript.

References

[1] Van Asselt MBA. Perspectives on uncertainty and risk: the PRIMA approach to decision support. Dordrecht, The Netherlands: Kluwer Academic Publishers; 2000.
[2] Barclay J, Haynes K, Mitchell T, Solana C, Teeuw R, Darnell A, et al. Framing volcanic risk communication within disaster risk reduction: finding ways for the social and physical sciences to work together, 305. London: Geological Society; 163–77 (Special Publications).
[3] Bordens KS, Abbott BB. Research design and methods: a process approach. 4th ed.Mountain View, California: Mayfield Publishing Company; 1999.
[4] Borodzicz E, van Haperen K. Individual and group learning in crisis simulations. J Conting Crisis Manag 2002;10(3):139–47.
[5] Bostrom A, French S, Gottlieb S. Risk assessment, modeling, and decision support. In: Mumpower Jeryl, Renn Ortwin, editors. Risk, Governance and Society, vol. 14. Berlin: Springer Verlag; 2008 (341 p.).
[6] Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006;3(2):77–101.
[7] Braun V, Clarke V. Successful qualitative research: a practical guide for beginners.London: Sage Publications Ltd.; 2013 (400 p.).
[8] Bruine De Bruin W, Fischhoff B, Millstein SG, Halpern-Felsher BL, et al. Verbal and numerical expressions of probability: it's a fifty-fifty chance. Organ Behav Hum Decis Process 2000;81(1):115–31.

[9] Budescu DV, Broomell S, Por H-H. Improving communication of uncertainty in the reports of the Intergovernmental Panel on Climate Change. Psychol Sci 2009;20(3):299–308.
[10] Cannon-Bowers JA, Bell HE. Training decision makers for complex environments: implications of the naturalistic decision making perspective. In: Zsambok CE, Klein G, editors. Naturalistic Decision Making. Mahwah, NJ: Lawrence Erlbaum Associates; 1997. p. 99–110.
[11] Chaiken S, Trope Y. Dual process theories in social psychology.New York: Guilford Press; 1999.
[12] Chester DK. Volcanoes and Society.London: Edward Arnold; 1993.
[13] Chester DK, Dibben CJL, Duncan AM. Volcanic hazard assessment in Western Europe. J Volcanol Geotherm Res 2002;115:411–35.
[14] CPVAG, 2009. Central Plateau Volcanic Advisory Group Strategy, report No 2010/EXT/1117, Morris B, editor. Horizons Regional Council, Palmerston North, NZ. ISBN: 978-1-877556-42-5. Accessed from: 〈http://www.horizons.govt.nz/assets/publications/keeping-people-safe-publications/Central-Plateau-Volcanic-Advisory-Group-Volcanic-Strategy.pdf〉.
[15] Crego J, Spinks T. Critical incident management simulation. In: Flin R, Salas E, Strub M, Martin L, et al., editors. Decision making under stress: emerging themes and applications. Aldershot, England: Ashgate Publishing Limited; 1997. p. 85–94.
[16] Dalla RL. Night moves: a qualitative investigation of street-level sex work. Psychol Women Q 2002;26:63–73.
[17] Dhami MK, Wallsten TS. Interpersonal comparison of subjective probabilities: toward translating linguistic probabilities. Mem Cogn 2005;33(6):1057–68.
[18] Doyle JK. Judging cumulative risk. J Appl Soc Psychol 1997;27(6): 500–24.
[19] Doyle EE, Johnston DM. Science advice for critical decision-making. In: Paton D, Violanti J, editors. Working in high risk environments: developing sustained resilience. Springfield, Illinois, USA: Charles C Thomas; 2011. p. 69–92.
[20] Doyle EEH, McClure J, Johnston DM, Paton D, et al. Communicating likelihoods and probabilities in forecasts of volcanic Eruptions. J Volcanol Geotherm Res 2014;272:1–15.
[21] Doyle EEH, Johnston DM, McClure J, Paton D. The communication of uncertain scientific advice during natural hazard events. N Z J Psychol 2011;40(4):39–50.
[22] Eiser JR, Bostrom A, Burton I, Johnston DM, McClure J, Paton D, et al. Risk interpretation and action: a conceptual framework for responses to natural hazards. Int J Disaster Risk Reduct 2012;1: 5–16.
[23] Elsworth D, Mattioli G, Taron J, Voight B, Herd R, et al. Implications of magma transfer between multiple reservoirs on eruption cycling. Science 2008;322(5899):246–8.
[24] Epstein S. Integration of the cognitive and the psychodynamic unconscious. Am Psychol 1994;49(8):709–24.
[25] Evans J St BT, Stanovich KE. Dual process theories of cognition: advancing the debate. Perspect Psychol Sci 2013;8:223–41.
[26] Evans St. JSBT, Stanovich KE. Theory and metatheory in the study of dual processing: reply to comments. Perspect Psychol Sci 2013;8: 263–71.
[27] Farley JE, Barlow HD, Finkelstein MS, Riley L, et al. Earthquake hysteria, before and after: a survey and follow-up on public response to the browning forecast. Int J Mass Emerg Disasters 1993;11(3):305–21.
[28] Fiske, RS, Volcanologists, journalists, and the concerned local public: a tale of two crises in the eastern Caribbean. In: Studies in Geophysics: Explosive volcanism: inception, evolution and hazards, 1984, National Academy Press; Washington, 170–176.
[29] Folke C, Colding J, Berkes F. Building resilience and adaptive capacity in social-ecological systems. In: Berkes F, Colding J, Folke C, editors. Navigating social-ecological systems: building resilience for complexity an change. Cambridge: Cambridge University Press; 2003.
[30] Fischhoff B. What forecasts (seem to) mean. Int J Forecast 1994;10: 387–403.
[31] Garcia C, Fearnley CJ. Evaluating critical links in early warning systems for natural hazards. Env Hazards 2012;11(2):123–37.
[32] Gigerenzer G, Edwards A. Simple tools for understanding risks: from innumeracy to insight. Br Med J 2003;327(7417):741.
[33] Gill J, Rubiera H, Martin C, Cacic I, Mylne K, Dehui C, et al. World Meteorological Organization guidelines on communicating forecast uncertainty. Organization 2008.
[34] Green CH, Tunstall SM, Fordham MH. The risk from flooding: which risks and whose perception. Disasters 1992;15:227–36.
[35] Haynes K, Barclay J, Pidgeon N. The issue of trust and its influence on risk communication during a volcanic crisis. Bull Volcanol 2007;70(5):605–21.
[36] Haynes K, Barclay J, Pidgeon N. Whose reality counts? Factors affecting the perception of volcanic risk J Volcanol Geotherm Res 2008;172:259–72.
[37] Heath R, Lee J, Ni L. Crisis and risk approaches to emergency management planning and communication: the role of similarity and sensitivity. J Public Relat Res 2009;21(2):123–41.
[38] Honda H, Yamagishi K. Directional verbal probabilities. Exp Psychol (Z Exp Psychol) 2006;53(3):161–70.
[39] Houghton JT, Ding Y, Griggs DJ, Noguer M, van der Linden PJ, Dai X, et al. Climate change 2001: the scientific basis.Cambridge, U.K.: Cambridge Universtiy Press; 2002.
[40] IAVCEI, Subcomittee for Crisis Protocols. Professional conduct of scientists during volcanic crises. Bull Volcanol 1999;60:323–34.
[41] Johnson BB. Further notes on public response to uncertainty in risks and science. Risk Anal: Off Publ Soc Risk Anal 2003;23(4):781–9.
[42] Johnson BB, Slovic P. Lay views on uncertainty in environmental health risk assessment. J Risk Res 1998;1(4):261–79.
[43] Johnson BB, Slovic P. Presenting uncertainty in health risk assessment: initial studies of its effects on risk perception and trust. Risk Anal 1995;15(4):485–94.
[44] Jordan TH, Chen Y-T, Gasparini P, Madariaga R, Main I, Marzocchi W, et al. Operational earthuqake forecasting: state of knowledge and guidelines for utilization. A report by the International Commission on Earthquake Forecasting for Civil Protection. Ann Geophys 2011;54(4). http://dx.doi.org/10.4401/ag-5350.
[45] Jordan TH. Opinion: lessons of L'Aquila for operational earthquake forecasting. Seismol Res Lett 2013;84(1):4–7.
[46] Joslyn S, Pak K, Jones D, Pyles J, Hunt E, et al. The effect of probabilistic information on threshold forecasts. Weather Forecast 2007;22(4):804.
[47] Joslyn SL, Nadav-Greenberg L, Taing MU, Nichols RM, et al. The effects of wording on the understanding and use of uncertainty information in a threshold forecasting decision. Appl Cognit Psychol 2009;23(1):55–72.
[48] Karelitz TM, Budescu DV. You say probable and I say likely: improving interpersonal communication with verbal probability phrases. J Exp Psychol: Appl 2004;10(1):25–41.
[49] Keren G, Schul Y. Two is not always better than one: a critical evaluation of two-system theories. Perspect Psychol Sci 2009;4: 533–50.
[50] Keren G. A tale of two systems: a scientific advance or a theoretical stone soup? Commentary on Evans & Stanovich (2013) Perspect Psychol Sci 2013;8:257–62.
[51] Kilburn CRJ. Volcanoes and the fate of forecasting. New Sci 1978;80:511–3.
[52] Kim JL. Asian American women's retrospective reports of their sexual socialisation. Psychol Women Q 2009;33:334–50.
[53] Kruglanski AW, Gigerenzer G. Intuitive and deliberative judgements are based on common principles. Psychol Rev 2011;118: 97–109.
[54] Kuhn K. Message format and audience values: interactive effects of uncertainty information and environmental attitudes on perceived risk. J Environ Psychol 2000;20(1):41–51.
[55] Leonard GS, Johnston DM, Paton D, Christianson A, Becker J, Keys H, et al. Developing effective warning systems: ongoing research at Ruapehu volcano, New Zealand. J Volcanol Geotherm Res 2008;172 (3–4):199–215.
[56] Lindsay J, Marzocchi W, Jolly G, Constantinescu R, Selva J, Sandri L. Towards real-time eruption forecasting in the Auckland Volcanic Field: application of BET_EF during the New Zealand National Disaster Exercise "Ruaumoko". Bull Volcanol 2009;72(2):185–204.
[57] Lipkus IM. Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations. Med Decis Mak: Int J Soc Med Decis Mak 2010;27(5):696–713.
[58] Lipshitz R, Klein G, Orasanu J, Salas E, et al. Focus article: taking stock of naturalistic decision making. J Behav Decis Mak 2001;14 (5):331–52.
[59] Lipshitz R, Strauss O. Coping with uncertainty: a naturalistic decision-making analysis. Organ Behavior and Hum Decis Process 1997;69(2):149–63.
[60] Loewenstein GF, Weber EU, Hsee CK, Welch N, et al. Risk as feelings. Psychol Bull 2001;127(2):267–86.
[61] Mackie B. Warning fatigue – myth or misunderstanding: insights from the Australian bushfires. Can Risk Hazards Netw 2013;5(1):51–5.

[62] Mader CG, Blair ML. Living with a volcanic threat: response to volcanic hazards, Long Valley, Caldera.Portola Valley, California: William Spangle and Associates; 1987.
[63] Marrero JM, Garcia A, Llinares A, De la Cruz-Renya S, Ramos S, Ortiz R, et al. Virtual tools for volcanic crisis management and evacuation decision support: applications to El Chichón volcano (Chiapas, México). Nat Hazards 2013;68:955–80.
[64] Mastrandrea MD, Field CB, Stocker TF, Edenhofer O, Ebi KL, Frame DJ, et al., 2010. Guidance note for lead authors of the IPCC fifth assessment report on consistent treatment of uncertainties. IPCC cross-working group meeting on consistent treatment of uncertainties. Jasper Ridge, CA, USA; 6–7 July 2010. Accessed from: ⟨http://www.ipcc.ch/pdf/supporting-material/uncertainty-guidance-note.pdf⟩.
[65] MCDEM, 2008. Exercise Ruaumoko '08 Final Exercise Report. Wellington, NZ.
[66] McKee CO, Johnson RW, Lowenstein PL, Riley SJ, Blong RJ, De Saint Ours P, et al. Rabaul Caldera, Papua New Guinea: volcanic hazards, surveillance, and eruption contingency planning. J Volcanol Geotherm Res 1985;23(3–4):195–237.
[67] McGuire WJ, Solana MC, Kilburn CRJ, Sanderson D. Improving communications during volcanic crises on small, vulnerable islands. J Volcanol Geotherm Res 2009;183:63–75.
[68] Miles S, Frewer LJ. Public perception of scientific uncertainty in relation to food hazards. J Risk Res 2003;6(3):267–83.
[69] Morgan MG, Fischoff B, Bostrom A, Atman CJ, et al. Risk communication: a mental models approach.London: Cambridge University Press; 2001.
[70] Morss RE, Demuth JL, Lazo JK. Communicating uncertainty in weather forecasts: a survey of the U.S. public. Weather Forecast 2008;23:974–91.
[71] Moss RH, Schneider SH, 2000. Uncertainties in the IPCC TAR: recommendations to lead authors for more consistent assessment and reporting. In: Pachauri R, Taniguchi T, Tanaka K, editors. IPCC Supporting Material, Guidance Papers on the Cross Cutting Issues of the Third Assessment Report of the IPCC. p. 33–51.
[72] Newhall CG, Punongbayan RS. The narrow margin of successful volcanic-risk mitigation. In: Scarpa R, Tilling RI, editors. Monitoring and mitigation of volcanic hazards. Berlin Heidelberg: Spinger Verlag; 1996. p. 807–38.
[73] NRC. Completing the forecast: characterizing and communicating uncertainty for better decisions using weather and climate forecasts.Washington D. C.: National Academies Press; 2006 (124 p.).
[74] Omodei MM, McLennan J, Elliott GC, Wearing AJ, Clancy JM, et al. "More is better?": a bias toward overuse of resources in Naturalistic decision making settings In: Montgomery H, Lipshitz R, Brehmer B, editors. How professionals make decisions. Mahwah, NJ: Lawrence Erlbaum Associates; 2005. p. 29–41.
[75] Osman M. A case study: dual-process theories of higher cognition—commentary on Evans & Stanovich (2013). Perspect Psychol Sci 2013;8:248–52.
[76] Paton D, Johnston DM, Houghton B, Smith LM, et al. Managing the effects of a volcanic eruption: psychological perspectives on integrated emergency management. J Am Soc Prof Emerg Plan 1998;5:59–69.
[77] Paton D, Johnston D, Houghton B, Flin R, Ronan K, Scott B. Managing natural hazard consequences: information management and decision making. J Am Soc Prof Emerg Managers 1999;6:37–48.
[78] Paton D, Auld T. Resilience in emergency management: managing the flood. In: Paton D, Johnston D, editors. Disaster resilience: an intergrated approach. Springfield, Illinois, USA: Charles C Thomas; 2006.
[79] Paton D, Jackson D. Developing disaster management capability: an assessment centre approach. Disaster Prev Manag 2002;11(2):115–22.
[80] Paton D, McClure J. Preparing for disaster: building household and community capacity. Illinois, USA: Charles C Thomas Pub Ltd. Springfield; 2013 (245 p.).
[81] Paton D, Smith L, Violanti J. Disaster response: risk, vulnerability and resilience. Disaster Prev Manag 2000;9(3):173–80.
[82] Patt A, Dessai S. Communicating uncertainty: lessons learned and suggestions for climate change assessment. Comptes Rendus Geosci 2005;337(4):425–41.
[83] Patt A, Schrag DP. Using specific language to describe risk and probability. Clim Change 2003;61:17–30.
[84] Peterson DW, Tilling RI. Interactions between scientists, civil authorities and the public at hazardous volcanoes. In: Kilburn CRJ, Luongo G, editors. Monitoring active lavas. London: UCL Press; 1993. p. 339–65.
[85] Pliske RM, McCloskey MJ, Klein G. Decision skills training: facilitating learning from experience. In: Salas E, Klein G, editors. Linking expertise and naturalistic decision making. Mahwah, NJ: Lawrence Erlbaum Associates; 2001. p. 37–53.
[86] Punongbayan RS, Newhall CG, Bautista MLP, Garcia D, Harlow DH, Hoblitt RP, et al. Eruption hazard assessments and warnings. In: Newhall CG, Punongbayan RS, editors. FIRE and MUD: eruptions and Lahars of Mount Pinatubo, Philippines. Seattle, London: University of Washington Press; 1996. p. 67–85.
[87] Quarantelli EL. Ten criteria for evaluating the management of community disasters. Disasters 1997;21(1):39–56.
[88] Reagan RT, Mosteller F, Youtz C. Quantitative meanings of verbal probability expressions. J Appl Psychol 1989;74:433–42.
[89] Risbey JS, Kandlikar M. Expressions of likelihood and confidence in the IPCC uncertainty assessment process. Clim Change 2007;85(1–2):19–31.
[90] Rolandi G. Volcanic hazard at Vesuvius: an analysis for the revision of the current emergency plan. J Volcanol Geotherm Res 2010;189:347–62, http://dx.doi.org/10.1016/j.jvolgeores.2009.08.007.
[91] Ryan GW, Bernard HR. Data management and analysis methods. In: Denzin NK, Lincoln YS, editors. Handbook of Qualitative Research. Thousand Oaks, London; New Delhi: SAGE Publications, Ltd; 2000. p. 769–802.
[92] Salas E, Stout RJ, Cannon-Bowers JA. The role of shared mental models in developing shared situational awareness. In: Gilson RD, Garland DJ, Koonce JM, editors. Situational awareness in complex systems: proceedings of a cahfa conference. Daytona Beach, FL: Embry-Riddle Aeronautical University Press; 1994. p. 298–304.
[93] Sharma U, Patt A. Disaster warning response: the effects of different types of personal experience. Nat Hazards 2011;60(2):409–23.
[94] Showalter PS. Prognostication of doom: an earthquake prediction's effect on four small communities. Int J Mass Emerg Disasters 1993;11(3):279–92.
[95] Siegrist M, Gutscher H, Earle T. Perception of risk: the influence of general trust, and general confidence. J Risk Res 2005;8(2):145–56.
[96] Sloman SA. The empirical case for two systems of reasoning. Psychol Bull 1996;119(1):3–22.
[97] Slovic P, Finucane ML, Peters E, MacGregor DG, et al. Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. Risk Anal: Off Publ Soc Risk Anal 2004;24(2):311–22.
[98] Smith R. Research, science and emergency management: partnering for resilience. In: Tephra, editor. Community resilience: research, planning and civil defence emergency management. Wellington, New Zealand: Ministry of Civil Defence & Emergency Management; 2009. p. 71–8.
[99] Smithson M. Conflict aversion: preference for ambiguity vs conflict in sources and evidence. Organ Behav Hum Decis Process 1999;79(3):179–98.
[100] Sparks RSJ. Forecasting volcanic eruptions. Earth Planet Sci Lett 2003;210(1–2):1–15.
[101] Solana MC, Kilburn CRJ, Rolandi G. Communicating eruption and hazard forecasts on Vesuvius, Southern Italy. J Volcanol Geotherm Res 2008;172:308–14.
[102] Tazieff H. La Soufrière, volcanology and forecasting. Nature 1977;269:96–7.
[103] Teigen KH, Brun W. The directionality of verbal probability expressions: effects on decisions, predictions, and probabilistic reasoning. Organ Behav Hum Decis Process 1999;80(2):155–90.
[104] Thompson VA. Why it matters: the implications of autonomous processes for dual-process theories – commentary on Evans & Stanovich (2013). Perspect Psychol Sci 2013;8:253–6.
[105] Tilling RI. Extreme environmental events. In: Tilling RI, Meyers RA, editors. Complexity in forecasting and early warning volcanic hazards and early warning, 1. Springer-Verlag; 2011. p. 1135–45 (978-1-4419-7695-6).
[106] Vogt WP. Dictionary of statistics & methodology. 3rd ed.California, USA: SAGE Publications, Ltd. Thousand Oaks; 2005 (456 p.).
[107] Wallsten TS, Fillenbaum S, Cox JA. Base rate effects on the interpretations frequency expressions of probability. J Mem Lang 1986;25:571–87.
[108] Weber EU. Experience-based and description-based perceptions of long-term risk: why global warming does not scare us (yet). Clim Change 2006;77(1–2):103–20.
[109] Wagenaar WA, Keren G, Lichtenstein S. Islanders and hostages: deep and surface structures of decision problems. Acta Psychol 1998;67:175–89.

[110] Wiedemann P, Borner F, Schultz H. Lessons learned: recommendations for communicating conflicting evidence for risk characterization. In: Wiedemann PM, Schultz H, editors. The role of evidence in risk characterisation: making sense of conflicting data. Weinheim: Wilery-VCH Verlag GmbH & Co. KGaA; 2008. p. 205–13.
[111] Woo G. Probabilistic criteria for volcano evacuation decisions. Nat Hazards 2008;45(1):87–97.
[112] Fearnley CJ. Assigning a volcano alert level: negotiating uncertainty, risk, and complexity in decision-making processes. Environment and Planning A 2013;45:1891–911, http://dx.doi.org/10.1068/a4542.
[113] Budescu DV, Wallsten TS. Consistency in interpretation of probabilistic phrases. Organizational behavior and human decision processes 1985;36:391–405.