Maintaining Credibility When Communicating Uncertainty: The Role of Communication Format

Sarah C. Jenkins (s.jenkins.12@ucl.ac.uk) and Adam J. L. Harris (adam.harris@ucl.ac.uk)
Department of Experimental Psychology, University College London,
26 Bedford Way, London, WC1H 0AP, UK

R.M. Lark (mlark@bgs.ac.uk)
British Geological Survey, Keyworth, Nottingham, NG12 5GG, UK

Abstract

Research into risk communication has commonly highlighted the disparity between the meaning intended by the communicator and what is understood by the recipient. Such miscommunications will have implications for perceived trust and expertise of the communicator, but it is not known whether this differs according to the communication format. We examined the effect of using verbal, numerical and mixed communication formats on perceptions of credibility and correctness, as well as whether they influenced a decision to evacuate, both before and after an 'erroneous' prediction (i.e. an 'unlikely' event occurs, or a 'likely' event does not occur). We observed no effect of communication format on any of the measures pre-outcome, but found the numerical format was perceived as less incorrect, as well as more credible than the other formats after an 'erroneous' prediction, but only when low probability expressions were used. Our findings suggest numbers should be used in consequential risk communications.

Keywords: verbal probability expressions; numerical probabilities; risk communication; trust; expertise; credibility

Introduction

Science is suffering from a 'crisis of trust' (House of Lords, 2000); preserving and cultivating the public's trust has never been more important for the scientific community (Nature, 2010). Uncertainty is an inescapable part of any scientific endeavour, but the presence of it creates doubt in the minds of the public and it is often used as a reason to delay taking action (Lewandowsky, Ballard, & Pancost, 2015). Effectively communicating information regarding risk and uncertainty thus represents a significant problem for scientists.

Methods for communicating risk and uncertainty include using verbal probability expressions (VPEs; e.g. 'possible', 'likely'), numerical expressions (e.g. '20% likelihood'), or mixed expressions (e.g. 'unlikely [20% likelihood]'). Budescu and Wallsten (1995) proposed that the choice of format for communicating likelihood information should be governed by the congruence principle: the precision of the communication should be consistent with the degree of certainty that can reasonably be expected for estimates about the event described. Much research has investigated the pitfalls of using VPEs to communicate uncertainty using the 'how likely' translation approach, whereby people are asked to translate a VPE to a corresponding numerical probability This has highlighted the variability in people's usage and interpretations (e.g., Budescu & Wallsten, 1985), as well as the influence of other contextual and cultural factors (e.g., Bonnefon & Villejoubert, 2006; Harris & Corner, 2011; Harris, Corner, Xu, & Du, 2013; Teigen & Brun, 1999, 2003; Weber & Hilton, 1990). Such variability clearly highlights the potential for a reduction in perceived credibility of the communicator, if there is a disparity between the meaning intended by the communicator and that which is understood by the recipient.

A commonly suggested solution to the problems of miscommunication is to use a dual-scale, mixed format expression to communicate risk and uncertainty, for example 'It is unlikely (less than 33%)' (e.g., Budescu, Broomell, & Por, 2009; Budescu, Por, Broomell, & Smithson, 2014; Harris & Corner, 2011; Harris et al., 2013; Patt & Dessai, 2005; Witteman & Renooij, 2003). Using such a 'verbal-numerical' (V-N) format was found to increase correspondence between people's interpretations and the IPCC guidelines, an effect that replicated across 24 countries (Budescu et al., 2014). However, when shown a histogram of potential outcomes and asked to complete probability statements (e.g., "It is *unlikely* that the lava flow will extend to a distance of __km"), the so-called 'which outcome' approach to studying VPEs (e.g., Teigen, Juanchich, & Riege, 2013), participants tended to complete the sentence with a distance that exceeded any represented in the histogram, both for 'unlikely' *and* 'unlikely (20% chance)' (Jenkins, Harris, & Lark, 2016; see also Juanchich & Sirota, 2016). If such phrases are seen as appropriate for communicating an outcome with a 0% chance of occurring, the mismatch between this and an intended communication of '20% likelihood' could adversely affect confidence in subsequent communications.

Aside from the terms used, a further problem arises from people's general understanding of uncertainty and probability. Uncertainty is often perceived by the public as an 'indicator of ignorance', when in fact it should be seen as a source of actionable knowledge (Lewandowsky et al., 2015). Scientific forecasts are probabilistic (at best) and thus it is, for example, not possible to predict with certainty the probability of a volcanic eruption on a given day. Even if an event is predicted to be 'likely' to occur, the very fact it is not certain means that it might still not happen. In the same vein, even if an event is predicted as 'unlikely' to occur (e.g. 20% likelihood; Theil, 2002), it does not mean the event will definitely not occur, given that one in five times it will (on a frequentist interpretation of probability). The expectation of what will happen is largely driven by the directionality of the expression (Teigen & Brun, 1995, 1999); in that phrases which have negative directionality (e.g. 'unlikely') focus one's attention on the non-occurrence of the event, whereas those with positive directionality (e.g. 'likely') focus on the occurrence of the event. If the outcome is 'opposite' to what was predicted, the predictions are often seen as 'erroneous', which could have a knock-on effect on perceived credibility.

Despite recent calls to use a dual-scale communication format, research has yet to explore the effect of using mixed expressions on the perceived credibility of the communicator. Neither, perhaps more importantly, has it investigated the consequences of 'erroneous' predictions on credibility. Given a major function of risk communication is providing trustworthy information, confidence in the source of the information is key (Kasperson, 2014). After all, even if the information is understood as intended, it is of no use if the communicator is not perceived as credible and thus is not trusted enough to inspire action on the basis of the communication. Indeed, credibility has been found to influence risk perceptions. Trust is negatively associated with perceived risk (Sjöberg, 2001), as well as directly affecting behaviour (Wachinger, Renn, Begg, & Kuhlicke, 2013).

Longman, Turner, King, & McCaffery (2012) explored the effect of numerical formats on accuracy of understanding, perceived risk, and source credibility judgements for two different sources of risk information (clinician / pharmaceutical company). The risk estimate was presented either as a either a point (20 out of 100), small range (16 – 24 out of 100) or large range (8 – 32 out of 100). Range information resulted in reduced understanding and the large range was perceived as more risky compared to a point estimate. Experts using point estimates were viewed as more credible. Gurmankin, Baron and Armstrong (2004) investigated the effect of verbal and numerical statements of risk (percentage / fraction) on trust and comfort in a physician in a hypothetical medical communication. They found subjects were more trusting of, and more comfortable with, numerical versions of the information, though this effect decreased with lowering levels of numeracy, highlighting the importance of including a numeracy measure in the current study.

The importance of investigating the credibility of the communicator cannot be understated. Whilst an accurate understanding of information is clearly desirable, it is people's actions (on the basis of the communication) which matter, given they will have the most consequences for the individual. Therefore an investigation into the effects of communication format should also consider the effect of communication format on people's actions. Doyle, McClure, Paton, & Johnston (2014) found that fewer people suggested evacuating when the risk of a volcanic eruption was described using verbal terms than when using numerically equivalent terms, suggested to be a result of the fact that VPEs are viewed as more ambiguous, though again the study did not consider mixed-formats, or the influence of 'erroneous' predictions.

Although previous research has demonstrated the V-N format aids understanding in risk communications (Budescu et al., 2014), it may not be the preferred format for the recipient. Indeed, there may be a discrepancy between what people favour (for instance the preference for receiving information in numerical form, Erev & Cohen, 1990) and what experts can suitably provide. Using a numerical point estimate (e.g. 15%) to describe the chance of a natural hazard (which are, by nature, highly uncertain) might be perceived as overly precise according to the congruence principle (Budescu & Wallsten, 1995) and thus not credible.

A deeper understanding of the effects of using different communication formats and the consequences of 'erroneous' predictions is therefore clearly required, such that the public's trust in science can be built and maintained. We thus sought to examine whether initial perceptions of credibility in the communicator differed according to communication format over two studies featuring low and high probability events. We also investigated whether these perceptions changed after an 'erroneous' prediction (i.e. the 'unlikely' outcome occurred, in Study 1, or the 'likely' outcome did not occur, in Study 2). Ascertaining the effect of these factors is instructive for developing effective risk communication strategies.

Study 1

Method

Participants

300 Native English speakers (146 male) aged between 18 – 72 (*Mdn*= 33.5) were recruited from Prolific Academic (PA; www.prolific.ac). Participants received £0.75 for participating.

Design

A 4 × 2 mixed design was used. Communication format was in the low probability domain and had four levels, manipulated between participants: verbal- "unlikely", numerical- "20% likelihood", V-N- "unlikely (20% likelihood)" and N-V- "20% likelihood (unlikely)." Outcome (pre/post) was a within-participants variable.

Perceptions of trust, expertise, correctness and decision to evacuate were rated on five-point scales. Expertise was operationalised as 'How knowledgeable does the expert seem?' from 1 – 'Not at all knowledgeable' to 5 – 'Extremely knowledgeable'. Trust was operationalised as 'How much do you trust that the expert is giving you complete and unbiased information?' (Dieckmann, Slovic, & Peters, 2009), from 1 – 'Not at all' to 5 – 'A great deal'. Decision to evacuate, based on Doyle et al. (2014), was rated from 1– 'Definitely should evacuate today' to 5 – 'Definitely should not evacuate today'. Participants also then had to indicate why they made that decision. Correctness was rated from 1 – 'Not at all correct' to 5 – 'Completely correct'.

Materials and Procedure

After consenting to participate, participants indicated their age, gender and Prolific ID before reading the introductory text. The introductory text informed participants that they would see a geological scenario and be asked to make a series of judgements about this. On the next screen, participants read a vignette about a current volcanic eruption, in which lava flows were expected. A volcanologist presented a communication about the probability of the lava flows travelling a certain distance:

> "Mount Ablon has a history of explosive eruptions that have produced lava flows. An eruption is currently underway and lava flows are expected. Volcanologists from Ablon Geological Centre are communicating information about the volcano. A volcanologist has suggested that, given the volcano's recent history, there is a **20% likelihood (unlikely)** that the lava flow will extend 3.5km from the point of eruption."

Participants then provided initial ratings of expertise and trust in the expert's prediction of events. On the subsequent screen, participants were informed that the capital city was at risk of the volcanic eruption and asked to rate whether to evacuate the city today or not (Doyle et al., 2014). A mass evacuation was described as being 'very expensive and extremely disruptive to residents'.

Participants were then informed on the following screen that the unlikely outcome did in fact occur. They were asked to provide further trust and expertise ratings, as well as rating how correct the volcanologist's prediction was in light of the outcome. The next screen then showed a similar communication by a volcanologist about Mount Ablon, set two years on, with participants asked the two evacuation questions, as before.

Finally participants completed a numeracy scale (Lipkus, Samsa, & Rimer, 2001), with two additional questions from the Berlin Numeracy Test (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012) included to increase variability in scores, given previous studies using PA have found it to be a highly numerate sample. After completing the study, participants were given a code to claim their reward, thanked and debriefed.

Results

There was a significant correlation between trust and expertise ratings, both pre-outcome, r = .69, p < .001 and post-outcome, r = .74, p < .001. For ease of exposition, we averaged the measures to create a single measure of credibility. The data were entered into a 4 (communication format) × 2 (outcome) × 2 (numeracy) ANOVA, unless stated otherwise.

Given the highly skewed distribution of responses, participants with scores of eight or under were classed as low numeracy and those with nine or above classed as high numeracy. However, given there was only one effect of (or interaction involving) numeracy across Studies 1 and 2, this variable is only considered further in that single instance.

Credibility Ratings

Mean credibility ratings, by communication format, are plotted in Figure 1, which suggests that pre-outcome there was little difference between formats. All communication formats suffered from a loss of perceived credibility post-outcome, but there was less of a reduction in the numerical format. Correspondingly, there was a main effect of outcome, F (1, 292) = 218.60, p < .001, η_p^2= .43, and format, F (3, 292) = 5.77, p < .01, η_p^2= .06, but this was qualified by a significant interaction between outcome and format, F (3, 292) = 6.87, p < .001, η_p^2= .07. Simple effects analyses confirmed no effect of format pre-outcome F (3, 296) = 0.38, p =.77, and a significant effect of format post-outcome F (3, 292) = 8.02, p < .001. It is worth noting, however, that the reduction in credibility was still significant even in the numerical condition, t(73)= 3.66, p < .001, d= 0.43.

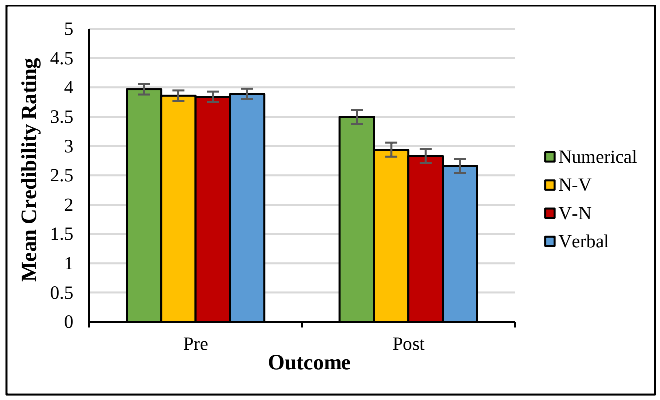

Figure 1. Effect of Communication Format on Perceptions of Credibility Before and After an 'Erroneous' Prediction (Error Bars Represent ±1SE) – Study 1 – Low Probability.

Decision to Evacuate

Mean evacuation ratings both pre- and post-outcome, by communication format, are displayed in Table 1, which shows a slight difference between communication formats prior to the outcome. Post-outcome, there was a shift to being more certain about evacuating today. There was a main effect of outcome, F (1, 292) = 98.19, p < .001, η_p^2= .25 and format, F (3, 292) = 5.59, p < .01, η_p^2= .05. Participants were more certain about evacuating today in the verbal condition and least certain decision in the N-V condition. There were no significant interactions (all ps > .12).

Correctness Ratings

A one-way ANOVA revealed a significant effect of communication format on correctness ratings, F (3, 292) = 26.32, p < .001, η_p^2= .22, corresponding to the differences in the credibility ratings. From Table 1, the numerical format was seen as 'least incorrect' and the verbal format seen as most incorrect.

Study 2

Method

Participants

299 Native English speakers were recruited from Amazon MTurk. 17 cases were removed for failing the attention check, leaving a final sample of 281 participants (138 male) aged between 18 – 74 (*Mdn*= 32). Participants received $0.60 for participating.

Design, Materials and Procedure

As in Study 1, except communication format was set in the high probability domain: verbal – "likely", numerical – "80% likelihood", V-N – "likely (80% likelihood)" and N-V– "80% likelihood (likely)". In addition, post-outcome, the likely event did *not* occur.

Results

Trust and expertise ratings were again correlated (pre-outcome: $r = .60$, $p < .001$; post-outcome: $r = .74$, $p < .001$). We combined the two measures as in Study 1. The data were analysed as in Study 1.

Credibility Ratings

Mean credibility ratings, by communication format, are plotted in Figure 2, which shows before the outcome there was little difference between formats, as in Study 1. Post-outcome, all communication formats suffered from a loss of perceived credibility, with no notable difference between formats. The outcome and format interaction of Study 1 was not replicated, F (3, 273) = 2.53, p = .06. The main effect of outcome was significant, F (1, 273) = 221.23, $p < .001$, η_p^2= .45, and the effect of format was marginally significant, F (3, 273) = 2.59, p = .053, η_p^2= .03. A post-hoc Gabriel test revealed there were no significant differences between formats (all *p*s > .08). Highest perceptions of credibility were in the numerical condition (M= 3.91, SE= 0.08), and the lowest were in the verbal condition (M= 3.63, SE= 0.08).

Decision to Evacuate

Mean evacuation ratings for both pre and post-outcome (by communication format) are displayed in Table 1, which shows little difference between formats both pre and post-outcome. Indeed, there was no significant effect of outcome (p = .07) nor format (p = .20) on the decision to evacuate. There was a significant effect of numeracy, F (1, 273) = 5.08, $p < .05$, η_p^2= .02, with the high numeracy group more certain about evacuating (M= 2.08, SE= 0.10), compared to the low numeracy group (M= 2.39, SE= 0.10). There were no significant interactions (all *p*s > .15).

Correctness Ratings

Again there was a significant effect of communication format on correctness ratings F (3, 273) = 4.90, $p < .01$, η^2_p= .05. As in Study 1, the numerical format was seen as 'least incorrect' and the verbal format seen as most incorrect (see Table 1).

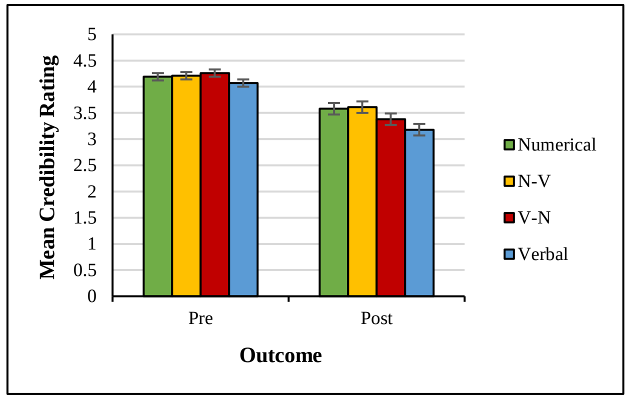

Figure 2. Effect of Communication Format on Perceptions of Credibility Before and After an 'Erroneous' Prediction- (Error Bars Represent ±1SE) – Study 2 – High Probability.

General Discussion

Pre-outcome, people did not perceive any of the volcanologists to be more credible than others using different communication formats, nor was there an effect of format on decision to evacuate. However, post-outcome, credibility was sensitive to an 'erroneous' prediction, with lower ratings in all formats. In Study 1 (low probability), the numerical format was affected least by this, and there was a trend for numerical-led communications (numerical and N-V) to be least affected in Study 2.

It is surprising that there was no initial difference between communication formats on perceptions of credibility in either probability domain, given the findings of Longman et al. (2012) that an expert who used a point estimate was seen as more credible. We would have expected numerical communications to have been rated as more credible, as the decision to use a precise numerical estimate could be thought to reflect a level of confidence and certainty in the prediction. Indeed, people expect experts to provide their knowledge in a precise manner (Shanteau, 1992).

In Study 1, the finding of most interest was the presence of a format × outcome interaction, whereby the numerical format lost least credibility following the occurrence of the unlikely event. These findings could be partly attributed to the directionality of the expression (Teigen & Brun, 1995, 1999). Although both V-N and N-V formats featured a negatively directional expression ('unlikely'), it was accompanied by the positively directional phrase '20% likelihood', which may have cancelled out the effect of the negative directionality. Although no significant interaction was observed with high probability expressions, the results followed a similar trend, with numerical and N-V expressions least affected.

We were surprised not to replicate Doyle et al.'s (2014) findings that more people chose to evacuate when given a risk communication featuring a numerical expression as opposed to a VPE. Although we found an effect of format in Study 1, it was in the opposite direction to the original study. A large number of responses to the question of 'why' people made their evacuation decision mentioned themes such as 'better to

Table 1. *Evacuation and Correctness Ratings for Studies 1 & 2*

Study	Measure	Outcome	Communication Format- Mean Rating (SE)			
			V	V-N	N	N-V
1 (Low Probability)	Evacuation Decision	Pre	3.87 (0.09)	3.87 (0.09)	3.93 (0.09)	3.87 (0.09)
		Post	2.41 (0.13)	2.65 (0.13)	3.31 (0.13)	2.83 (0.12)
	Correctness	Post	1.61 (0.15)	2.06 (0.14)	3.35 (0.14)	2.51 (0.14)
2 (High Probability)	Evacuation Decision	Pre	1.99 (0.17)	2.19 (0.16)	2.42 (0.16)	2.11 (0.16)
		Post	2.08 (0.16)	2.34 (0.15)	2.47 (0.15)	2.32 (0.15)
	Correctness	Post	1.80 (0.13)	2.08 (0.13)	2.45 (0.13)	2.31 (0.13)

be safe than sorry'. There was little cost to the participant to adopt this approach, which could have been a factor in the high proportion of people choosing to evacuate immediately. Whilst Doyle et al. (2014) attributed their results to the ambiguity of VPEs, we argue that our results could also be explained using this reasoning. Participants may have felt that the choice to use a VPE in the risk communication reflected a level of uncertainty in the outcome, with the communicator 'hedging their bets', and thus felt that it was better to adopt a conservative stance and evacuate, 'just in case'. Indeed, this is in line with the appropriate response of increased uncertainty providing an impetus to be concerned and an even greater reason to act (Lewandowsky et al., 2015). Additionally, if an 'unlikely' event were to occur, it would be far more consequential than if a 'likely' event did not occur.

The lack of an influence of numeracy on nearly all of our measures was somewhat unexpected, given the fact that numeracy has been demonstrated to influence effects of communication format (Gurmankin et al., 2004), and information format (e.g. frequencies versus percentages, Reyna, Nelson, Han, & Dieckmann, 2010).

Further research should seek to explore the effect of the precision of the communication format. Chess, Hance & Sandman (1988) claimed being open about levels of uncertainty would lead to enhanced credibility and trustworthiness. The current study only explored point numerical estimates (e.g. '20% likelihood'), rather than more specific point estimates (e.g. '23% likelihood'). Including range estimates (both small and large) would allow for a better understanding of the benefits of including numbers in risk communications. Whilst Longman et al.'s (2014) findings suggest that range estimates will have a negative effect on understanding and perceived credibility, others have found that range estimates are perceived as more useful and more honest (Dieckmann, Mauro, & Slovic, 2010; Johnson & Slovic, 1995).

Conclusion

This study provides a different perspective to examining the effectiveness of risk and uncertainty communications, moving away from merely how the information is understood. Trust is fundamental to improving these communications (Slovic, 1993), and our work contributes to this somewhat neglected area of research.

The present research provided a systematic comparison of the effect of differing communication formats on the credibility of the communicator in the context of geological risk communications. Identifying instances in which the communication format has a significant impact on the audience's perceptions of the communicator is key to building and maintaining public trust in science, as well as improving the effectiveness of risk communication. Our findings show that the numerical format is viewed as more correct and is most robust against reductions in credibility following an 'erroneous' prediction. The present results thus suggest numbers should be included in these communications wherever possible.

Acknowledgements

SJ was supported by a UCL IMPACT studentship, half funded by the British Geological Survey (BGS). RML's contribution is published with permission of the Director, BGS. We are grateful to Dr Charlotte Vye-Brown (BGS) for assistance with creating the vignette used in the studies.

References

Bonnefon, J. F., & Villejoubert, G. (2006). Tactful or doubtful? Expectations of politeness explain the severity bias in the interpretation of probability phrases. *Psychological Science*, *17*(9), 747–751.

Budescu, D. V, Broomell, S. B., & Por, H. H. (2009). Improving communication of uncertainty in the reports of the intergovernmental panel on climate change. *Psychological Science*, *20*(3), 299–308.

Budescu, D. V, Por, H. H., Broomell, S. B., & Smithson, M. (2014). The interpretation of IPCC probabilistic statements around the world. *Nature Climate Change*, *4*(6), 508–512.

Budescu, D. V, & Wallsten, T. S. (1985). Consistency in interpretation of probabilistic phrases. *Organizational Behavior and Human Decision Processes*, *36*(3), 391–405.

Budescu, D. V, & Wallsten, T. S. (1995). Processing linguistic probabilities: General principles and empirical evidence. *Psychology of Learning and Motivation*, *32*(2), 275–318.

Chess, C., Hance, B. J., & Sandman, P. M. (1988). *Improving dialogue with communities: a short guide for government risk communication*. Division of Science and Research, New Jersey Department of Environmental Protection.

Cokely, E., Galesic, M., Schulz, E., Ghazal, S., & Garcia-Retamero, R. (2012). Measuring risk literacy: the Berlin numeracy test. *Judgment and Decision Making*, *7*(1), 25–47.

Dieckmann, N., Mauro, R., & Slovic, P. (2010). The effects of presenting imprecise probabilities in intelligence forecasts. *Risk Analysis*, *30*(6), 987–1001.

Dieckmann, N., Slovic, P., & Peters, E. M. (2009). The Use of narrative evidence and explicit likelihood by decisionmakers varying in numeracy. *Risk Analysis*, *29*(10), 1473–1488.

Doyle, E. E. H., McClure, J., Paton, D., & Johnston, D. M. (2014). Uncertainty and decision making: Volcanic crisis scenarios. *International Journal of Disaster Risk Reduction*, *10*(PA), 75–101.

Erev, I., & Cohen, B. L. (1990). Verbal versus numerical probabilities: Efficiency, biases, and the preference paradox. *Organizational Behavior and Human Decision Processes*, *45*(1), 1–18.

Gurmankin, A. D., Baron, J., & Armstrong, K. (2004). The effect of numerical statements of risk on trust and comfort with hypothetical physician risk communication. *Medical Decision Making : An International Journal of the Society for Medical Decision Making*, *24*(3), 265–271.

Harris, A. J. L., & Corner, A. (2011). Communicating environmental risks: Clarifying the severity effect in interpretations of verbal probability expressions. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *37*(6), 1571–8.

Harris, A. J. L., Corner, A., Xu, J., & Du, X. (2013). Lost in translation? Interpretations of the probability phrases used by the Intergovernmental Panel on Climate Change in China and the UK. *Climatic Change*, *121*(2), 415–425.

House of Lords. (2000). *Science and Technology - Third Report*. Retrieved from http://www.publications.parliament.uk/pa/ld199900/ldselect/ldsctech/38/3801.htm

Jenkins, S., Harris, A. J. L., & Lark, R. M. (2016). “Unlikely” Outcomes Might Never Occur, But What About “Unlikely (20 % Chance)” Outcomes ? In & J. . T. A. Papafragou., D. Grodner., D. Mirman. (Ed.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 390–395). Austin, TX: Cognitive Science Society.

Johnson, B. B., & Slovic, P. (1995). Presenting uncertainty in health risk assessment: initial studies of its effects on risk perception and trust. *Risk Analysis: An Official Publication of the Society for Risk Analysis*, *15*(4), 485–494.

Juanchich, M., & Sirota, M. (2016). *How much will the sea level rise ? Outcome selection and subjective probability in climate change predictions*. Manuscript in progress.

Kasperson, R. (2014). Four questions for risk communication. *Journal of Risk Research*, *17*(10), 1233–1239.

Lewandowsky, S., Ballard, T., & Pancost, R. D. (2015). Uncertainty as knowledge. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *373*(2055).

Lipkus, I. M., Samsa, G., & Rimer, B. K. (2001). General performance on a numeracy scale among highly educated samples. *Medical Decision Making : An International Journal of the Society for Medical Decision Making*, *21*(1), 37–44.

Longman, T., Turner, R., King, M., & McCaffery, K. J. (2012). The effects of communicating uncertainty in quantitative health risk estimates. *Patient Education and Counseling*, *89*(2), 252–259.

Nature. (2010). A Question of Trust. *Nature*, *466*(7302).

Patt, A. G., & Dessai, S. (2005). Communicating uncertainty: Lessons learned and suggestions for climate change assessment. *Comptes Rendus - Geoscience*, *337*(4), 425–441.

Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. (2010). How Numeracy Influences Risk Comprehension and Medical Decision Making. *Psychological Bulletin*, *135*(6), 943–973.

Shanteau, J. (1992). Competence in Experts: The Role of Task Characteristics. *Organizational Behavior And Human Decision Processes*, *53*, 252–266.

Sjöberg, L. (2001). Limits of knowledge and the limited importance of trust. *Risk Analysis*, *21*(1), 189–198.

Slovic, P. (1993). Perceived risk, trust, and democracy. *Risk Analysis*, *13*(6), 675–682.

Teigen, K. H., & Brun, W. (1995). Yes, but it is uncertain: Direction and communicative intention of verbal probabilistic terms. *Acta Psychologica*, *88*(3), 233–258.

Teigen, K. H., & Brun, W. (1999). The Directionality of Verbal Probability Expressions: Effects on Decisions, Predictions, and Probabilistic Reasoning. *Organizational Behavior and Human Decision Processes*, *80*(2), 155–190.

Teigen, K. H., & Brun, W. (2003). Verbal Probabilities: A Question of Frame? *Journal of Behavioral Decision Making*, *16*(1), 53–72.

Teigen, K. H., Juanchich, M., & Riege, A. H. (2013). Improbable outcomes: Infrequent or extraordinary? *Cognition*, *127*(1), 119–139.

Theil, M. (2002). The role of translations of verbal into numerical probability expressions in risk management: a meta-analysis. *Journal of Risk Research*, *5*(2), 177–186.

Wachinger, G., Renn, O., Begg, C., & Kuhlicke, C. (2013). The risk perception paradox-implications for governance and communication of natural hazards. *Risk Analysis*, *33*(6), 1049–1065.

Weber, E. U., & Hilton, D. J. (1990). Contextual effects in the interpretations of probability words: Perceived base rate and severity of events. *Journal of Experimental Psychology: Human Perception and Performance*, *16*(4), 781–789.

Witteman, C. L. M., & Renooij, S. (2003). Evaluation of a verbal-numerical probability scale. *International Journal of Approximate Reasoning*, *33*(2), 117–131.