This is a prepublication manuscript. To access the final version, please see the Journal of Experimental Psychology: Applied. http://www.apa.org/pubs/journals/xap/

How much will the sea level rise? Outcome selection and subjective probability in climate change predictions

Marie Juanchich[a] and Miroslav Sirota[a]

Authors' affiliations:

[a] Department of Psychology, University of Essex.

Corresponding author:
Marie Juanchich
Address: Department of Psychology, University of Essex, Wivenhoe Park, CO4 3SQ, Colchester, UK.
Email: m.juanchich@essex.ac.uk
Tel: (+44) 120 687 38 12

Miroslav Sirota
Address: Department of Psychology, University of Essex, Wivenhoe Park, CO4 3SQ, Colchester, UK.
Email: msirota@essex.ac.uk
Tel: (+44) 120 687 42 29

Acknowledgment

The authors are grateful to Professor Karl-Halvor Teigen for helping to shape our initial ideas and to Grantham Professor Theodore Shepherd for his useful insights into climate change science. We also thank reviewers for their suggestions.

This is a prepublication manuscript. To access the final version, please see the Journal of Experimental Psychology: Applied. http://www.apa.org/pubs/journals/xap/

Abstract

We tested whether people focus on extreme outcomes to predict climate change and assessed the gap between the frequency of the predicted outcome and its perceived probability while controlling for climate change beliefs. We also tested two cost-effective interventions to reduce the preference for extreme outcomes and the frequency-probability gap by manipulating the probabilistic format: numerical or dual-verbal-numerical. In four experiments, participants read a scenario featuring a distribution of sea level rises, selected a sea rise to complete a prediction (e.g., "It is 'unlikely' that the sea level will rise … inches") and judged the likelihood of this sea rise occurring. Results showed that people have a preference for predicting extreme climate change outcomes in verbal predictions (59% in Experiments 1-4) and that this preference was not predicted by climate change beliefs. Results also showed an important gap between the predicted outcome frequency and participants' perception of the probability that it would occur. The dual-format reduced the preference for extreme outcomes for low and medium probability predictions but not for high ones and none of the formats consistently reduced the frequency-probability gap.

Key words: Uncertainty, prediction, probability, frequency, climate change.

Public Significance Statement: Four experiments advance our knowledge of the way people predict climate change based on a distribution of sea rise projections. People tend to predict extreme outcomes and to over-estimate their chance of occurrence. To reduce miscommunications, forecasters should state whether they predict a minimal or maximal outcome and should express their predictions in such a way that the probability of the outcome matches its frequency of occurrence in the projections.

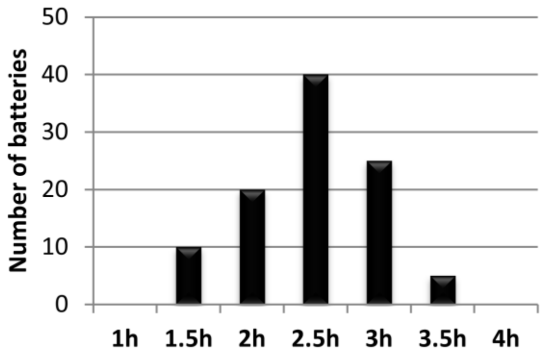
Introduction

Experts are almost unanimous: human-made climate change is happening right now (Cook et al., 2013). The public have developed a better awareness of climate change and its risks (Read, Bostrom, Morgan, Fischhoff, & Smuts, 1994; Reynolds, Bostrom, Read, & Morgan, 2010) although they may not see climate change as a priority (Leiserowitz, Maibach, Roser-Renouf, Feinberg, & Howe, 2013). A large body of research has focused on the ways that people understand climate change predictions (e.g., Budescu, Por, & Broomell, 2012; Budescu, Por, Broomell, & Smithson, 2014). However, we know very little about how people form climate change predictions. Recent evidence suggests that people tend to focus their predictions on extreme and rare outcomes rather than moderate outcomes (Juanchich, Teigen, & Gourdon, 2013; Teigen, Juanchich, & Filkuková, 2014; Teigen, Juanchich, & Riege, 2013). This preference for predicting extreme and rare outcomes can be detrimental to the accurate communication of climate change events because extreme outcomes are, by definition, very unlikely. Here, we studied the extent to which people's preference for extreme outcomes occurs in climate change predictions. We investigated the beliefs and attitudes that may underpin the preference for extreme outcomes and tested two cost-effective interventions to reduce this preference.

The preference for extreme and rare outcomes in Outcome Completion Tasks

A recent research trend focuses on how people form verbal probability predictions (i.e., predictions formed from a verbal probability and an outcome). This approach typically uses an Outcome Completion Task (also called a Which Outcome Task). In this task, participants are typically provided with an approximately Gaussian frequency distribution of continuous outcomes and are asked to complete an unfinished prediction with an appropriate quantity (see Figure 1; Juanchich et al., 2013; Teigen & Filkuková, 2013; Teigen et al., 2013; Teigen et al., 2014).

A sample of computers of the brand "Comfor" were tested to check how long the batteries last before they need to be recharged. The figure below shows how many batteries lasted how many hours. Please complete the prediction below with the outcome that seems most appropriate in this context.



"It is unlikely that the battery will last ... hours."

Figure 1. Example of an Outcome Selection Task used in Teigen et al., 2013.

In the above-mentioned example, when selecting an outcome to form a prediction, participants had a strong and consistent preference for extreme and rare outcomes. For example, whether they predicted a battery duration that was "unlikely", "possible" or "virtually certain", participants selected an outcome that was located in the extreme ends of the distribution, or even beyond the distribution, and therefore had a very low frequency (0%-10%). This preference for predicting extreme and rare outcomes was consistently observed when the distribution of outcomes was presented as a narrative, a table or a graph (Teigen et al., 2014), or when the distribution was not normal but bimodal or skewed (Juanchich et al., 2013).

How likely are extreme outcomes?

Extreme outcomes and the IPCC guidelines. The preference for extreme and rare outcomes for low and high probability predictions is not in line with the recommended usage of verbal probability issued by the Intergovernmental Panel on Climate Change (IPCC) (shown in Table 1). According to the IPCC, "unlikely" and "virtually certain" should be used to convey probabilities of around 20% and over 99% and should therefore be used to qualify outcomes that have a 20% or 99% frequency. Using "unlikely" for a 0% frequency outcome is not technically incorrect, although this frequency would better match the verbal probability "exceptionally unlikely".

Table 1

Likelihood communication guidelines of the Intergovernmental Panel on Climate Change (IPCC, 2007, 2010, 2013).

Verbal probability	Likelihood of occurrence
Virtually certain	99-100%
Very likely	90-100%
Likely	66-100%
About as likely as not	33-66%
Unlikely	0-33%
Very unlikely	0-10%
Exceptionally unlikely	0-1%

Although it does not appear in the IPCC-recommended probability lexicon, "possible" is often used to convey uncertainty in the IPCC reports (Intergovernmental Panel on Climate Change, 2014). "Possible" means, on average, a probability range between 35 and 60%, (Budescu & Wallsten, 1985; Juanchich, Sirota, & Butler, 2012; Reagan, Mosteller, & Youtz, 1989). It should therefore be used for outcomes that have a similar frequency of 35-60%. In the IPCC report, "possible" seems to be used in place of the recommended medium verbal probability "as likely as not", given that they both communicate a medium probability magnitude.

The focus on extreme outcomes leads to probability over-estimations. When an outcome is extreme and rare (i.e., ≤ 10% frequent), it should follow that this outcome is very unlikely (e.g., Lichtenstein, Fischhoff, & Phillips, 1982). Therefore, when speakers are predicting an extreme outcome (0%-10% frequency), we would expect that they believe that these outcomes are 0% to 10% likely and we would hope that recipients of such a prediction would believe that those outcomes are 0% to 10% likely. Yet, recent evidence points to a gap between the frequency of the outcome and its perceived probability. Participants who selected an outcome that had a 0% frequency, in order to make an "unlikely" prediction, believed that the outcome had a 20% chance of occurring – instead of believing that its chances of occurring were close to 0 (Teigen et al., 2013, Experiment 5).

The selection of extreme and rare outcomes can also create some communication issues between speakers and recipients. The probabilities typically perceived by recipients for "unlikely", "possible" and "certain" were much higher than the frequencies of the outcomes they qualified (Juanchich et al., 2013; Teigen et al., 2014; Teigen et al., 2013). For example, "possible" conveys a 50% probability (Juanchich et al., 2012), yet, participants predicted that an outcome that had a 10% frequency was "possible" (Teigen et al., 2014). Experiments varying the shape of the outcome distribution indicated that the preference for extreme outcomes varied as a function of the frequency of the outcome but did not fully disappear. For example, "possible" was less often associated with the maximal outcome when participants formulated their prediction based on a U-shaped outcome distribution with the minimal outcome being the most frequent (Juanchich et al., 2013, Experiment 2).

Uncertainty in climate change predictions

Predicting how much the climate will change, and how much this will affect humans and their environment, is a key issue for climate change scientists. Their goal is not to prove whether or not the climate is changing but to communicate the degree to which it is changing, and the extent to which people will have to adapt to these changes. When choosing predictions relating to the rise of sea levels, in order to make a 20% likely prediction, participants should ideally choose a non-extreme sea level rise which occurs in 20% of the projections. This outcome should then be perceived as having a 20% probability of occurring. However, based on past research, we expect that participants will select an extreme outcome and then believe that this outcome is quite likely. If proved true, this finding could have important consequences for climate change communication. For example, let’s assume that Paul reads the IPCC report and discovers that a sea level rise of 5 cm is the minimum sea rise that can be expected over the next 50 years and that this has occurred in 10% of the climate change projections. Paul might then tell his friend Simon that “a sea level rise of 5 cm is certain”. Simon could therefore believe that this sea level rise has a 90-100% chance of occurring, thus largely over-estimating the probability of observing such a sea rise. Paul’s probability perception itself could also be distorted. For example, he could subsequently believe that a 5 cm sea level rise has a 90-100% chance of occurring although it actually has a 10% probability of occurring. If either Paul or Simon wants to buy a house on the seafront, they will need to assess whether the house is at risk from the effects of the sea level rise. If they have a calibrated perception of the probability, they will consider that the sea is likely to rise more than 5 cm, but if their perception is distorted, they will consider that “it is certain” that the sea will rise by *only* 5 cm. Decision quality can only be as good as the elements that are used to achieve the decision. Therefore, a biased subjective probability of climate change outcomes would lead to proportionally poor climate change decisions (Yates et al., 1989).

Climate change frequencies rely on climate change projections which are based on past climate change conditions and on a set of assumptions about the future (e.g., the trend of CO_2 emissions). Climate change is a single event and the projection frequencies do not represent the actual frequencies of occurrence of climate change outcomes in the past. Climate change probabilities are therefore single event probabilities and do not *have to* be formally derived from projected frequencies. Yet, projections are, to date, the best tool available with which to model the climate of the future and its consequences (IPCC, 2013). Refusing to use climate change predictions to predict the climate of the future would be similar to refusing to light a candle while searching a needle in a dark room. Further, when projections are considered to be reliable, the IPCC recommends their use to form probabilities of climate change outcomes (Intergovernmental Panel on Climate Change, 2013a). Therefore, our position is that projected frequencies should be considered as a benchmark for forming climate change predictions and, therefore, it is assumed that the perceived probabilities of climate change outcomes should reflect their frequencies of occurrence in projections.

The preference for extreme outcomes: does it apply to a real-life issue?

The preference for extreme outcomes in the Outcome Completion Task was replicated across a range of outcomes (Juanchich et al., 2013; Teigen et al., 2014; Teigen et al., 2013). However, participants dealt with hypothetical situations (e.g., the battery life of a hypothetical computer brand). The hypothetical nature of those outcomes may have hindered participants' motivation to match the outcome frequency with the probability conveyed by the prediction. For instance, participants may have selected an extreme outcome due to lack of motivation. Extreme and rare information is considered to be more important, attracts more attention (Fiske, 1980) and is deemed to be more informative (McKenzie & Mikkelsen, 2000). People even prefer incorrect extreme predictions to less extreme correct ones (McKenzie & Amin, 2002). Forming a prediction regarding a relevant and topical real-life outcome (Hallegatte et al., 2016) could increase people's interest in the task (Hidi, 2001) and their efforts to select an outcome with a frequency matching the probability conveyed by the prediction. Replicating previous findings in a climate change setting would strengthen the generalised nature of the prior results and eliminate a lack of cognitive motivation.

Interpretation of outcome distributions: frequencies or cumulative frequencies?

In the Outcome Completion Task, an outcome value was considered to be rare when it had a frequency equal to or below 10%. The frequency of outcomes was simply derived from the number of times it occurred in the sample of outcomes (normalised in percent). However, one could also consider the cumulative frequency of the outcomes. As shown in Figure 2, according to a cumulative frequency reading of a typical distribution of outcomes, the minimal outcome occurs 100% of the time (as indicated in the grey columns). Cumulative frequencies apply well to outcomes that can be ranked on a scale and for which larger outcomes entail the occurrence of smaller ones[1].

[1] Wind strength is an example of an ordinal non-scalar outcome: a wind speed of 20 knots does not mean the wind was blowing at 5 knots at any point.

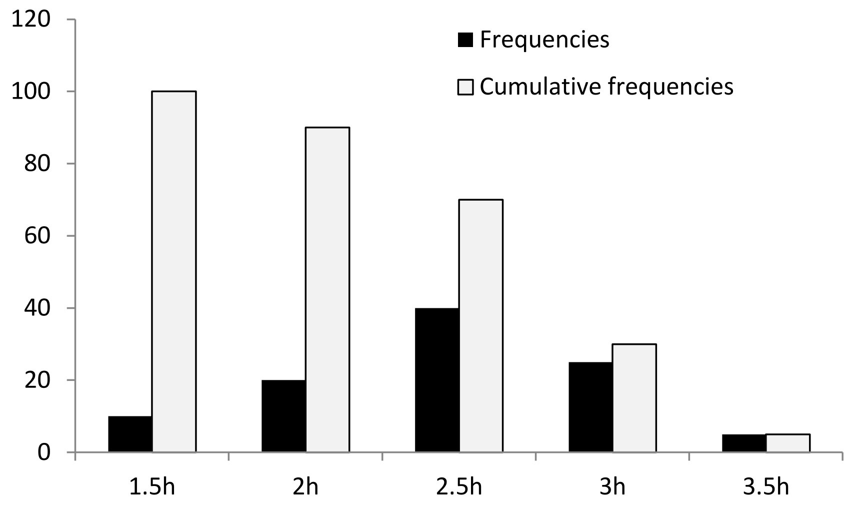


Figure 2. Frequency (black) and cumulative frequency (grey) of the duration of a computer battery; adapted from Teigen et al. (2014).

The use of cumulative frequencies may explain the preference for minimal outcomes for “certain” predictions. The minimal outcome has a 100% cumulative frequency and could therefore be considered to be certain. Consistent with this possibility, “certain” predictions often feature a lower bound interval marker (e.g., “It is certain that the battery will last *at least* one hour”; Juanchich et al., 2013). This is also in line with psycholinguistic findings on numerical quantities showing that people often infer that quantities stated are the minimal quantities to be expected. For example, “Two of the students did well in the test” does not exclude the fact that more than two students did well in the test (Horn, 1984). For a review on the exact vs. at least interpretations of numerical quantities, see Breheny (2007).

However, the use of cumulative frequencies does not explain the preference for outcomes for “unlikely” and “possible” predictions: their cumulative frequency is still much lower (0% and 10% respectively) than their subjective probabilities (20% and 50%). Further, results showing that the preference for the maximal outcome can be swapped to the minimal outcome indicate that the cumulative frequency may not be used to select the outcomes for “possible” predictions. For example, when making a prediction for a house-buyer, participants focused on the possible maximum price for a house (10% cumulative frequency), whereas when talking to a buyer, they focused on the "possible" minimum one (100% cumulative frequency (Teigen et al., 2014).

The IPCC does not specify whether scientists should use the frequencies of precise outcomes or the cumulative frequencies of outcome ranges to derive probability estimates. However, "virtually certain" does not seem to be used to qualify the minimum outcome to be expected in the summary for policy-makers (e.g., 0 occurrence in the Intergovernmental Panel on Climate Change, 2013b). In fact, there seems to be a range of practices, such as predicting a minimal outcome to be expected (e.g., global surface temperature change for the end of the 21st century is likely to exceed 1.5°C) or a range to be expected (e.g., the global mean surface temperature change will likely be in the range of 0.3°C to 0.7°C). In both cases, the implication seems to be that changes smaller than those ranges are "unlikely" – although in a strict cumulative reading of frequencies they would be considered to be "certain" (Intergovernmental Panel on Climate Change, 2013b, p 18).

In the present set of experiments, we examined both the gap between subjective probability and frequency along with the gap between subjective probability and cumulative frequency. The difference between the two gaps may be indicative of the frequency participants used to derive their probability estimates.

Who prefers to predict extreme climate change outcomes?

We have little knowledge of the mechanism underpinning the preference for extreme outcomes, nor the variability in outcome selection. A variable that may contribute to the selection of specific outcomes is people's personal beliefs. We hypothesised that people who believe more strongly in climate change would be more likely to select an extreme top climate change outcome. In support of this motivated reasoning hypothesis, evidence has indicated that people's prior beliefs affect the way they understand climate change predictions (Budescu, Broomell, & Por, 2009; Budescu et al., 2012; Budescu et al., 2014; Corner, Whitmarsh, & Xenias, 2012; Kahan, Jenkins-Smith, & Braman, 2011). For example, a global warming prediction has been perceived to have a greater probability by climate change believers than by sceptical individuals (Budescu et al., 2009; Budescu et al., 2012; Budescu et al., 2014). Kahan et al. (2011) proposed that the biased perceptions aimed to protect existing views and justify (potentially inadequate) behaviours.

Interventions to reduce the selection of extreme outcomes

When making an extreme prediction, participants are posited to rely on a dispositional form of uncertainty (Teigen et al., 2013), which is based on "tendencies or predispositions favoring specific outcomes" (Keren & Teigen, 2001; Popper, 1959, 1990). In contrast, when interpreting the meaning of a prediction, people are believed to rely on a distributional form of uncertainty, which is based on "the relative frequency of the occurrence of an attribute or an event approaching an infinite number of observations" (Keren & Teigen, 2001, p. 1,011). Therefore, it can be hypothesised that if we elicit a more distributional uncertainty, participants will be less likely to select an extreme outcome with which to form their predictions.

Intervention 1. Using a numerical format. Windschitl and Wells (1998) proposed that numbers elicit a rule-based uncertainty, which follows the rules of distributional uncertainty, whereas verbal probabilities elicit a more intuitive uncertainty, which is closer to a dispositional form of uncertainty. For example, participants produced more biased estimates on verbal probability scales than on numerical probability scales (Windschitl & Wells, 1996, 1998). If different probabilistic formats trigger different variants of uncertainty, then using numerical probabilities will reduce the selection of extreme outcomes and the gap between the frequency of the outcome and the subjective belief that the outcome will occur. Numerical probabilities also have the advantage of being more precise than verbal probabilities, which restrains the between-subjects variability in probabilistic interpretation (Budescu et al., 2009; Budescu et al., 2012; Budescu et al., 2014).

Intervention 2. Using a verbal-numerical dual-format. Evidence suggests that a dual-format – presenting verbal and numerical probabilities together – brings the best of both worlds: the ease of processing of verbal probabilities (Witteman & Renooij, 2003) and the precision and distributional uncertainty associated with numbers (Windschitl & Wells, 1996). Additionally, compared with verbal probability only, the dual verbal-numerical format decreased probability perception variability and increased the consistency with the guidelines of interpretations of verbal probabilities given by the IPCC and shown in Table 1 (Budescu et al., 2009; Budescu et al., 2012; Budescu et al., 2014).

Aims of our research

There were three main goals associated with the present research. Our first goal was to test whether people were biased towards selecting rare and extreme outcomes when making climate change predictions. First, we investigated which outcome participants selected to complete sea rise predictions conveying low (Exp. 1 and 4), medium (Exp. 2 and 4) and high probabilities (Exp. 3 and 4) in both between-subjects (Exp. 1-3) and within-subjects (Exp. 4) designs. We extended previous research by testing the preference for extreme outcomes in a real-life context that mattered to people, and by investigating whether individual differences in climate change beliefs and ecological attitudes predicted outcome selection. We expected to replicate the preference for extreme outcomes in climate change predictions and that stronger climate change beliefs and pro-ecological attitudes would be linked with more extreme high climate change outcomes. Second, we tested whether participants' perceptions of the probability that this outcome would occur were attuned with the frequency or to the cumulative frequency of this outcome, thereby assessing a frequency-probability gap. We expected that participants would over-estimate the probability of the event. Third, we tested two theory-driven and low-cost interventions based on the probabilistic format of the prediction. These interventions were expected to reduce the preference for extreme and rare outcomes and reduce the gap between frequency and subjective probabilities.

Experiment 1

Method, Experiments 1-3

Participants, Experiment 1. A sample of 101 participants from Amazon Mechanical Turk completed a web questionnaire. The participants were Americans who successfully completed more than 80% of their Mechanical Turk tasks (percentage of hit success). Eight participants were excluded from the sample because they either failed to respond correctly to an instructional manipulation check (Oppenheimer, Meyvis, & Davidenko, 2009) or completed the outcome question with a range of outcomes instead of a single outcome (e.g., "10 to 20 inches"). In the final sample of 93 participants, 41% were female and the mean age was 31.5 years (18-65, *SD* = 11.0). Most participants were White Caucasian (81%), had a higher education (at least two years of college, 61%), and were employed (66%). Participants reported their political leanings as Democrat (42%), Independent (32%), Republican (17%) or other (9%).

Participants, Experiment 2. A new sample of 103 participants from Amazon Mechanical Turk completed a web questionnaire. Five participants were excluded from the sample based on the same exclusion rule as before. In the sample of 98 participants, 32% were female and the mean age was 30.0 years (18-60, *SD* = 8.5). Most participants were

White Caucasian (85%), had a higher education (74%), and were employed (67%). Participants reported their political leanings as Democrat (45%), Independent (43%), Republican (9%) or other (3%).

Participants, Experiment 3. A new sample of 101 participants from Amazon Mechanical Turk completed a short web questionnaire. Thirteen participants were excluded from the sample based on the same exclusion rule as in previous experiments. In the final sample of 88 participants, 32% were female and the mean age was 28.8 years (18-72, *SD* = 9.8). Most participants were White Caucasian (85%), had a higher education (70%) and were employed (64%). Participants reported their political leanings as Democrat (49%), Independent (36%), Republican (9%) or other (6%).

Design. In a between-subjects design, the format of the probability used in the prediction was manipulated: verbal, numerical or dual (see Table 2). The numerical format was a probability range following the probability communication guidelines of the IPCC. In Experiment 1, the probability conveyed was low (<10%), in Experiment 2 the probability conveyed was medium (33%-66%) and in Experiment 3 the probability conveyed was very high (>99%).

Materials and procedure. After giving their informed consent (Appendix A), participants completed 19 randomly presented items measuring their beliefs and attitudes towards climate change (Appendix B). Four items originated from the *Belief in Global Climate Change Occurrence Scale* (e.g., "It seems to me that weather patterns have changed compared to when I was a child", Cronbach's α = .91, Heath & Gifford, 2006). The remaining 15 items originated from the *Revised New Ecological Paradigm* and measured pro-ecological attitudes (Dunlap, Van Liere, Mertig, & Jones, 2000, Cronbach's α = .85). The pro-ecological scale assessed people's attitudes towards the environment and the role of humans in environmental changes (e.g., "We are approaching the limit of the number of people the earth can support"). The scale also included an attention-checking question: "Please select 'strongly disagree' to show that you are reading the instructions".

Then, participants read a climate change vignette based on Harris and Corner (2011), describing how ice is currently melting in the Arctic. Participants read that experts from the Intergovernmental Panel on Climate Change conducted 100 projections of how much the sea level will rise by 2100 and were shown the distribution of sea rises (see Figure 3).

Table 2

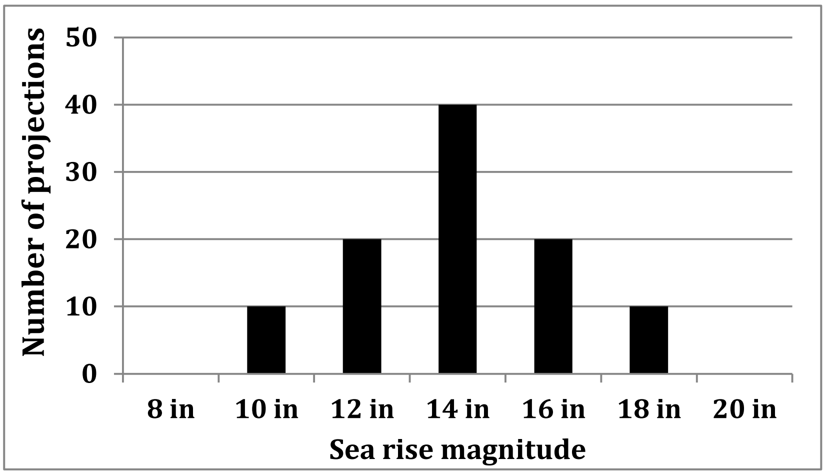
Predictions completed by participants in Experiments 1-4 according to the probability conveyed and the probabilistic format.

Probability conveyed	Verbal format	Numerical format	Dual-format
Low	It is unlikely that the sea will rise ... inches.	There is a probability between 10% and 33% that the sea level will rise ... inches.	It is unlikely (probability between 10% and 33%) that the sea level will rise ... inches.
Medium	It is possible that the sea will rise ... inches.	There is a probability between 33% and 66% that the sea level will rise ... inches.	It is possible (probability between 33% and 66%) that the sea level will rise ... inches.
High	It is virtually certain that the sea will rise ... inches.	There is a probability between 99% and 100% that the sea level will rise ... inches.	It is virtually certain (probability between 99% and 100%) that the sea level will rise ... inches.

The X-axis of the bar chart ranged from 8 to 20 inches (20 to 51 cm) with projections showing sea level rise values ranging from 10 to 18 inches (25 to 46 cm). This corresponds to the projections based on the Representative Concentration Pathway 2 scenarios provided by the IPCC in 2007 (RCP2.4). These values are lower than the values shown in projections based on more pessimistic RCP such as the RCP8.5, which showed a sea level rise from 20 to

39 inches (52 to 98 cm). Participants completed the low probability prediction of a sea rise magnitude by writing a number in a space provided. The probabilistic format of the prediction was either verbal, numerical or dual (see Table 2). Participants could easily infer the frequency of each outcome normalised as a percentage given that the Y-axis provided the number of projections in which the outcome was observed out of 100 projections.

Global warming is melting the ice at both the North Pole and the South Pole. The melting of ice leads to a rise in global sea levels. Experts from the Intergovernmental Panel on Climate Change (IPCC) conducted 150 projections of the sea rise that will be observed in 2100. The projections show that by 2100 the sea level will rise between 10 and 20 inches, as depicted below.



Please complete the following statement with the sea rise magnitude that you judge most appropriate in this context.

It is unlikely that the sea will rise .... inches.

Figure 3. Vignette used in Experiments 1-4 showing a distribution of sea level rises ranging from 8 to 20 inches (20 - 51 cm) in the low and verbal probability prediction condition.

Afterwards, participants judged to what extent they supported four governmental strategies to mitigate climate change: automobile use reduction, industry emissions reduction, alternative energy promotion and recycling promotion. These strategies were taken from Reynolds et al. (2010). Judgments were provided on a slider ranging from 0: *I do not at all support this action* to 100: *I definitely support this action* with increments of 1 (see Appendix

C). The judgments of support for climate change mitigation were aggregated into a score reflecting support for climate change mitigation, Cronbach's $\alpha = .84$; $M = 74.5$, $SD = 21.7$. The scores for this variable were consistently very high, possibly indicating a ceiling effect. Support for climate change mitigation was not related to the format of the prediction and to the sea rise probability perceptions in Experiments 1-3. Further, in those experiments, support for climate change mitigation was positively correlated with climate change belief and pro-ecological attitude, Pearson's r ranged between .44 and .71, all ps < .001. The three variables were used as covariates in our analyses.

Participants then assessed the probability associated with the prediction they completed ("Based on the forecast that the sea level will rise by [number provided by the participant] inches, please provide a number that you think matches the chances that this sea level rise will happen"). Participants expressed their answer by placing a cursor on a slider ranging from 0 to 100 with increments of 1 (0: *it is impossible*, 100: *it will definitely happen*). Finally, participants completed a socio-demographic questionnaire.

Variables and statistical analysis. Prior to the main statistical analyses, we created three variables: the outcome extremeness variable, the frequency-probability gap variable and the cumulative-frequency-probability gap variable.

The outcome extremeness variable. The outcomes selected by participants were coded as "extreme low" if they formed the low tail of the distribution (10 inches sea rise) or if they were lower than this tail outcome. The outcomes were coded as "extreme high" if they formed the high tail of the distribution (18 inches sea rise) or if they were higher. These extreme outcomes had a low frequency: between 0 and 10%. All the other outcomes were categorised as non-extreme and had a frequency between 20 and 40%. The proportion of exact outcomes selected by participants is available in the supplementary materials.

The frequency-probability gap variable. This variable represents the difference between the frequency of the outcome and participants' probability perception that this outcome could occur in the future. To compute this variable we subtracted the normalised frequency of the outcome selected (in %) from the probability score given by participants (also in %). Frequency-probability gap scores close to zero indicated a strong match between frequency and subjective probability, whereas scores far from zero indicated a larger frequency-probability gap, with negative scores indicating a probability under-estimation and positive scores a probability over-estimation. For example, a person who chose a 10 inch sea rise (which occurred in 10% of the projections) and who reported that this sea rise had a 20% probability of occurring, had a frequency-probability gap score of 10% (20% probability – 10% frequency).

The cumulative-frequency-probability gap variable. We subtracted the cumulative frequency of the outcome chosen by participants from the probability given by participants. Minimal or lower outcomes had a cumulative frequency of 100% given that they happened in all of the projections. Higher outcomes were allocated a frequency of 100 minus the sum frequency of the lower outcomes. For example, the 12 inches outcome has a 20% frequency and a 90% cumulative frequency; a 20% frequency because it occurred in 20% of the projections, and a 90% cumulative frequency because a 12-inch sea rise also happened in all the sea level rises equal to or above 12 inches, which adds up to a total frequency of 90%.

The two gap variables reflect the degree to which the frequency of sea level rises observed in a set of projections match (or do not match) the outcome's perceived probability of occurrence. The comparison of the magnitude of the two gaps can be used as an indicator of whether participants have used the frequency of the outcomes or their cumulative frequency to derive their probability perceptions.

Across the experiments, the two gap variables had a satisfactory skew and kurtosis but were sometimes bimodal. We investigated the effect of format on the two frequency gap variables using parametric and non-parametric analyses. For Experiments 1-3, we reported the parametric analyses because they offer the flexibility to include the covariates belief in climate change, pro-ecological attitudes and climate change mitigation support. However, we also reported the non-parametric results (Kruskal Wallis) in the single instance where the non-parametric results departed from the parametric results (Experiment 1 for the frequency-probability gap).

Results

Preference for extreme outcomes in low probability predictions. In order to form a prediction about what is "unlikely", participants most often selected an extreme outcome from beyond the distribution range, mostly from beyond the maximal outcome (Table 3). We compared the selection of extreme outcomes according to the prediction format using a chi-square test followed by Bonferroni adjusted pairwise comparisons (visible with subscript letters in Table 3). The analysis showed that the format affected the choice of the outcome in the prediction completion, χ^2 (4, N = 93) = 43.64, p < .001, Cramer's V = .48. Compared to the verbal condition, participants selected an extreme outcome less often in the numerical and the dual conditions. Further, participants selected an extreme outcome more often in the dual condition than in the numerical one.

A multinomial logistic regression analysis tested the relationship between outcome extremeness and climate change belief, mitigation support and pro-ecological attitude. The model had a fair fit, χ^2 (176, N = 93) = 185.01, p = .306. Altogether, the variables in the model did not predict variance in outcome selection, supporting the fact that climate change belief, mitigation support and pro-ecological attitude did not guide the outcome selected to complete the prediction, χ^2 (6, N = 93) = 3.25, p = .778, Cox and Snell Pseudo R-square = .03.

Table 3

Outcome selected (in %) with its average frequency, cumulative frequency and subjective probability according to the format for Experiments 1, 2, and 3 (n per cell varying from 25 to 35 cases).

	Prediction format		
Exp./ Outcome	**Verbal**	**Numerical**	**Dual**
Exp. 1 (low probability)			
Outcome	% selection	% selection	% selection
Moderate	10%$_a$	85%$_b$	38%$_c$
Extreme low	13%$_a$	12%$_a$	21%$_a$
Extreme high	77%$_a$	3%$_b$	41%$_c$
Total extreme	90%	15%	62%
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Frequency	5.2 (12.1)	23.0 (12.3)	14.1 (11.3)
Cumulative frequency	20.6 (37.3)	75.8 (24.9)	45.2 (40.1)
Subjective probability	24.6 (26.8)	54.0 (26.2)	29.3 (21.8)
Exp. 2 (medium probability)			
Outcome	% selection	% selection	% selection
Moderate	53%$_a$	82%$_b$	97%$_b$
Extreme low	10%$_a$	6%$_a$	0%$_a$
Extreme high	37%$_a$	12%$_{a b}$	3%$_b$
Total extreme	47%	18%	3%
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Frequency	25.3 (15.0)	31.5 (14.8)	36.9 (8.0)
Cumulative frequency	49.7 (33.0)	63.3 (26.4)	70.6 (12.4)
Subjective probability	57.8 (20.7)	53.5 (14.6)	54.7 (18.0)
Exp. 3 (high probability)			
Outcome	% selection	% selection	% selection
Moderate	80%$_a$	68%$_a$	53%$_a$
Extreme low	20%$_a$	26%$_a$	44%$_a$
Extreme high	0.0%$_a$	7%$_a$	3%$_a$
Total extreme	20%	33%	47%
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Frequency	31.6 (12.8)	28.7 (14.8)	24.4 (16.3)
Cumulative frequency	73.6 (18.0)	74.8 (22.9)	74.2 (22.4)
Subjective probability	67.3 (18.4)	74.2 (22.4)	75.3 (22.4)

Note: Superscript letter shows the pair of proportions that were statistically different at $p < .05$ with Bonferroni adjustment.

The frequency-probability gaps. On average, participants chose outcomes with a 13.9% frequency, but they believed that those outcomes had a 36.0% probability of occurring, hence over-estimating the probability of those outcomes and showing an average frequency-probability gap of 22.1%. This is clearly visible in Figure 4, which shows that the frequency and subjective probability distributions do not overlap well, and in Table 4, which shows the average differences (i.e., the gaps). The gap between the cumulative frequency of the chosen outcomes and participants' probability perceptions was also fairly wide (-11 to -22). The difference between cumulative frequency and probability perception was smaller than the difference between frequency and probability perception for 26% of participants, hinting that this set of participants may have relied more on cumulative frequencies than frequencies to form their probability perceptions.

Our goal was to test the effect of the format together with the climate change individual differences on the difference between frequencies and probability judgments. In order to do this we conducted a multivariate analysis of variance with the frequency-probability gap and the cumulative-frequency-probability gap as the dependent variables, the prediction format as the independent variable and climate change belief, climate change mitigation support and pro-ecological attitude as the covariate. The format of the prediction had an effect on the frequency-probability gap (see the results of the omnibus test and the pairwise comparisons in Table 4). Compared to the verbal format, the dual format reduced the frequency-probability gap, whereas the numerical format increased it. However, the non-parametric analyses showed no effect of the format on the frequency-probability gap, χ^2 (2, N = 64) = 4.75, p = .093.

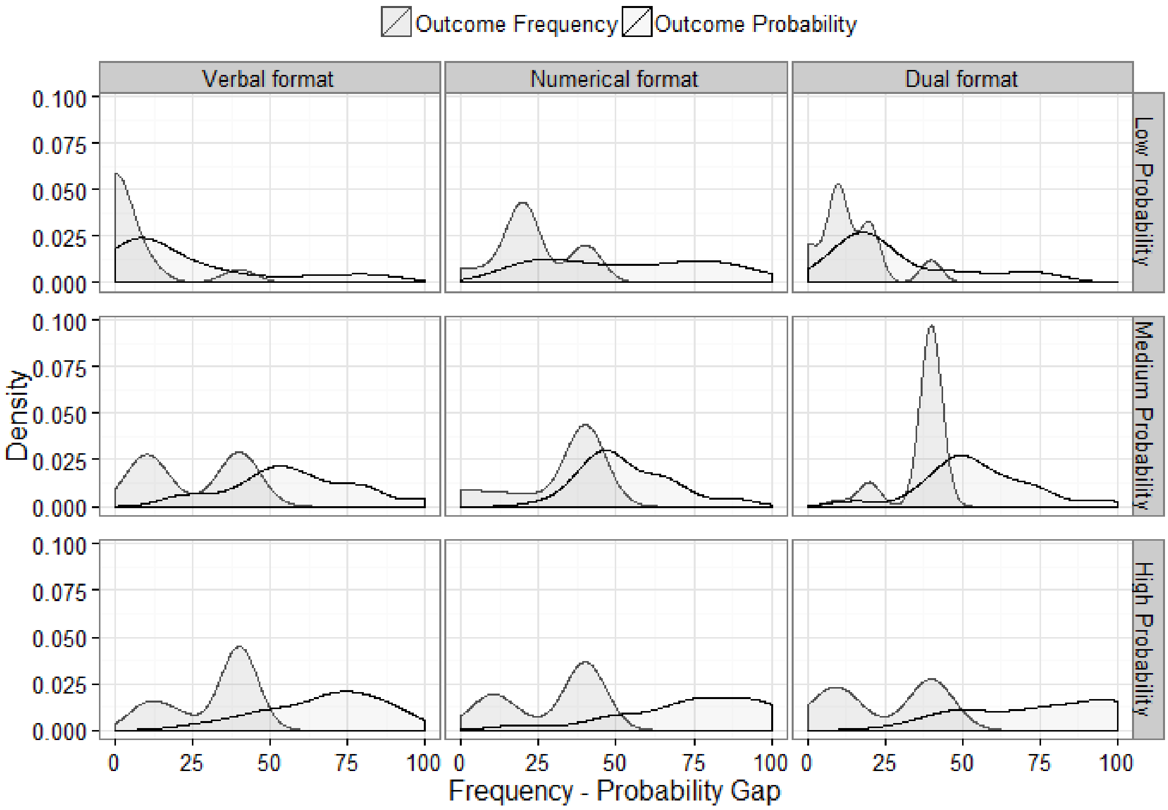


Figure 4. Distribution of the frequency (dark grey) and subjective probability (light grey) of the outcome selected by participants according to the format of the prediction for low, medium and high probability predictions respectively tested in Experiments 1-3.

The format of the prediction had an effect on the cumulative-frequency-probability gap (see Table 4). The gap was at its widest in the dual-format condition and at its smallest in the verbal condition. Pairwise comparisons showed only one statistically significant difference, between the verbal and the numerical format condition. The gap in the dual-format condition was in between the gap in the verbal and numerical conditions and was not statistically different to those. Finally, the multivariate tests showed that climate change belief, climate change mitigation support and pro-ecological attitude did not predict the magnitude of the two gap variables, F (2, 86) = 2.16, p = .121, $\eta_p^2 = .05$, F (2, 86) < 1, $\eta_p^2 < .01$ and F (2, 86) = 2.03, $p = .138$, $\eta_p^2 = .05$.

Table 4

Mean frequency-probability and cumulative-frequency-probability gaps for low probability predictions according to the probabilistic format along with pairwise comparisons according to format (V: verbal, N: numerical, and D: dual; n per cell between 29 and 33; *N* = 93,).

				Omnibus			Pairwise comparison Mdiff and 95% CI		
Gaps	V	N	D	F (2, 92)	*p*	η_p^2	**N** *vs* V	**D** *vs* V	**D** *vs* N
Freq-proba	19.5 (24.3)	31.0 (27.4)	15.2 (19.6)	4.64#	.012	.10	13.8* [0.7, 27.5]	-2.4^{ns} [-16.5, 11.8]	16.2* [1.9, 30.4]
C-freq-proba	4.0 (37.9)	-21.8 (33.6)	-15.9 (38.3)	3.93	.023	.08	-23.8* [-45.8, -1.9]	19.1^{ns} [-3.5, 41.8]	4.7^{ns} [-18.1, 27.5].

Note. *: *p* < .05. Freq-proba: frequency-probability gap and C-freq-proba: cumulative-frequency-probability gap). # This result was not consistent with the non-parametric analysis, which showed no effect of format in this instance. Mse for freq-proba and C-freq-proba in the MANCOVA: 490.04 and 1255.38;

Many participants selected an outcome that was so extreme that it did not occur at all in the sea rise projections. The preference for extreme outcomes was reduced by using either a numerical or a dual verbal-numerical prediction. However, compared to the verbal format, the numerical and the dual format failed to reduce the frequency-probability gap. In fact, the numerical format actually increased the magnitude of the gap. The gap between probability and cumulative frequencies was also large, which suggests that participants did not rely on a cumulative reading of the distribution to choose their outcomes or to shape their probability perceptions. Yet, the format may have influenced the way participants read the distribution: the difference between cumulative frequency and probability perception was low in the verbal probability condition, which may indicate a cumulative frequency reading of the distribution. Finally, climate change belief, climate change mitigation support and pro-ecological attitude did not guide the selection of a particular outcome nor did it shape the gaps between frequencies and subjective probabilities.

Experiment 2 – Medium probability predictions

The IPCC guidelines do not include the probability term "possible", but this expression is used in the IPCC reports to describe climate change outcomes. For example, "possible" occurs 355 times in the climate change mitigation report (2013), whereas "unlikely" appears only 80 times. The term may refer to the notion of feasibility (e.g., "this report was made possible"). However, it is also used as a way to quantify probability as in "non-climate effects could include possible depletion of stratospheric ozone by stratospheric aerosol injections" (p. 61) or "it is possible that the reduction of the ice sheet to a much smaller extent would be irreversible" (IPCC, 2007, p. 152). The term "possible" is also often used to describe the likelihood of climate change outcomes by governmental agencies, in the media or by people in daily life. For example, in a report, the Department of Energy & Climate Change (2014) states that "it is possible that the ice sheet would not be able to regrow…" (p. 6).

Results and Discussion

Preference for extreme outcomes in medium probability predictions. When describing a "possible" sea rise, half of the participants selected an extreme outcome taken from the tails of the distribution (see Table 3). This means that half of the participants chose an outcome that occurred in 10% of the sea rise projections to form a prediction that typically conveys a 50% probability.

We compared the selection of extreme outcomes according to the prediction format using a chi square followed by Bonferroni adjusted pairwise comparisons (visible with subscript letters in Table 3). The analysis showed that format affected the choice of the outcome in the prediction completion, χ^2 (4, N = 98) = 19.12, p = .001, Cramer's V = .31. The preference for one of the extreme outcomes dropped from 47% in the verbal format condition to 6% in the numerical probability prediction and 3% in the dual verbal-numerical condition. Compared to the verbal condition, participants selected less extreme outcomes in the numerical and the dual conditions. Further, participants selected an extreme high outcome less often in the dual condition than in the numerical one.

A multinomial logistic regression tested the relationship between outcome extremeness and climate change belief, mitigation support and pro-ecological attitude. The model had a fair fit, χ^2 (188, N = 98) = 192.25, p = .401. Altogether, the variables in the model did not predict variance in outcome selection, supporting the fact that climate change belief, mitigation support and pro-ecological attitude did not guide the outcome selected to complete the prediction, χ^2 (6, N = 98) = 4.40, p = .623, Cox and Snell Pseudo R-square = .04.

The frequency-probability gap. On average, participants chose outcomes that had a 31.2% frequency but believed that the outcome had a 55.3% probability of occurring, hence over-estimating the probability of the outcome and showing an average frequency-probability gap of 24.1%. This is quite obvious in Figure 4, which shows the frequency and probability distributions, and in Table 5, which shows the average differences. The gap between the cumulative frequency of the outcome chosen and participants' probability perception was between -16 and +8. The gap between probability and cumulative frequency was smaller than the difference between probability and frequency for 39% of participants. This indirectly indicates that those participants may have relied more on cumulative frequencies than frequencies to form their probability perceptions.

Table 5

Mean frequency-probability and cumulative-frequency-probability gaps for medium probability predictions according to the probabilistic format along with pairwise comparisons according to format (V: verbal, N: numerical, and D: dual) (N = 98, *n* per cell between 30 and 35).

				Omnibus			Pairwise comparison Mdiff and 95% CI		
gaps	V	N	D	F (2, 97)	*p*	η_p^2	**N** *vs* V	**D** *vs* V	**D** *vs* N
freq-proba	32.4 (18.9)	22.0 (20.9)	17.8 (19.1)	6.11	.003	.12	-10.2 [-21.6, 1.2]	-16.1* [-27.4, -4.8]	-5.9 [-16.7, 4.9]
C-freq-proba	8.1 (29.5)	-9.8 (31.9)	-15.9 (24.2)	7.99	.001	.15	-18.8* [-35.7, -1.9]	-27.0* [-43.7, -10.3]	-8.3^{ns} [-24.2, 7.7].

Note. *: *p* < .05. Freq-proba : frequency-probability gap and C-freq-proba: cumulative-frequency-probability gap). Mse for freq-proba and C-freq-proba in the MANCOVA: 2002.00 and 5731.94

The format of the prediction had an effect on the frequency-probability gap (Table 5). Pairwise comparisons showed that compared to the verbal format, the frequency-probability gap was reduced in the dual-format condition. The frequency-probability gap was similar in the verbal condition compared with the numerical condition and in the dual condition compared with the dual one. The cumulative-frequency-probability gap was also determined by the format of the prediction (Table 5). The gap was wider in the numerical and dual-format condition compared to the verbal condition. The gap was similar between the numerical and the dual condition.

As in Experiment 1, climate change mitigation support and pro-ecological attitude were not related to the average magnitude of the two gap variables, F (2, 91) = 1.29, p = .280, η_p^2 = .03 and F (2, 91) < 1, η_p^2 < .01. However, climate change belief was related to outcome choice, F (2, 91) = 4.29, p = .017, η_p^2 = .09. Aligned with this finding, participants who believed more in climate change also exhibited bigger frequency-probability and cumulative-frequency-probability gaps: r = 0.39, p < .001 and r = .33, p = .001.

Half of the participants made a medium verbal probability prediction with an extreme outcome that occurred in only 10% of the projections. The preference for extreme outcomes was reduced by using either a numerical or dual verbal-numerical prediction. Further, the frequency-probability gap was reduced by the use of the dual format. The majority of the participants seem to have relied more on frequencies than cumulative frequencies to form their probability perceptions. The preference for those extreme outcomes was not predicted by climate change belief, mitigation support and pro-ecological attitude. The difference between probability and frequencies or cumulative frequencies was related to climate change belief but not mitigation support nor pro-ecological attitude.

Experiment 3

Given that the most frequent sea rise in the distribution presented to participants was only 40% frequent and based on the “at least” reading of quantities, we had a different hypothesis for the effect of the format compared to Experiments 1 and 2 (i.e., numerical and dual-format predictions would reduce the preference for extreme outcomes). We hypothesised that if the numerical and dual-format interventions re-enforced the cumulative reading of the sea rise distribution it would lead to an increased preference for extreme outcomes and a decrease in the cumulative-frequency-probability gap. In addition, this would also lead to an increase in the frequency-probability gap. In contrast, if the interventions re-enforced a frequency reading of the distribution, they should decrease the preference for extreme outcomes and reduce the gap between frequencies and probabilities.

Results and Discussion

Preference for extreme outcomes in very high probability predictions. To make a verbal high probability prediction, only 20% of the participants selected an extreme outcome (see Table 3). We compared the selection of extreme outcomes according to the prediction format as used in Experiments 1 and 2. Participants chose an extreme outcome twice more in the dual condition than in the verbal one, but the analysis showed no effect of the format, χ^2 (4, N = 88) = 6.22, p = .184, Cramer's V = .19.

A multinomial logistic regression analysis tested the relationship between extremeness of outcome and climate change belief, mitigation support and pro-ecological attitude. The model had a fair fit, χ^2 (166, N = 87) = 161.55, p =.583 and showed that climate change belief, mitigation support and pro-ecological attitude did not explain a significant proportion of variance in outcome selection, χ^2 (6, N = 87) = 7.89, p = .246, Cox and Snell Pseudo R-square = .09.

The frequency-probability gap. On average, participants chose outcomes that were 28% frequent but believed that these outcomes had a 72% probability of occurring, hence over-estimating the probability of those outcomes and showing an average frequency-probability gap of 44%. This is clearly visible in Figure 3, which shows the frequency and probability distributions, and in Table 4, which shows the average differences. The gap between the cumulative frequency of the outcome chosen and participants' probability perception is much smaller (between 0% and 6%). For 72% of participants, the difference between frequency and probability perception was larger than the difference between cumulative frequency and probability perception, indicating that participants may have relied more on cumulative frequencies rather than frequencies to form their probability perceptions.

We tested the effect of the format together with climate change belief on the difference between frequencies and probability judgments as in Experiments 1 and 2. Results (depicted in Table 6) showed no effect of the prediction format on the frequency-probability gap and on the cumulative-frequency-probability gap. Finally, the multivariate tests showed that climate change belief, climate change mitigation support and pro-ecological attitude were not related to the average magnitude of the two gaps, $F(2, 81) < 1$, $\eta_p^2 = .02$, $F(2, 81) = 2.54$, $p = .085$, $\eta_p^2 = .06$ and $F(2, 81) = 1.96$, $p = .147$, $\eta_p^2 = .05$.

Table 6

Mean frequency-probability and cumulative-frequency-probability gaps for low probability predictions according to the probabilistic format (V: verbal, N: numerical, and D: dual) (N = 98, *n* per cell between 25 and 32).

gaps	V	N	D	F (2, 92)	*p*	η_p^2
freq-proba	35.7 (23.8)	45.5 (32.0)	50.9 (29.2)	1.31	.275	.03
C-freq-proba	-6.3 (20.1)	-0.6 (30.8)	-6.3 (25.3)	0.80	.455	.02

Note. *: $p < .05$. Freq-proba: frequency-probability gap and C-freq-proba: cumulative-frequency-probability gap). Mse for freq-proba and C-freq-proba in the MANCOVA: 758.4 and 636.6.

In the present experiment, between 20 and 40% of the participants selected a rare and extreme outcome in order to make a high probability prediction. Despite choosing outcomes with a fairly low frequency (around 30%), participants felt that this outcome was actually likely (around 70%). Participants' subjective probabilities matched better with the cumulative frequency of the outcomes. Finally, individual differences regarding climate change did not impact outcome selection nor the frequency-probability gap.

Experiment 4

Experiments 1 and 2 showed that the numerical and the dual format reduced the rate of extreme outcomes for low and medium predictions in favour of outcomes that had a frequency better suited to conveying low and medium probabilities. Further, none of the formats consistently helped to reduce the gap between subjective probability and frequency or cumulative frequency. In the present experiment we had two aims. First, we aimed to test the robustness of the observed effects in a more controlled environment (i.e., in lab settings).

Second, we aimed to assess the effect of the probability conveyed by the prediction and whether this interacts with its format to determine outcome selection.

Method

Participants. A total of 89 psychology students from Kingston University, London, completed the web questionnaire in exchange for course credits. The sample included 79% females and participants had a mean age of 22.4 years (18-51, *SD* = 5.68). The sample was diverse in terms of ethnicity with 27% White British, 20% Black British, 18% White Other, 14% Asian British, 15% Asian/Other or Black/Other and 6% Other. Most students were in their first year of study (81%).

Design and materials. This experiment features a 3 (Probability conveyed: low, medium and high probability) × 3 (Format: numerical, verbal and dual) mixed design where the probability conveyed by the prediction was manipulated within-subjects and the format of the prediction was manipulated between-subjects. Participants completed an online questionnaire on individual computers in small groups overseen by two research assistants. Participants read the same sea rise vignette as in Experiments 1-3 but with three incomplete predictions conveying a low, medium and very high probability of occurrence. The order of presentation of the three predictions was randomised for each participant. The probabilistic format of the prediction was manipulated as in Experiments 1-3. Then, participants were shown their three predictions on three different web pages and rated the likelihood that they conveyed on a 0%-100% slider scale with increments of 1.

Results and Discussion

Preference for extreme outcomes. In the verbal condition, participants selected an extreme outcome about 2/3 of the time. Table 7 shows the proportion of selection of extreme outcome. This preference appeared to decrease when the probability increased: 90% of participants selected an extreme outcome for “unlikely”, 65% for “possible” and 47% for “virtually certain”. A Friedman test confirmed this trend, χ^2 (2, *N* = 89) = 28.79, *p* < .001. The numerical and the dual format decreased the selection of extreme outcomes for low and medium probability predictions but not for high probability predictions as indicated by three chi square analyses. The format of the prediction had an impact on the extremeness of the outcome selected by participants for low probability predictions, but not for medium and high probability predictions, χ^2 (4, *N* = 89) = 13.83, *p* = .008, Cramer’s *V* = .28, χ^2 (4, *N* = 89) = 7.80, *p* = .099, Cramer’s *V* = .21, χ^2 (4, *N* = 89) = 3.50, *p* = .477, Cramer’s *V* = .14. The pairwise comparisons showed that participants chose an extreme high outcome less often in the numerical condition. Despite the main effect of the format not being significant for medium probability predictions, the pairwise comparisons showed that the dual format reduced the selection of extreme outcomes (see subscript letter in Table 7).

The frequency-probability gap.

Participants globally and consistently over-estimated the probability of those outcomes and showed an average frequency-probability gap that ranged from 20% to more than 50%. This is plain to see in Figure 5, which shows that the frequency and the probability distributions do not overlap well. Table 8 shows the average differences between subjective probability and frequency or cumulative frequency (i.e., the gap variables).

Table 7

Outcome selected (in %) together with average frequency, cumulative frequency and subjective probability according to the format in Experiment 4 (N = 89).

Probability/Outcome	Prediction format: Verbal	Numerical	Dual
Low probability	*% selection*	*% selection*	*% selection*
Moderate	11% $_a$	33% $_a$	19% $_a$
Extreme low	21% $_a$	47% $_b$	32% $_{ab}$
Extreme high	68% $_a$	20% $_b$	48% $_{ab}$
Total extreme	89%	67%	80%
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Frequency	6.4 (11.0)	12.3 (11.9)	9.0 (10.8)
Cum. frequency	32.5 (42.7)	70.0 (38.6)	48.4 (45.8)
Subjective probability	31.1 (21.9)	33.1 (16.8)	29.7 (25.6)
Medium probability	*% selection*	*% selection*	*% selection*
Moderate	36% $_a$	63% $_{a\,b}$	68% $_b$
Extreme low	21% $_a$	13% $_a$	16% $_a$
Extreme high	43% $_a$	23% $_a$	16% $_a$
Total extreme	64%	36%	32%
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Frequency	15.7 (11.4)	18.0 (15.0)	18.8 (11.8)
Cum. frequency	20.4 (40.3)	62.3 (37.5)	60.7 (36.4)
Subjective probability	49.0 (21.3)	47.0 (19.6)	53.1 (18.6)
High probability	*% selection*	*% selection*	*% selection*
Moderate	54% $_a$	33% $_a$	55% $_a$
Extreme low	39% $_a$	57% $_a$	39% $_a$
Extreme high	7% $_a$	10% $_a$	7% $_a$
Total extreme	46%	67%	46%
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
Frequency	22.9 (15.8)	15.7 (15.9)	24.5 (15.7)
Cum. frequency	75.0 (28.1)	79.7 (31.0)	74.8 (26.9)
Subjective probability	70.5 (23.9)	71.3 (32.0)	78.1 (26.8)

Note: Superscript letter shows the pair of proportions that were statistically different at $p < .05$ with Bonferroni adjustment.

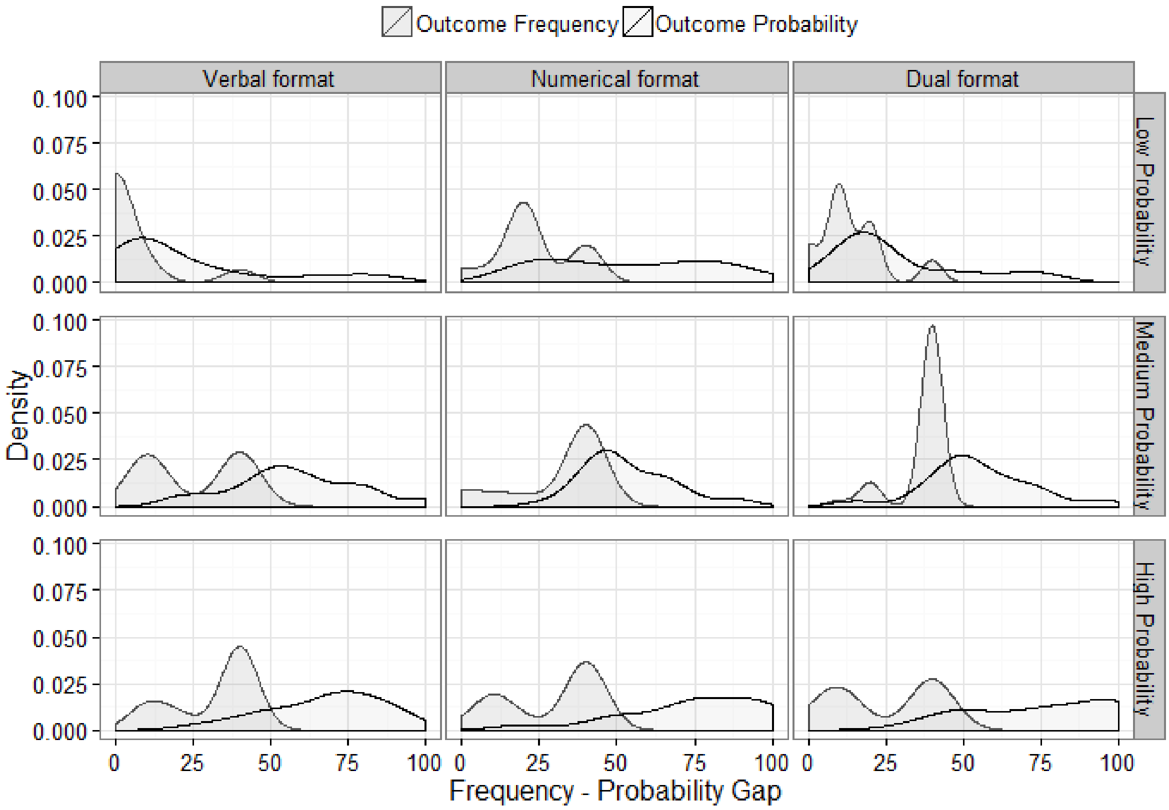


Figure 5. Distribution of the frequency (dark grey) and subjective probability (light grey) of the outcome selected by participants according to the format of the prediction for low, medium and high probability predictions found in Experiment 4 (with probability conveyed as a within-subjects independent variable).

The probability conveyed by the prediction had a positive effect on the magnitude of the frequency-probability gap: the gap was larger for high probability predictions. We observed the opposite pattern for the cumulative-frequency-probability gap: the gap was smaller for higher probability predictions. The format did not seem to affect the frequency-probability gap but seemed to affect the cumulative-frequency-probability gap which was larger for the numerical and the dual-format predictions than for the verbal probability predictions.

We tested the effect of the probability conveyed and the format of the prediction, along with their interaction, with a Generalized Estimating Equation (GEE) for ordinal variables with probability conveyed as a within-subject independent variable and format as a between-subjects independent variable for both gap variables, separately. The results showed that the probability conveyed had an effect on the magnitude of the frequency-probability gap but that the format did not, Wald χ^2 (2, N = 89) = 57.94, p < .001, Wald χ^2 (2, N = 89) = 0.07, p = .964. The probability conveyed and the format did not interact, Wald χ^2 (4, N = 89) = 1.78, p = .777.

Table 8

Mean outcome frequency, subjective probability and frequency-probability gap as a function of the probabilistic format of the prediction in Experiment 4 (verbal, numerical, and dual; n per cell varied between 28 and 31).

	Low probability			Medium probability			High probability		
Gaps	Verbal	Num.	Dual	Verbal	Num.	Dual	Verbal	Num.	Dual
Freq-proba	24.7 (23.7)	20.7 (21.8)	20.6 (23.7)	33.3 (27.6)	29.0 (21.1)	34.4 (22.7)	47.6 (23.6)	55. 7 (33.9)	53.6 (29.5)
C-freq-proba	-1.4 (42.5)	-36.9 (42.4)	-18.7 (56.7)	-1.3 (48.0)	-15.3 (40.1)	-7.6 (37.6)	-4.5 (36.1)	-8.3 (45.0)	3.3 (32.0)

Note. freq-proba (frequency-probability gap) and C-freq-proba (cumulative-frequency-probability gap).

For the cumulative-frequency-probability gap, the GEE results showed the probability conveyed had an effect on the magnitude of the frequency-probability gap but the format did not, Wald χ^2 (2, N = 89) = 106.78, p < .001, Wald χ^2 (2, N = 89) = 2.58, p = .275. However, we also observed that the probability conveyed and the format interacted to determine the cumulative-frequency-probability gap, Wald χ^2 (4, N = 89) = 10.22, p = .037. The interaction reflected the fact that the cumulative-frequency-probability gap varied more according to the format in the low probability condition than in the medium and high probability conditions (see bottom row of Table 8). We conducted three Kruskal-Wallis tests to investigate the effects of format within each probability magnitude. The results showed that the format did not have an effect on cumulative-frequency-probability for low, medium and high probability predictions (after adjusting for the p value threshold), Kruskal-Wallis χ^2 (2, N = 89) = 6.28, p = .043, χ^2 (2, N = 89) = 0.75, p = .688 and χ^2 (2, N = 89) = 1.50, p = .472.

The results of Experiment 4 replicated the pattern of outcome selection observed in Experiments 1, 2 and 3 in the verbal probability condition, with a preference for extreme outcomes for low and moderate probability predictions. The findings replicate the fact that the numerical format decreased the preference for extreme outcomes in the low probability condition and that the dual format decreased the selection of extreme outcomes for medium probability predictions. Despite results being in the expected direction, we did not replicate the finding that the numerical format reduced the selection of extreme outcomes for medium probability predictions and that the dual format reduced the preference for extreme outcomes for low probability predictions. Further, the experiment replicated the fact that the format did not affect the frequency-probability gap for low and high probability predictions but did not replicate the gap reduction in medium probability predictions.

General Discussion

In the four experiments presented here, we found that participants exhibited a preference for predicting rare and extreme sea rise outcomes for low and medium verbal probability predictions, but not for high probability predictions. Participants' beliefs in climate change, pro-ecological attitudes and support for climate change mitigation did not lead to a greater preference for predicting extremely large climate change outcomes. We have also shown that participants over-estimated the probability of the outcomes they predicted, as shown by a wide frequency-probability gap. Despite knowing the frequency of the outcome they predicted, participants largely over-estimated its probability of occurrence. Participants' probability perceptions also departed substantially from the outcomes' cumulative frequencies for low and medium probability predictions. However, participants' probability perceptions were fairly close to the outcome cumulative frequency for very high probabilities, indicating that participants may have relied on a cumulative reading of the distribution to make their predictions. Finally, the format-based interventions only brought some limited benefits. The numerical format reduced the preference for extreme outcomes for low probability predictions (Experiments 1 and 4) and the dual format reduced the selection of extreme outcomes for medium probability predictions (Experiments 2 and 4). However, none of the formats consistently reduced the frequency-probability gaps.

Preference for predicting extreme outcomes

Unlikely and Possible outcomes. Participants consistently exhibited a strong preference for predicting extreme outcomes for low and moderate probabilities in the verbal format conditions. Predicting an extreme outcome may be perceived as valuable by recipients: extreme predictions are judged more positively than timid predictions, even when the extreme predictions are wrong (McKenzie & Amin, 2002). The extreme outcomes selected for "unlikely" and "possible" were taken from the upper end of the distribution and can be considered to stem from the informativeness of the most extreme outcome. The selection of maximal outcomes for "possible" also correspond to their frequent matching with an upper interval single bound marker (Teigen et al., 2014). When prompted to complete a "possible" prediction with an outcome and a modifier of their choice (e.g., "at least", "up to", "around" or "exactly") 75% of the participants selected the maximal outcome and the modifier "up to".

Virtually certain. The selection of the minimal outcome for "virtually certain" predictions may have stemmed from a cumulative reading of the frequency distribution whereby the minimal continuous outcome is indeed certain. The gap between the frequency and the probability of the outcome was quite large for the high probability predictions. However, when we examined the difference between the subjective probability and cumulative frequencies of the outcomes, the gap was much smaller, indicating that participants may have relied on a cumulative reading of the frequencies to choose their outcome and assess its probability of occurring. This is consistent with findings that participants associated lower bound markers such as "at least" in "certain" predictions (Juanchich et al., 2013). This finding also explains why in past research many participants selected the minimal outcome to predict what "will" happen (Juanchich et al., 2013; Teigen & Filkuková, 2013; Teigen et al., 2014).

Minimum and maximum outcome predictions seen as single bound interval predictions. As described above, for "certain" predictions, minimal values may have been chosen with an implicit "at least" in mind, and for "possible", maximal values may have been chosen with an implicit "up to" in mind. When "certain" predictions of minimal outcomes are considered to predict the lower bound of an interval, they should be considered to be 100% likely. In this case, "certain" hints at a minimal sea level rise that could happen and entails the possibility of larger sea rise magnitudes. Following the same reasoning, when "certain" predictions of maximal outcomes are considered to predict the upper bound of an interval, they should also be considered to be 100% likely. In this case, "possible" hints at a maximum sea rise, and entails the possibility of the occurrence of a smaller sea level rise. For example, saying that the sea level will rise "up to" 200 m in the next five years is 100% sure, given that any sea level rise that will happen in the next few years will fall below this upper bound. In line with the "at least" cumulative reading of "certain", minimal outcomes were judged to be "very likely". However, in contrast with the "up to" reading of "possible" predictions, participants did not believe that maximal outcomes had a high chance of representing the maximal bound of the interval; possible outcomes were believed to be about 50% likely.

The prediction of implicit single bound intervals could also mean trouble for recipients of "certain" and "possible" predictions. Given the implicit nature of the position of the outcome in the outcome distribution, recipients may not be aware that the outcomes serve as the upper or lower bounds of an interval being predicted. Instead, recipients could consider that the phrase indicates a degree of certainty regarding a precise outcome, hence creating a misunderstanding between the speaker and the recipient. To reduce the risk of miscommunication, climate change communicators should specify the position of the outcome they are describing in the outcome distribution: is it an exact outcome, a minimal one or a maximal one? To do so, they could use appropriate linguistic markers (e.g., "it is likely that the sea will rise *at least* 10 inches").

Contrast with past findings. For the high probability predictions, participants' selection of extreme outcomes in our experiments was not as gloomy as in previous experiments. In previous research more than half of the participants chose an extreme outcome as being "certain" (Juanchich et al., 2014; Teigen et al., 2013, Teigen et al., 2014), whereas in our experiments only 20% did so for the verbal probability "virtually certain". The difference could be explained by the effect of the relevance of the context. Previous studies were conducted on hypothetical events that may not have elicited a great deal of motivation to formulate a prediction, whereas climate change is considered to be an important societal issue (albeit not enough). The design also seems to account for some variations in the preference for extreme outcomes. When predictions conveying different probabilities were presented in a within-subjects design (Experiment 4), the preference for extreme outcomes went up by 15% for "virtually certain" compared to results from experiments in which participants made a single prediction. It is important to note that the difference may have been driven by a difference in the words used. We used the expression "virtually certain" to characterise a sea rise, whereas previous research focused on "certain" (e.g., 60% in Teigen et al., 2014, Study 1; 80% in Juanchich et al., 2013, Study 1). Based on the linguistic maxim of informativeness within the principle of cooperation (Grice, 1975), participants may have seen the adverb "virtually" as a marker that the outcome was *not* entirely certain. This marker could have deterred participants from choosing the minimal outcome which was indeed 100% certain according to a cumulative reading of the frequencies.

Further, the small selection rate of the minimal outcome partly contrasts with the literature on quantities, which posits that quantities are usually minimal quantities to be expected. The lower-bounding of quantities may not be as frequent in speakers as it is inferred by recipients. This finding is in line with the view of Geurts (2006) that quantities can simply be given different interpretations ("at least", "exactly", "at most") according to the context. However, here we also see that, in the same context, different interpretations can be selected by different speakers. Individual differences, such as numeracy, may account for the different preferences of speakers.

Role of individual differences

Participants who believed more strongly in climate change did not choose to predict larger climate change outcomes. This appears to be in contrast with findings showing that participants who believed more strongly in climate change associated higher probabilities with climate change predictions than people who believed less strongly in climate change (Budescu, et al., 2009; Budescu, et al., 2012). This indirectly indicates that, when provided with climate change projection data, the choice of an outcome to formulate a prediction is less sensitive to individual differences than probability perception given a prediction. The way speakers form predictions may be more anchored to implicit conversational rules and therefore less influenced by motivated reasoning than probability perception. In our studies, participants did not have a particular communication goal; they simply used some data to form a prediction. A change in the speaker's intention could affect the relationship between outcome predicted and personal belief. For example, if participants formulated a prediction to warn recipients about the negative potential effects of climate change, we could expect that personal belief would have a stronger effect.

The preference for extreme outcomes and the gap between frequencies and subjective probabilities may stem from other individual differences, such as graph literacy. Indeed, research shows that people tend to find graphical representations of uncertain quantities hard to interpret (Ibrekk & Morgan, 1987). This is unlikely to be the main factor explaining the preference for extreme outcomes, given that similar findings were found in vignettes where the frequencies were presented in a table or a narrative (Teigen et al., 2014; Teigen et al., 2013). However, future research could assess the respective impact of the ability to understand frequencies and graphics on the selection of extreme outcomes using measures of numeracy (e.g., Lipkus, 2007) and graphical literacy (e.g., Okan, Garcia-Retamero, Cokely, & Maldonado, 2011).

Probability over-estimation

A good strategy to predict the probability of occurrence of an outcome is to rely on how often it has occurred in the past. An event that is 20% frequent should be predicted to be 20% likely and should be associated with a 20% probability. However, in our studies we consistently found that there was an extensive gap between the frequency of the predicted outcome and its subjective probability.

Participants selected outcomes that had a low frequency to convey low, medium and high probability predictions and then believed that those outcomes had a low, medium or high probability perception. Hence, there was a gap between the actual probability that an outcome would occur (based on its frequency in the projections) and participants' perception of the probability that this outcome would occur. Participants largely over-estimated the probability of occurrence of an outcome after predicting it was "unlikely", "possible" or "virtually certain". Our data indicate that speakers making a prediction seem to be influenced by the probability they convey in their predictions rather than by the frequentist evidence provided in the sea rise projections. For example, when participants predict that "it is possible that the sea level will rise 18 inches", they then believe that this sea rise has a 50% chance of occurring, despite the fact that it occurred in only 10% of the projections. This finding replicates the results of Teigen et al. (2013) who found the frequency-probability gap with "unlikely" predictions. We have extended the finding to the climate change context and to medium and high probability predictions.

Furthermore, the present results indirectly highlight a miscommunication issue between speakers and recipients. We have shown that people made the "unlikely", "possible" and "virtually certain" predictions with outcomes that occurred in only 5%, 25% and 32% of the projections (Experiments 1-3) whereas those phrases are typically understood to mean 20%, 50% and 90% (Harris, Corner, Xu, & Du, 2013; Juanchich et al., 2012).

Prediction format can help decrease the selection of extreme outcomes.

We tested two cost-effective interventions to reduce the preference for extreme outcomes in predictions. The dual format helped participants to choose outcomes that were less extreme for medium probability predictions (Experiments 2 and 4) and the numerical format helped participants choose less extreme outcomes for low probability predictions

(Experiments 1 and 4). However, in contrast with our expectations, the use of a numerical and dual format did not consistently reduce the frequency-probability gap. The only two effects of the format on the frequency-probability gap found in Experiments 1 and 2 were not replicated in Experiment 4. The frequency-probability gap therefore proves hard to eliminate. The change of format did not help participants to resolve their internal "frequency-probability" gap, whether they relied on frequencies or cumulative frequencies. The fact that the dual format can help to reduce the selection of rare and extreme outcomes complements previous research showing that the dual format helps recipients to interpret the meaning of predictions more in line with the IPCC guidelines (e.g., Budescu et al., 2009; Budescu et al., 2012; Budescu et al., 2014).

A "possible" recommendation

The verbal probability "possible" is not part of the recommended vocabulary to convey probabilities in the IPCC guidelines on uncertainty communication. Nevertheless, "possible" is quite often used in the latest IPCC report (IPCC, 2013) with 319 occurrences, some of which aim to convey a moderate degree of certainty (e.g., "[it] may also lessen uncertainty in the assessment of possible and probable impacts", IPCC 2013, p. 138). We argue that there is a case to discourage the use of this term because it is associated with maximal outcomes. We recommend that the IPCC guides authors to limit the use of "possible" by using the recommended term "as likely as not" instead.

Conclusion

Our experiments provide, to our knowledge, the first empirical evidence on how people form climate change predictions based on climate change projection data. The present paper makes a theoretical and applied contribution. The theoretical contribution is to provide evidence that people form climate change predictions that are consistent with a dispositional conception of probabilities and less consistent with a frequentist conception of probabilities. The applied contribution is to show that people have a preference for predicting extreme climate change outcomes and that, although people are aware those outcomes are rare, they tend to over-estimate their probability of occurrence. We have complemented existing research mainly concentrating on the way that people understand predictions given to them by focusing on the way that people predict climate change.

References

Breheny, R. (2007). A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics*, *25*, 93–139. doi: http://dx.doi.org/10.1093/jos/ffm016

Budescu, D. V., Broomell, S. B., & Por, H. H. (2009). Improving communication of uncertainty in the ipcc reports. *Psychological Science, 20,* 299-308. doi: http://dx.doi.org/10.1111/j.1467-9280.2009.02284.x

Budescu, D. V., Por, H.-H., & Broomell, S. B. (2012). Effective communication of uncertainty in the ipcc reports. *Climatic Change, 113*, 181-200. doi: 10.1007/s10584-011-0330-3

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,* 508-512. doi: 10.1038/nclimate2194

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

Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., et al. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters, 8,* 1-7. doi: http://dx.doi.org/10.1088/1748-9326/8/2/024024

Corner, A., Whitmarsh, L., & Xenias, D. (2012). Uncertainty, scepticism and attitudes towards climate change: Biased assimilation and attitude polarisation. *Climatic Change, 114*, 463-478. doi: 10.1007/s10584-012-0424-6

Department of Energy & Climate Change. (2014). Climate change and energy – guidance. Climate change explained: United Kingdom Government.

Dunlap, R. E., Van Liere, K. D., Mertig, A. G., & Jones, R. E. (2000). Measuring endorsement of the new ecological paradigm: A revised nep scale. *Journal of Social Issues, 56*, 425–442-425–442.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology, 38*, 889-906.

Geurts, B. (2006). Take ‘five’: The meaning and use of a number word. In S. Vogeleer & L. Tasmowski (Eds.), *Non-definiteness and plurality* (pp. 311-329). Amsterdam/Philadelphia.: Benjamins.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), (Vol. 3): New York: Academic Press.

Hallegatte, S., Rogelj, J., Allen, M., Clarke, L., Edenhofer, O., Field, C. B., et al. (2016). Mapping the climate change challenge. *Nature Clim. Change, 6*, 663-668. doi: 10.1038/nclimate3057

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*, 1571-1578. doi: http://dx.doi.org/10.1037/a0024195

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*, 415–425-415–425. doi: http://dx.doi.org/10.1007/s10584-013-0975-1

Heath, Y., & Gifford, R. (2006). Free-market ideology and environmental degradation. *Environment and Behavior, 38*, 48-71. doi: http://dx.doi.org/10.1177/0013916505277998

Hidi, S. (2001). Interest, reading, and learning: Theoretical and practical considerations. *Educational Psychology Review, 13*, 191-209.

Horn, L. R. (1984). A new taxonomy for pragmatic inference: Q-based and r-based implicature. In D. Schiffrin (Ed.), *Meaning, form and use in context* (pp. 11-42). Washington: Georgetown University Press.

Ibrekk, H., & Morgan, M. G. (1987). Graphical communication of uncertain quantities to nontechnical people. *Risk analysis, 7*. doi: 10.1111/j.1539-6924.1987.tb00488.x

Intergovernmental Panel on Climate Change. (2007). Climate change 2007. The climate change physical science basis.

Intergovernmental Panel on Climate Change. (2013a). Climate change 2013. The physical science basis. Working group i contribution to the fifth assessment report of the intergovernmental panel on climate change. New York: United States of America.

Intergovernmental Panel on Climate Change. (2013b). Fifth assessment report climate change 2013: The physical science basis. Summary for policymakers.

Intergovernmental Panel on Climate Change. (2014). Climate change 2014 mitigation of climate change working group iii contribution to the fifth assessment report of the intergovernmental panel on climate change.

Juanchich, M., Sirota, M., & Butler, C. L. (2012). The perceived functions of linguistic risk quantifiers and their effect on risk, negativity perception and decision making. *Organizational Behavior and Human Decision Processes, 118*, 72-81. doi: http://dx.doi.org/10.1016/j.obhdp.2012.01.002

Juanchich, M., Teigen, K. H., & Gourdon, A. (2013). Top scores are possible, bottom scores are certain (and middle scores are not worth mentioning): A pragmatic view of verbal probabilities. *Judgment and Decision Making, 8*, 345-364.

Kahan, D. M., Jenkins-Smith, H., & Braman, D. (2011). Cultural cognition of scientific consensus. *Journal of Risk Research, 14*, 147-174.

Keren, G., & Teigen, K. H. (2001). The probability - outcome correspondence principle: A dispositional view of the interpretation of probability statements. *Memory and Cognition, 29*, 1010-1021.

Leiserowitz, A., Maibach, E. W., Roser-Renouf, C., Feinberg, G., & Howe, P. (2013). Climate change in the american mind: Americans' global warming beliefs and attitudes in april 2013. *Available at SSRN 2298705*.

Lichtenstein, S., Fischhoff, B., & Phillips, L. D. (1982). Calibration of probabilities: The state of the art to 1980. In D. Kahneman & A. Tversky (Eds.), (pp. 306-334): Cambridge: Cambridge University Press.

Lipkus, I. M. (2007). Numeric, verbal, and visual formats of conveying health risks: Suggested best practices and future recommendations. *Medical Decision Making, 27*, 696-713. doi: http://dx.doi.org/10.1177/0272989X07307271

McKenzie, C. R., & Mikkelsen, L. A. (2000). The psychological side of hempel's paradox of confirmation. *Psychonomic Bulletin & Review, 7*, 360-366.

McKenzie, C. R. M., & Amin, M. B. (2002). When wrong predictions provide more support than right ones. *Psychonomic Bulletin & Review, 9*, 821-828. doi: 10.3758/bf03196341

Okan, Y., Garcia-Retamero, R., Cokely, E. T., & Maldonado, A. (2011). Individual differences in graph literacy: Overcoming denominator neglect in risk comprehension. *Journal of Behavioral Decision Making, online publication*. doi: http://dx.doi.org/10.1002/bdm.751

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*, 867-872. doi: http://dx.doi.org/10.1016/j.jesp.2009.03.009

Popper, K. (1959). The propensity interpretation of probability. *British Journal for the Philosophy of Science, 10*, 25-42.

Popper, K. (1990). *A world of propensities: Two new views of causality*: Bristol: Thoemmes.

Read, D., Bostrom, A., Morgan, M. G., Fischhoff, B., & Smuts, T. (1994). What do people know about global climate change? 1. Mental models. *Risk Analysis, 14*, 959–970-959–970. doi: http://dx.doi.org/DOI : 10.1111/j.1539-6924.1994.tb00065.x

Reagan, R. T., Mosteller, F., & Youtz, C. (1989). Quantitative meanings of verbal probability expressions. *Journal of Applied Psychology, 74*, 433-442.

Reynolds, T. W., Bostrom, A., Read, D., & Morgan, M. G. (2010). Now what do people know about global climate change? Survey studies of educated laypeople. *Risk Analysis, published online*. doi: http://dx.doi.org/DOI : 10.1111/j.1539-6924.2010.01448.x

Teigen, K. H., & Filkuková, P. (2013). Can > will: Predictions of what can happen are extreme, but believed to be probable. *Journal of Behavioral Decision Making, 26*, 68-78. doi: 10.1002/bdm.761

Teigen, K. H., Juanchich, M., & Filkuková, P. (2014). Verbal probabilities: An alternative approach. *Quarterly Journal of Experimental Psychology*. doi: http://dx.doi.org/10.1080/17470218.2013.793731

Teigen, K. H., Juanchich, M., & Riege, A. (2013). Improbable outcomes: Infrequent or extraordinary? *Cognition, 127*, 119-139. doi: http://dx.doi.org/10.1016/j.cognition.2012.12.005

Windschitl, P. D., & Wells, G. L. (1996). Measuring psychological uncertainty: Verbal versus numeric methods. *Journal of Experimental Psychology-Applied, 2,* 343-364. doi: http://dx.doi.org/10.1037//1076-898X.2.4.343

Windschitl, P. D., & Wells, G. L. (1998). The alternative-outcomes effect. *Journal of Personality and Social Psychology, 75,* 1411-1423. doi: http://dx.doi.org/10.1037/0022-3514.75.6.1411

Witteman, C. L. M., & Renooij, S. (2003). Evaluation of a verbal-numerical probability scale. *International Journal of Approximate Reasoning, 33*, 117-131. doi: http://dx.doi.org/10.1016/S0888-613X(02)00151-2

Yates, J. F., Zhu, Y., Ronis, D. L., Wang, D. F., Shinotsuka, H., & Toda, M. (1989). Probability judgment accuracy: China, japan and the united states. *Organization Behavior and Human Decision Processes, 43,* 145-171.