Climatic Change (2012) 113:181–200
DOI 10.1007/s10584-011-0330-3

Effective communication of uncertainty in the IPCC reports

David V. Budescu · Han-Hui Por · Stephen B. Broomell

Received: 21 June 2010 / Accepted: 19 October 2011 / Published online: 23 November 2011

Abstract The Intergovernmental Panel on Climate Change (IPCC) publishes periodical assessment reports informing policymakers and the public on issues relevant to the understanding of human induced climate change. The IPCC uses a set of 7 verbal descriptions of uncertainty, such as *unlikely* and *very likely* to convey the underlying imprecision of its forecasts and conclusions. We report results of an experiment comparing the effectiveness of communication using these words and their numerical counterparts. We show that the public consistently misinterprets the probabilistic statements in the IPCC report in a regressive fashion, and that there are large individual differences in the interpretation of these statements, which are associated with the respondents' ideology and their views and beliefs about climate change issues. Most importantly our results suggest that using a dual (verbal—numerical) scale would be superior to the current mode of communication as it (a) increases the level of differentiation between the various terms, (b) increases the consistency of interpretation of these terms, and (c) increases the level of consistency with the IPCC guidelines. Most importantly, these positive effects are independent of the respondents' ideological and environmental views.

1 Introduction

The Intergovernmental Panel on Climate Change (IPCC) was established by the World Meteorological Organization and the United Nations Environmental Programme to assemble and disseminate information about global climate change (GCC). The panel

Electronic supplementary material The online version of this article (doi:10.1007/s10584-011-0330-3) contains supplementary material, which is available to authorized users.

D. V. Budescu (✉) · H.-H. Por
Department of Psychology, Fordham University, 441 East Fordham Road, Bronx, NY 10458, USA
e-mail: budescu@fordham.edu

S. B. Broomell
Department of Social and Decision Sciences, Carnegie Mellon University, Pittsburgh, PA 15213, USA

S. B. Broomell
College of Information Sciences and Technology, Pennsylvania State University, University Park, PA 16802, USA

compiles periodical Assessment Reports that inform policymakers and the public on issues relevant to the understanding of human induced GCC. The panel's efforts were recognized by the 2007 Nobel Prize for Peace committee which commended it for "lay(ing) the foundations for the measures that are needed to counteract such change" (Nobelprize.org 2008).

1.1 Communication of uncertainty in the IPCC reports

One important issue facing the writers of the IPCC reports is to communicate the inherent uncertainties in climate measurements and models to the public. This challenge applies to all risk communications, but the debates surrounding GCC have reached a much higher level of intensity. A variety of questions—Is GCC happening? How severe is it? Is it driven by human activities? How best to address the problems it causes?—come up frequently in public and political debates. Some of the debates are related to uncertainties that are inherent in climate science. Others reflect the public's imperfect understanding of climate-related issues and its misperceptions about the scientific consensus on the topic (Pew Center 2009).

A key challenge facing the IPCC is to convey information with the level of precision warranted by the available data and evidence (Budescu and Wallsten 1987). Using precise (numerical) probabilities to convey predictions in climate change models could be misleading as it would imply too high a level of precision of the estimates, and a high consensus among experts. In recent assessments the IPCC decided to use verbal descriptions of uncertainty such as *very unlikely* and *likely* to convey the underlying imprecision (see Ha-Duong et al. 2007 for a history).

Verbal descriptions of uncertainty reduce the unwarranted excessive precision inherent in numerical probabilities but they induce other possible problems. Empirical research by psychologists and decision theorists has documented large individual differences in the ways people understand, communicate and use these terms as well as their sensitivity to various factors (see review by Wallsten and Budescu 1995). To illustrate the challenge of using verbal descriptions to convey uncertainty, we will first review briefly some of the findings and the principles identified by this line of research and then summarize an earlier study where we examined the understanding of the probabilistic words as used by IPCC in one community.

The representation and interpretation of probability terms depend on context Context effects on the interpretation of probability terms are pervasive. For example, Fischer and Jungermann (1996), Pepper and Prytulak (1974), Wallsten et al. (1986), and Weber and Hilton (1990) have shown that the numerical interpretation of a term used to describe the chances of a given event occurring correlates positively with the event's perceived base rate. Furthermore, Weber and Hilton (1990) found that mean probability interpretations of phrases were positively correlated with event severity. Patt and Schrag (2003) asked respondents to assign verbal probability labels to events where the numerical likelihoods of occurrence were given. They found that subjects were more likely to use descriptors implying higher likelihood (e.g., *very likely* as opposed to *exceptionally unlikely*) to describe a hurricane than to describe snow flurries even when the numerical likelihood of occurrence for both events were the same, effectively demonstrating the role of context in interpretation of probabilistic terms.

The interpretation of probability terms depends on role of the individual in the communication process Budescu and Wallsten (1990), and Fillenbaum et al. (1991)

observed that recipients of the verbal forecasts consistently interpreted probability phrases as being less extreme and more imprecise than intended by the communicators.

Communication mode choices are sensitive to the degrees of vagueness inherent in the events being described, the source of the uncertainty and the nature of the communication task In general, people prefer to communicate their opinions verbally under conditions of imprecision (e.g. Brun and Teigen 1988; Erev and Cohen 1990; Wallsten et al. 1993). Verbal communication is preferred unless the underlying opinions are based on solid quantitative evidence about aleatory events (e.g., Olson and Budescu 1997), or there is some clear incentive to be precise (Erev et al. 1991). On the other hand, most people prefer receiving precise numerical information (Brun and Teigen 1988; Erev and Cohen 1990; Fischer and Jungermann 1996; Wallsten et al. 1993). In fact, the modal pattern of responses in the Wallsten et al. (1993) survey and the Erev and Cohen (1990) study, was preference for communicating to others verbally *and* receiving information from others numerically. People prefer to communicate verbally because this modality is more natural, more personal and easier, and prefer to receive numerical information because it is more precise (Wallsten et al. 1993).

Probability terms are susceptible to self-serving interpretations Fox and Irwin (1998) emphasized the role of the social and motivational context and noted that emotions and political implications can induce contextual effects. Recent work by Piercey (2009) in the context of accounting auditing demonstrates how probability words can be used effectively to reach and justify one's preferred conclusions.

The verbal lexicons vary drastically across individuals The use of verbal terms can induce an 'illusion of communication' (Budescu and Wallsten 1985) that stems from the (natural, but false) assumption that everyone interprets the terms similarly. Over their lifetime, people develop preferences for specific terms and tend to avoid others. Consequently, when they need to choose terms to describe uncertainty, different individuals will spontaneously pick different words. For example, Budescu et al. (1988) report that the 20 participants in their study produced 111 distinct probability phrases when asked to describe 11 different, graphically displayed, probabilities (see also Dhami and Wallsten 2005; Erev and Cohen 1990; Karelitz and Budescu 2004; and Zwick and Wallsten 1989).

There is large inter-personal variability in the interpretation of verbal probabilities There is a long line of studies mapping phrases into numerical probabilities. Participants were asked to rank order, compare or simply convert phrases into numbers, and vice versa, in various ways and under various contexts (see Wallsten and Budescu 1995, for a partial list of these studies). The most robust findings are that between-individuals variability is very high and that it is considerably larger than the variance observed within-individuals, when judging the same terms (Beyth-Marom 1982; Budescu and Wallsten 1985; Clarke et al. 1992; Johnson 1973; Mullet and Rivet 1991 and Reagan et al. 1989). This suggests that most people perceive the meaning of verbal probabilities consistently and reliably, but vary differently from each other. Numerous studies found considerable interpersonal variability in interpreting probability phrases among experts operating within their professional domains such as military intelligence (e.g., Beyth-Marom 1982; Johnson 1973), accounting (e.g., Chesley 1985), and especially medicine (e.g., Bryant and Norman 1980; Kong et al. 1986; Mapes 1979; Merz et al. 1991; Nakao and Axelrod 1983; Sutherland et al. 1991). Many of these studies caution against the use of probability phrases in their respective domains.

It is difficult to "standardize" the meaning of verbal probability terms Researchers (e.g., Beyth-Marom 1982; Hamm 1991; Mosteller and Youtz 1990) have suggested using a fixed list of terms to reduce errors in communication of uncertainty. The major drawback of such scales is the difficulty of most people to suppress the meanings they normally associate with these terms. For example, Wallsten et al. (1986) demonstrated that National Weather Service forecasters could not transfer the imposed meanings of some words from one domain (e.g. forecasting rain) to another (e.g. interpreting medical advice). These forecasters are trained to use a fixed set of phrases to communicate their estimates for meteorological events. However, when the same phrases were presented in a different domain, the forecasters reverted to their natural (colloquial) meaning indicating that it is extremely difficult to "legislate meaning". Another problem is that, as we mentioned earlier, the strong context effects can undermine the validity of such scales.

Despite the inherent subjectivity, large inter-personal variability and the difficulty of legislating language, this is the solution that was adopted by the IPCC to communicate uncertainty. Authors of the IPCC reports are instructed to use a scale with seven verbal terms to convey uncertainties (see Table 1 from IPCC 2005). Even if all the authors comply with these instructions the results cited earlier should lead one to question whether policy makers and the general public—the target readership of the reports—understand these words as intended and form correct impressions. This is a universal concern in risk communication but it is more prominent in the GCC context because some critics have used the uncertainty underlying climate science as a convenient excuse to dismiss the findings all together (e.g., Begley 2007).

Budescu et al. (2009) ran several experiments to study whether readers understand the words as intended by the IPCC authors. Participants read 13 sentences containing probabilistic terms extracted from the IPCC report and provided their (a) best estimate of the probability intended by the report's authors, (b) the lowest and (c) the highest possible values they thought were consistent with the authors' intentions. The consistency between readers and authors (as measured by the degree to which the participants' judgments of (a) the best estimates and (b) ranges of possible values matched the relations outlined in Table 1) was low and the variability in the readers' interpretations far exceeded the uncertainty in the predictions.

Access to the translation table (Table 1) improved consistency slightly, but the addition of numerical probability ranges to the verbal probabilistic terms in the statements themselves was more effective in improving consistency. Furthermore, participants were sensitive to the width of the numerical probability ranges: the responses range was wider for participants presented with the wide ranges than for participants presented with the narrow ranges. This suggests that the respondents paid attention to both the numerical and the verbal information and that effectiveness of the uncertainty communication in the IPCC

Table 1 IPCC guidelines for translation of probabilities into words

Phrase	Likelihood of occurrence/outcome
Virtually certain	>99%
Very likely	>90%
Likely	>66%
About as likely as not	33% to 66%
Unlikely	<33%
Very unlikely	<10%
Exceptionally unlikely	<1%

reports can be improved significantly by relatively minor adjustments of the communication format.

An intriguing possibility is that interpretation of the IPCC pronouncements is related to systematic individual differences between the respondents. We distinguish between several sources of differences: Numeracy, attitudes and beliefs. Reyna and Brainerd (2007) claim that numeracy is a separate construct from education or general intelligence. Numeracy is analogous to literacy in that both are necessary for making judgments and decisions in everyday life. Their review suggests that highly numerate individuals can assess health-related information more accurately and as a result, make better health related decisions and Peters et al. (2006) show that highly numerate individuals are more likely to apply appropriate numerical principles and are less susceptible to framing effects, and tend to draw stronger and more precise affective meaning from numerical information. The variance in numeracy in the population may account for the superiority of the joint verbal-numerical presentation in the sense that highly numerate people benefit from such a presentation.

The other potential source of differences is selective attention to, and interpretation of, evidence. Nickerson (1998) refers to the patterns of search, evaluation, weighting and interpretation of evidence that are systematically biased in the direction of one's expectations, hypotheses and beliefs, collectively, as the *confirmation bias*. This tendency is unconscious and non-deliberate. Some researchers (e.g., Klayman and Ha 1987; Nickerson 1998), stress the cognitive source of these processes, but an alternative view emphasizes their motivational nature (Kunda 1990). Several manifestations of these tendencies are directly relevant to the interpretation of the IPCC statements (see also Piercey 2009): (a) People tend to seek and retrieve information they consider supportive of their expectations; (b) They overweight information that is supportive of their beliefs, compared to contradictory evidence; (c) People seek and overweight positive confirmatory cases; (d) They tend to be more accepting (critical) of research methods and results that are in line with (contradict) their hypotheses. For example, a recent paper by Bastardi et al. (2011) suggests that people interpret seemingly ambiguous scientific evidence in a manner that is congruent with their desires. The confirmation tendency is not applied indiscriminately, but only in some cases, and in moderation. Kunda's (1990) explanation for this apparent self-restraint is that most people attempt to find rational arguments and justifications that would appeal to dispassionate objective observers. Thus, the phenomenon would be most likely to be found when people can maintain an "illusion of objectivity" (Pyszczynski and Greenberg 1987). In our case, respondents might interpret the probabilistic terms in the statements extracted from the IPCC reports in line with their beliefs about GCC: people who, based on their ideology, political orientation and other background variables (e.g. Leiserowitz et al. 2011), tend to believe in GCC and its dangers would tend to assign more extreme probabilities to various GCC related events than others who, by and large, do not believe in GCC.

1.2 The present study

The participants in the Budescu et al. (2009) studies were volunteers from the University of Illinois community, and 60% of them were students. In an effort to extend and generalize the conclusions of their study we replicated portions of the experiment using a large random sample representative of the US population. In addition to asking the respondents to interpret probabilistic sentences from the IPCC we also measured the respondents' attitudes towards, and beliefs about, GCC and collected extensive demographic information. These covariates would allow us to test Kunda's motivated reasoning hypothesis. Budescu et al. (2009) did not find strong systematic relationships between ideological and demographic factors and the interpretation of the terms but these potential correlates could prove to be more relevant in a more diverse sample. We also used a short numeracy scale to investigate the relation between numeracy and interpretation of scientific evidence in climate change. We hypothesize that the main results from Budescu et al. (2009) will be replicated, namely:

(a) The public interprets the probabilistic terms in the IPCC reports in a regressive manner, i.e., underestimate high probabilities and overestimate low probabilities; and
(b) The level of consistency is higher when the probability words are accompanied by numerical ranges.

In addition, we will be able to determine if

(c) The interpretation of the probability words is correlated with the respondents' attitudes toward GCC and other ideological variables;
(d) The benefits of the new dual (verbal—numerical) scale persist after taking the differences in attitudes toward GCC into account; and
(d) The benefits of the dual scale are more pronounced for highly numerate respondents.

2 Method

The sample The study was conducted using the online Time-sharing Experiments for the Social Sciences (TESS) platform that offers researchers access to nationally representative samples through Knowledge Networks (KN), which conducts online surveys using its KnowledgePanel®. KN supplements the national random-digit dial sampling frame with the address-based sampling frame to also include cell-phone only households in their respondent pool. Since surveys are administered via the internet, all recruited households without a computer and/or internet access were given a WebTV or laptop and free internet access. While care is taken to minimize selection bias, there is no guarantee that all sources of potential bias are eliminated. The sampling methodology is described in Section 1 of the Online Resource. Between Dec 30, 2009 and Jan 13, 2010 a total of 841 panel members were invited to participate in the study and 556 (66%) completed the survey. Their ages ranged from 18 to 92 (mean=47.8, median=47.5, SD.=17.1), and 41.12% of the respondents were males.

The tasks All the participants performed three tasks:

(a) They read eight sentences extracted from the IPCC reports (IPCC 2001; 2007; see Online Resource, Section 2) which contained probabilistic terms pertaining to climate events. We selected two items using each of the following four terms *Very Likely, Likely, Unlikely* and *Very Unlikely*. One of the two items for each term involved a specific prediction (For example, "The Greenland ice sheet and other Arctic ice fields ***likely*** contributed no more than 4 m of the observed sea level rise."). The other item involved use of the same term in a more general statement about climate change, such as "Temperatures of the most extreme hot nights, cold nights and cold days are ***likely*** to have increased due to anthropogenic forcing." The respondents entered their best estimate[1] of the probability intended by the report's authors numerically, and those were displayed graphically in the form of a red 'Feeling Thermometer' (see Fig. 1). They could adjust these estimates until they were comfortable and proceeded to the next item.

(b) Respondents completed a short questionnaire measuring their attitudes toward climate change that consisted of a subset of the scales used by Heath and Gifford (2006) and Broomell et al. (2009)—Belief in GCC Scale—(BGCC)—2 items; Personal Experience with GCC (PE)—3 items; Perception of Causes of GCC (PCA)—4 items; and Perception of Consequences of GCC (PCO)—4 items. The items are listed in Online Resource (section 3). The motivated reasoning hypothesis suggests that we will find systematic correlations between these scales and the numerical estimates.

(c) Respondents completed a five item numeracy questionnaire including questions from Peters et al. (2006) and Frederick (2005). The numeracy items (e.g., A bat and a ball cost \$1.10 in total. The bat costs \$1.00 more than the ball. How much does the ball cost?) are listed in Online Resource (Section 4).

The order of administration of tasks (a) and (b) was randomized, and task (c) was always the last one to be completed. There were no time limits or constraints and respondents answered at their own pace.

The design Participants were randomly assigned to three conditions defined by the presentation form of the probability words in the IPCC sentences: In the *control* group participants were not given any instructions on their interpretation; in the *translation* group they were provided with the IPCC translation table (Table 1); and in the *verbal-numeric* (VN) group all the probabilistic terms were accompanied by numerical probabilities (the ranges listed in the IPCC guidelines) in the text. Table 2 lists the sample sizes in the various groups. We had access to detailed demographic information about the respondents from the Knowledge Network Public Affairs Profile (see details and distributions in Section 5 of the Online Resource).

3 Results

Preliminary considerations Eleven respondents responded 0% to all 4 IPCC items including the higher (and positive) terms, *Likely, Very Likely,* and 3 responded 100% to the 4 items using the low (negative) phrases, *Very Unlikely, Unlikely.* It appears that these respondents were using the survey as a platform to communicate their views and opinions rather than responding to the items as instructed. To minimize the impact of such (potentially biased) responses on the results, we adopted a conservative approach and recoded all extreme responses (either 0% or 100%) *of all the respondents* as missing. A total of 371 responses (i.e., 8.3% of the total $556\times8=4{,}448$) and 109 probabilistic words (i.e. 4.9% of the total $556\times4=2{,}224$) were re-coded as a result.[2] Most of the cases recoded affected the term *Very Likely* (32%) with the other three words being affected about equally.

[1] Budescu et al. (2009) also asked respondents to indicate a range of values that best describes the word. Our sample size was inversely proportional to the length of the questionnaire, so we decided to focus only on the mean values (as a proxy for the ranges), in an effort to maximize sample size.

[2] The two numbers are different because there were two items associated with each word.

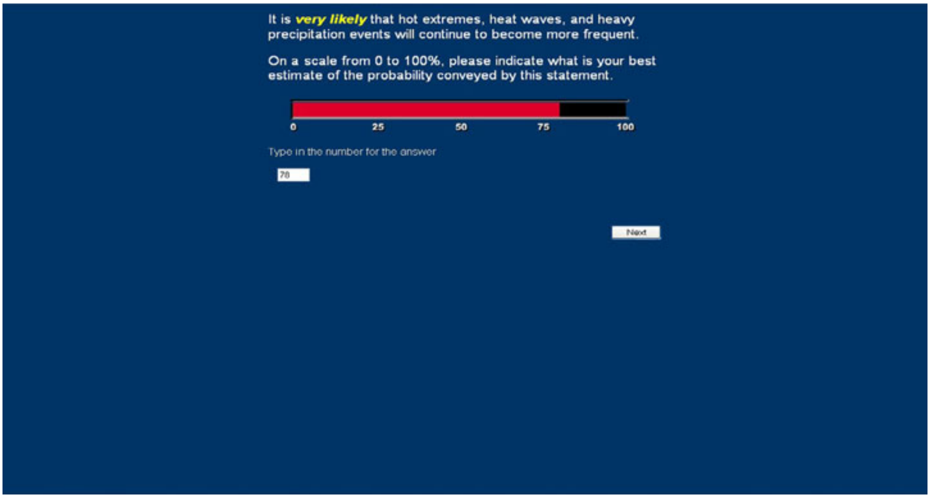


Fig. 1 Example of response screen for IPCC items using a feeling thermometer (Control condition)

Probability estimates We examined the numerical estimates provided as a function of the (a) word (*Very Likely, Likely, Unlikely, Very Unlikely*), (b) presentation format (control, translation, VN), (c) order of presentation (IPCC items first then attitude survey or vice-versa) and (d) item content (specific prediction, general statement) using a 4-way ANOVA. The format and order are between respondent factors and terms and content are repeated within-respondents factors. The order of the tasks and the content of the item did not affect the numerical assignment significantly. Figure 2 presents the mean values assigned to the four terms in the three groups averaged across items and order of tasks (The horizontal lines in the graph denote the thresholds used by IPCC and are included to facilitate interpretation).

We found significant differences between words ($F_{3,1293}=150.36$, $p<.001$). As expected, the mean probability estimates were the highest for *Very Likely*, followed by *Likely*, and then *Unlikely*, and lowest for *Very Unlikely* and all differences between adjacent words are significant, indicating that the respondents' answers were meaningful and sensible. We also found significant differences between presentation formats, ($F_{2,431}=3.54$, $p=.03$). Most importantly, we found an interaction between presentation format and words ($F_{6,1293}=3.30$, $p=0.003$). As Fig. 2 shows, the differentiation between the words is most (least) pronounced in the VN (Control) group. We found significant differences between VN and the average of the other two groups for *Unlikely* ($F_{1,497}=8.75$, $p=.003$) and *Very Unlikely* ($F_{1,497}=8.23$, $p=.004$), and between the control and translation groups for *Very Likely* ($F_{1,497}=5.61$, $p=.018$) and *Likely* ($F_{1,497}=6.10$, $p=.014$).[3]

Figure 3 shows the complete distribution of responses for each term by presentation format. The range of values assigned to the probabilistic words far exceeded what the IPCC authors intended. Table 3 presents the proportion of respondents whose estimates are consistent with the IPCC guidelines, for each word in every group. The rates are not high but, clearly, the VN presentation is superior, on average, as it reduces the number of out of

[3] We re-ran the same analysis with the complete data (including all the extreme responses), and found significant differences between words, presentation formats and an interaction between words and presentation formats. Similarly, the differences between words is most (least) pronounced in the VN (Control) group. This analysis confirms the robustness of our results. The detailed results are described in Section 6 of the Online Resource.

Table 2 Conditions and group sizes in the experiment

Group	Group	Order	Sample size	
1	Control	1	100	193
2		2	93	
3	Translation	1	89	175
4		2	86	
5	Verbal-Numerical	1	96	188
6		2	92	
		Total	556	556

range response from about 80% to approximately 70%. The advantage of this presentation is most pronounced for the extreme terms (*Very Likely* and *Very Unlikely*).

The mean estimates for the four words are 41 for *Very Unlikely*, 44 for *Unlikely*, 54 for *Likely*, and 62 for *Very Likely*. The most striking result is how regressive these estimates are and the low level of correspondence between them and the IPCC guidelines. These values are more regressive than in other studies of probability words (e.g., Budescu, et al. 2009). In section 7 of the Online Resources we analyze the data and seek to identify some of the sources of this difference. Recall, however, the main goal of our study is to compare between the various presentation methods and, indeed as predicted, the terms are better differentiated, and the number of cases out of range is reduced considerably (though not as much as one would hope), when they are presented in the newly proposed dual scale.

Consistency with IPCC guidelines We calculated for each respondent the consistency rate with the IPCC guidelines—the proportion of items whose interpretation was consistent with the IPCC guidelines.[4] This is a direct measure of the efficacy of the various communication modes. Consistency with the IPCC guidelines is quite low—24% of the respondents had no response consistent with the guidelines and about half of the respondents have only 1 or 2 responses consistent with the guidelines. Only 6% of the respondents had 6, or more, responses consistent with the guidelines. We examined the effects of two between respondent factors: (a) presentation format (control, translation, VN) and (b) order (IPCC items then attitude survey or vice-versa) on consistency rate. The order, as expected, was not significant. The presentation format was significant ($F_{2,542}=11.8$, $p<.001$)—The mean consistency rates were 20.76%, 18.81% and 30.12% for the control, translation and VN group respectively. The consistency rate in the VN group was higher than the control ($F_{1,542}=14.62$, $p=.001$) and the translation ($F_{1,542}=20.07$, $p<.001$) groups. Figure 4 displays the cumulative distributions of the consistency measures for the three groups and shows clearly the superiority of the VN presentation.

We also examined consistency with the guidelines at the item level and found that for all eight items respondents in the VN group were the most consistent. The probability of this pattern by chance is $\left(\frac{1}{3}\right)^8 < 0.0002$, suggesting that the VN presentation was more effective than both translation and control.

Another measure of consistency is the degree to which the estimates of the 4 terms comply with the ordering implied by the IPCC guidelines—Very Likely > Likely > Unlikely > Very Unlikely. We calculated for each respondent the Kendall τ_b rank-order

[4] Ten respondents were not included in this analysis because they used extreme values (either 0 or 100) for all items and these were coded as missing values.

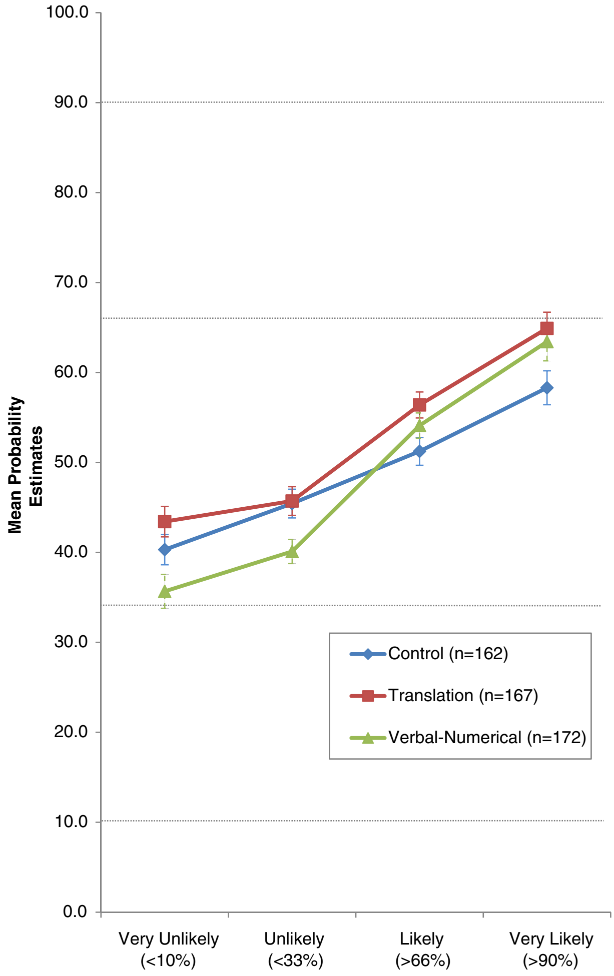


Fig. 2 Mean probability estimates (and SE bars) as a function of the presentation format and probability phrase

correlation between the order implied by the numerical estimates of each word and the IPCC conversion table. The highest median τ_b was found in the VN group (0.67) followed by the translation group (median τ_b=0.44), and the lowest was in the control group (0.33). These values correspond to order violation rates of 0.16, 0.28, and 0.33, respectively. A non-parametric (Kruskal-Wallis) ANOVA confirmed that the rate of order violations was lowest in the VN group (χ^2 (2, N=511)=7.17, p=0.028).

Correlations between probability estimates and covariates The first four columns of Table 4 display correlations between the probability estimates of the 4 words and the various covariates. The correlations are very similar in all experimental groups, so we focus our discussion on results based on the full sample. The probability estimates for *Very Likely, Likely* and *Unlikely* correlate positively and significantly with the Belief in Global Climate Change (BGCC), Personal Experience (PE), Perception of Causes (PCA) and Perception of Consequences (PCO), suggesting that respondents who believe more strongly in GCC assigned, on average, higher probability estimates to the words. The trend is more pronounced for the high probability words, *Very Likely* and *Likely.* This can be best

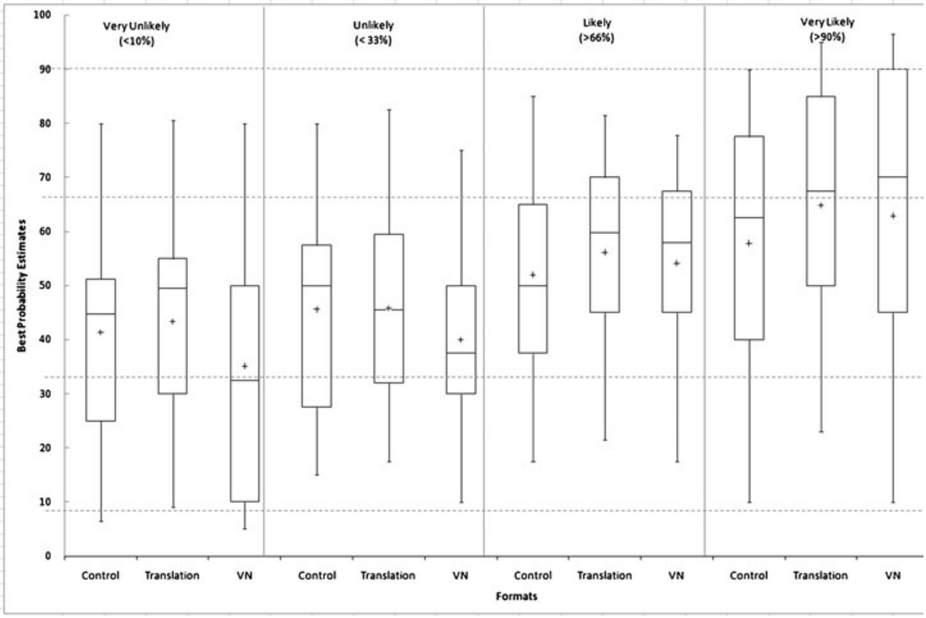


Fig. 3 The central 90% of the distribution of respondents' best estimates of the meanings of the four probability terms in the three conditions. Each box includes the central 50% of judgments, the solid lines in the boxes mark the medians and the '+'mark the means. The horizontal dotted lines represent the IPCC guidelines for interpreting the probability terms

explained by the fact that they are "positive" (Teigen and Brun 1999), i.e., they trigger thoughts about reasons that can cause the target events and the observation that there is a general "positivity bias" (Budescu et al. 2003).

The numeracy scores, and the respondents' education, were positively correlated with the estimates of the positive words—*Very Likely* and *Likely*—and negatively correlated with the negative words (significant only with *Very Unlikely*). There were no systematic correlations between the estimates and age, gender, location and religiosity, but there was a clear relationship with party affiliation, as shown in Fig. 5. Essentially, participants who identify strongly with the Democratic party assigned higher probabilities to all the words than the participants who identify strongly with the Republican party with the participants with weaker, or no, political affiliation in between (and showing much weaker differentiation between the words).

Correlations between consistency rate and covariates The last column in Table 4 shows that consistency rates with the IPCC guidelines correlated with 3 of the scales that measure attitudes toward climate change (the one exception being Personal Experience). Stronger beliefs in climate change were correlated with higher consistency with the IPCC guidelines. Consistency rates were also correlated (positively) with the numeracy score and education, and (negatively) correlated with the frequency of religious service attendance.

A closer examination of the correlations in each experimental group revealed that the correlations with Numeracy were significant only in the translation and VN groups, where respondents were provided with numerical values in addition to the probabilistic terms. Figure 6 also shows that those respondents with higher numeracy scores were more

Table 3 Proportion of respondents consistent with IPCC guidelines for each item

	Item	Control		Translation		Verbal-numerical	
		N	Consistency (%)	N	Consistency (%)	N	Consistency (%)
Very likely (>90%)	1	157	5.1	165	8.5	168	19.6
	2	165	3.0	167	10.2	169	18.9
Likely (>66%)	3	177	24.3	168	28.0	179	35.2
	4	166	37.3	167	30.5	171	36.8
Unlikely (<33%)	5	168	45.2	168	25.6	176	43.2
	6	159	33.3	166	31.3	174	43.1
Very unlikely (<10%)	7	181	7.2	168	8.9	176	20.4
	8	168	6.5	166	7.2	169	21.3
Mean			20.3		18.8		29.8

consistent only in the VN group, while the consistency of the respondents with low numeracy scores was not affected.[5]

Re-analysis with covariates In light of the results in the previous section it is natural to ask whether the effects of presentation format documented earlier are meaningful above and beyond the strong and persistent individual differences. With this in mind we re-examined the numerical estimates provided as a function of the (a) word (*Very Likely, Likely, Unlikely, Very Unlikely*), (b) presentation format (control, translation, verbal-numerical), (c) order (IPCC items first then attitude survey or vice-versa) and (d) item content (specific prediction, general statement) using a 4-way ANCOVA. The covariates used are those variables which had significant correlations with the estimates (see Table 4)—the four attitude scales, numeracy scores, education and party affiliation.

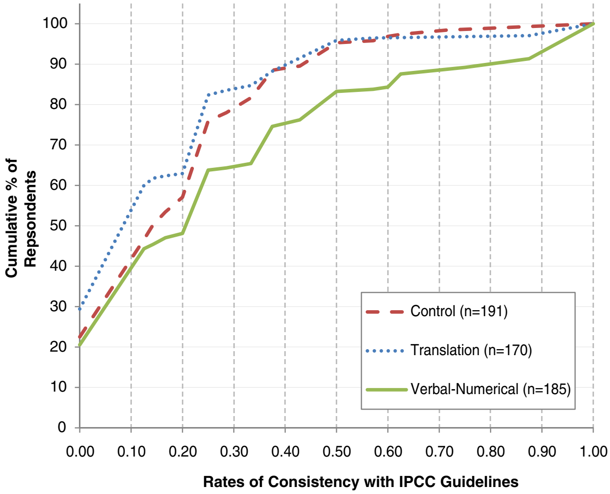
We replicated our key results—We found significant differences between words ($F_{3,1272}$= 25.19, p<.001): as expected, the mean probability estimates were the highest for *Very Likely*, followed by *Likely*, and then *Unlikely*, and lowest for *Very Unlikely*. We also replicated the significant differences between presentation formats, ($F_{2,424}$=3.16, p=.04) and, most importantly, the interaction between presentation format and words ($F_{6,1272}$=7.72, p<0.001). Figure 7 plots the least squares estimates of the mean probabilities (adjusted for all the covariates), and shows that the differentiation between the words was most pronounced in the VN group.[6]

We also re-examined the effects of the two between respondent factors: (a) presentation format (control, translation, verbal-numerical) and (b) order (IPCC items then attitude survey or vice-versa) on consistency rate using the same covariates. The presentation order, as expected, was not significant. The presentation format was significant ($F_{2,542}$=12.11, p<.001). The least square mean consistency rates after adjusting for the covariates were 20.5%, 18.3% and 30.7% for the control, translation and verbal-numerical group, respectively. The consistency

[5] Relatively fewer respondents scored 3 or more on the numeracy scale and this contributed to higher standard errors observed for these scores

[6] We replicated most of the key ANCOVA results using the complete data. We found significant differences between words, but not between formats. Most importantly, the interaction between words and formats is significant with the VN group showing the best differentiation between terms (see Section 6 in the Online Resource for the detailed results).

Fig. 4 Cumulative distributions of rates of individual consistency with IPCC guidelines as a function of presentation format



rate in the VN group was higher than the control ($F_{1,542}$=18.22, $p<.001$) and the translation ($F_{1,542}$=25.26, $p<.001$) groups.

4 Summary and discussion

This study replicated and generalized many of the results reported by Budescu et al. (2009) and its key points are easily summarized. We have shown that (a) the public consistently misinterprets the probabilistic statements in the IPCC report and (b) most misinterpretations are regressive, i.e., on the average the public interprets these pronouncements as less extreme than intended by the authors of the reports. This is especially true for the terms *Very Likely* and *Very Unlikely*. Most people, independent of their views of climate change and their political ideology, estimate the probabilities as substantially as less extreme than intended by the IPCC reports. For example, in the statement, "Continued greenhouse gas emissions at or above current rates would cause further warming and induce many changes in the global climate system during the 21st century that would *very likely* be larger than those observed during the 20th century.[7]", the IPCC intends to convey a probability of at least 0.90, but the typical respondent (i.e., the median response) interprets this to mean only about 0.65–0.75 (even in the VN condition). Moreover, (c) there are large individual differences in the interpretation of these statements, which are strongly associated with the respondents' ideology and their views and beliefs about climate change (e.g. Leiserowitz et al. 2011), as predicted by the literature on confirmation bias and motivated reasoning.

We demonstrated that (d) alternative communication formats can improve the quality of communication. More specifically, compared to the current presentation mode, the dual (verbal—numerical) scale (1) increases the level of differentiation between the various terms, (2) increases the consistency of interpretation of these terms, and (3) increases the level of consistency with the IPCC guidelines. Most importantly, (4) these positive effects are independent of the respondents' ideological and environmental views.

Two new findings that were made possible by the availability of a representative national sample are the clear documentation of the association between the interpretation of terms

[7] Extracted from IPCC AR4 Working Group I Report "The Physical Science Basis", Chapter 10.

Table 4 Correlations of probability estimates and consistency rates with covariates (Across all groups)

	Probability estimates				Consistency rate
	Very likely	Likely	Unlikely	Very unlikely	
Belief in global climate change	***0.56***	***0.52***	***0.29***	0.08	***0.15***
Personal experience	***0.49***	***0.46***	***0.33***	***0.15***	0.04
Perception of causes	***0.48***	***0.50***	***0.26***	0.01	***0.14***
Perception of consequences	***0.47***	***0.46***	***0.22***	−0.02	***0.14***
Numeracy score	***0.17***	***0.16***	−0.04	***−0.15***	***0.14***
Age	−0.05	***−0.13***	0.00	0.07	−0.06
Gender	0.03	−0.03	0.03	−0.04	0.00
Education	***0.11***	***0.09***	−0.07	***−0.11***	***0.17***
Location (Metro, Non-Metro)	0.02	0.01	−0.07	−0.03	0.04
Party	***0.29***	***0.28***	***0.21***	0.08	0.06
Religiosity	0.05	0.02	0.07	***0.11***	***−0.09***

Correlations in italicized bold are significant ($p<0.05$)

representing uncertainty and (a) numeracy and (b) political identification and ideological perspectives. The dual scale that presents both verbal and numerical information caters to people with different preferences for information. Providing numerical values helped increase the accuracy of interpretation among individuals with high numeracy scores, without having any negative effects on the respondents with lower numeracy scores. The differences in interpretation between respondents with various political views and attitudes towards GCC reiterate Piercey's (2009) observation that people tend to interpret these words in ways that are consistent with their beliefs (see also Bastardi et al. 2011) and expectations as predicted by the motivated reasoning argument (Kunda 1990), as well as Fox and Irwin's (1998) observation that uncertainties with political implications can increase the polarization in situations where listeners are less amendable to updating their prior beliefs. It is reassuring to see that, despite this polarity, the efficacy of communication was improved significantly by the dual scale indicating that this presentation mode affects positively all respondents, regardless of their political affiliation and their views.

It is surprising to find this level of inter-individual variability and inconsistency with the IPCC guidelines, even when they are explicitly stated in the text. We do not believe there is a single simple explanation for this pattern; rather, it is due to the joint effects of multiple factors. Some of them—the (constantly underestimated) vagueness and elasticity of the language, and the communicator-listener discrepancy—are cognitive in nature and are often cited in the literature on probability words. Others are motivational and are related to the specific context. They are special instances of motivated reasoning—judgments driven by pre-conceived beliefs and theories resistant to contradictory evidence—as illustrated by the differences between the respondents with different political orientations.

As a result, the gap between the authors' intentions and readers' understanding of the probabilistic communications are large and systematic, as demonstrated in this study. Climate change may be a particularly difficult domain to communicate risks because the public has only limited understanding of the relevant science (see recent surveys by Leiserowitz, and Smith 2010; Leiserowitz et al. 2010), and because our mental models are often inadequate for understanding complex dynamic systems (see Sterman 2011). It will take more than the guidelines table currently used by the IPCC to eliminate them. The VN

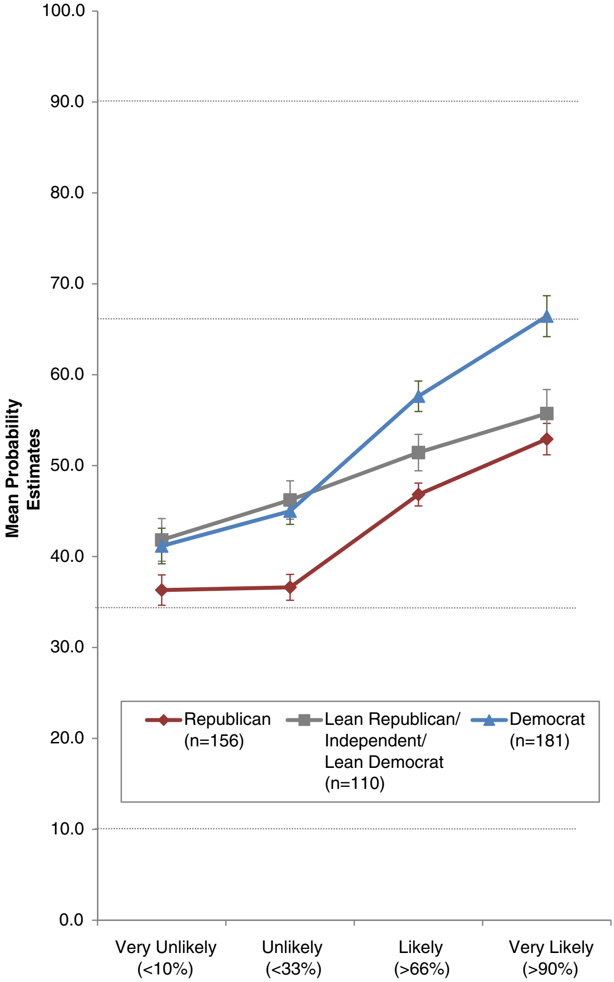
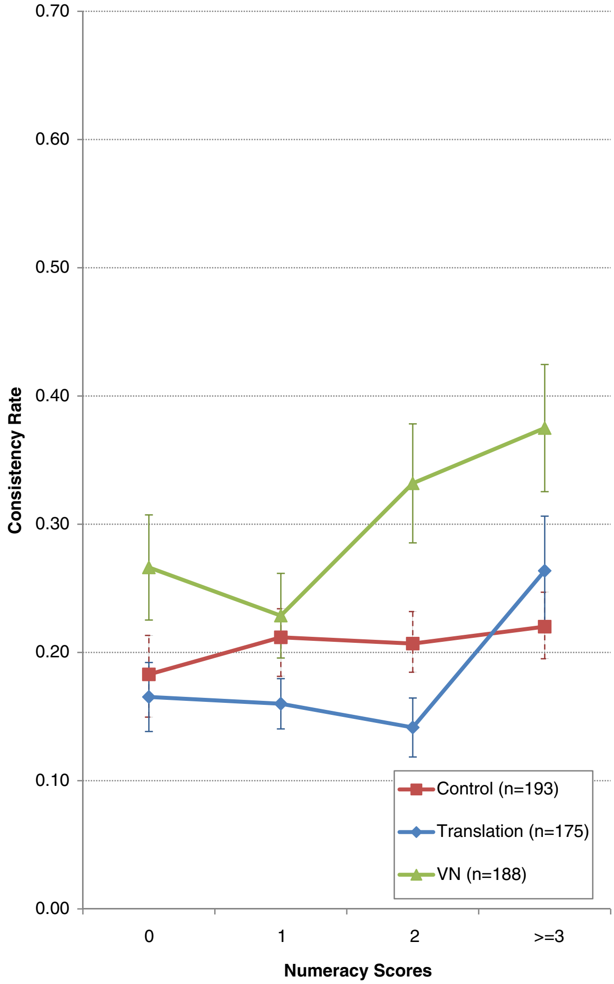


Fig. 5 Mean probability estimates for the four terms (and SE bars) as a function of party affiliation

format is a step in the right direction that can be implemented easily, quickly and with almost no cost. Unfortunately, it is just a first and, relatively, modest step and large under (over) estimation of high (low) probabilities persist.

Our results should not be interpreted as a criticism of the decisions made by the IPCC on communication of uncertainties—there is no perfect or optimal method, any choice made by the IPCC would have been shown to be suboptimal in some sense, and we do not believe that there is one single major change that could improve the quality of the communication drastically. The history of research in risk communication suggests that the joint and cumulative effect of simple remedies, such as the one we propose, could go a long way towards improving the quality of communication of uncertainty in the IPCC leading to a better understanding of the assessments, and strengthening the foundation for sound policy decisions. By improving the presentation and communication of results, the IPCC (and other agencies involved in risk communication) can better fulfill their mission to "provide the world with a clear scientific view on the current state of knowledge in climate change and its potential environmental and socio-economic impacts" (also see Sterman 2011).

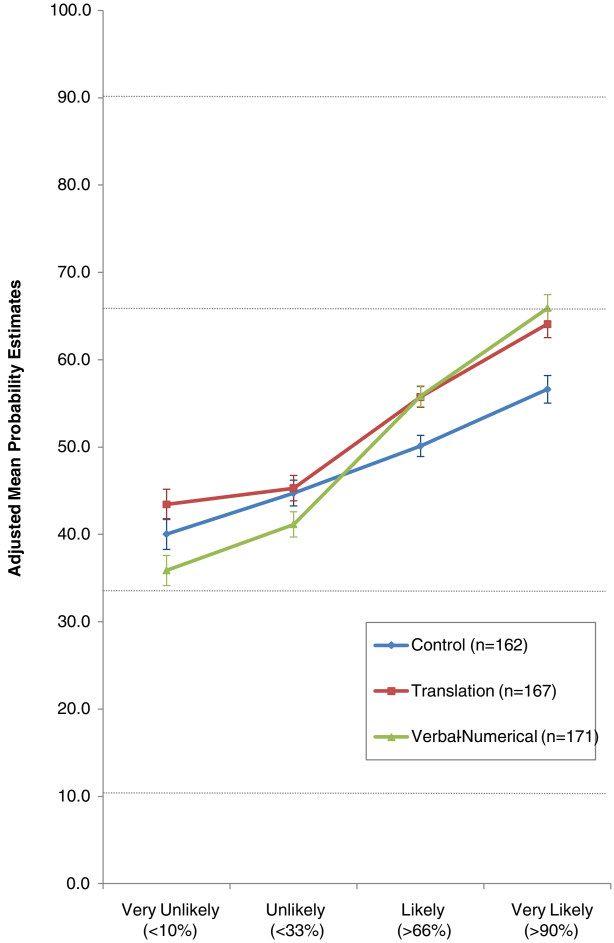
Fig. 6 Mean consistency rate (and SE bars) as a function of presentation format and numeracy score



We conclude this discussion by revisiting some of the recommendations from the Budescu et al. (2009) paper in light of our new results. Given the special role that these terms play in the report, we recommend that they be used selectively and judiciously. Probabilistic pronouncement should be used exclusively for precise and unambiguous events to avoid confounds between the ambiguity of the event and its underlying uncertainty (Wallsten and Budescu 1995; Fischhoff 1994). For example, when reading that "It is *very unlikely* that the MOC will undergo a *large abrup*t transition during the 21st century", people may not agree on what type of transition qualifies to be called *large* and/or *abrupt* and this ambiguity may affect their interpretations of the terms systematically. Authors should exercise caution when considering the use of negatively worded terms (e.g., *Unlikely*) in place of positively-worded terms (e.g., likely) because these phrases are not inverses of each other. A recent re-analysis of the Budescu et al. (2009) data (Smithson et al. 2011) showed that, everything else being equal, the negatively-worded terms resulted in greater variability in the responses than the positively-worded terms.

The key recommendation is to replace the current method that relies only on probability words with a dual scale that uses both *verbal terms and numerical values*.

Fig. 7 Adjusted mean probability estimates (and SE bars) as a function of presentation format and probability phrase



The report could maintain the current coarse seven-category classification with labels and define general guidelines for the use of these terms (as in Table 1) but, in addition, every term should be accompanied by a range of probabilities. The terms should rely on a common stem (*likely*) accompanied by modifiers that determine the terms' ordering (Lipkus 2007). This approach would assure continuity with previous reports (Ha-Duong, et al 2007) but we recommend that numerical boundaries associated with the various terms should be adjusted on the basis of the numerous published scaling studies (e.g., Wallsten and Budescu 1995; Mosteller and Youtz 1990). It is important to note that the vast majority of these studies were done in English and there is no research on the universality of these boundaries across languages.[8]

[8] The translation table is ambiguous as it does not specify whether the various ranges are mutually exclusive or overlapping (for example, it is unclear whether *likely* applies to all probabilities above 66%, or only to values between 67% and 90%). The recently published guidelines for the 5th assessment (Mastrandea, et al. 2010) explicitly state that they are *not exhaustive*, i.e. *likely* can be used for any probability between 66% and 100% (not only for the 67–90% range). Although we have no data on this point we suspect that most people will find this solution counterintuitive and confusing.

The dual approach provides more information, facilitates similar interpretation of the phrases, and caters to a broad and heterogeneous audience with various levels of expertise and preferences (see Witteman and Renooij 2003; Witteman et al. 2007). Our results show that supplementing verbal terms with numerical boundaries improved considerably the quality of communication. This proposed changed is very easy and inexpensive to implement: It does not require any additional research, nor any other special changes in the report. Another indirect advantage of this approach is that the presence of the relevant numerical ranges in each sentence would force the authors to think carefully about their choices of terms, and would reduce the number of instances where the authors ignore (or forget) the ranges prescribed by the IPCC for each term. This is also likely to increase precision and efficiency of the communication.

We reiterate that we do not believe that there is an "optimal" presentation format and no single change that would improve the quality of communication drastically. We think that improvement will be achieved through a combination of many small changes, such as those proposed in this paper. One insight from our work is that the communication can benefit from the combination of multiple communication modes and formats. This suggests the possibility that additional improvements could be achieved by adding other modalities and formats. For example, we speculate that in online versions of the reports in which the probability terms and their numerical bounds are supplemented (through hyperlinks) by *optional* graphical displays related to (a) the sources of uncertainty underlying a certain forecast and / or (b) the uncertainty scales could improve the overall quality of communication.

Acknowledgements This work was supported by the National Science Foundation under Grant No. 0345925, and the data were collected by Time-sharing Experiments for the Social Sciences, NSF Grant 0818839 (Jeremy Freese and Penny Visser, Principal Investigators). The third author was supported by the National Geospatial-Intelligence Agency (under Grant No. HM1582-09-1-0020). The opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the funding entity.

Many thanks to Profs. Klaus Keller and Thomas Wallsten and three reviewers for useful comments and suggestions on an earlier version.

References

Bastardi A, Uhlmann EL, Ross L (2011) Wishful thinking: belief, desire, and the motivated evaluation of scientific evidence. Psychol Sci 22:731–732

Begley S (2007) The truth about denial. Newsweek 150:20–29

Beyth-Marom R (1982) How probable is probable? A numerical translation of verbal probability expressions. J Forecast 1:257–269

Broomell S, Budescu DV, Por H (2009) Beliefs and actions: an analysis of the public's attitudes toward global climate change. Paper presented at the Annual Meeting of the Society of Judgment and Decision Making, Boston

Brun W, Teigen KH (1988) Verbal probabilities: ambiguous, context-dependent, or both? Organ Behav Hum Decis Process 41:390–414

Bryant G, Norman G (1980) Expressions of probability: words and numbers. N Engl J Med 302:411

Budescu DV, Broomell SB, Por HH (2009) Improving communication of uncertainty in the reports of the Intergovernmental Panel on Climate Change. Psychol Sci 20:299–308

Budescu DV, Wallsten TS (1990) Dyadic decisions with numerical and verbal probabilities. Organ Behav Hum Decis Process 46:240–263

Budescu DV, Wallsten TS (1985) Consistency in interpretation of probabilistic phrases. Organ Behav Hum Decis Process 36:391–405

Budescu DV, Wallsten TS (1987) Subjective estimation of precise and vague uncertainties. In: Wright G, Ayton P (eds) Judgmental forecasting. Wiley, New York, pp 63–82

Budescu DV, Weinberg S, Wallsten TS (1988) Decisions based on numerically and verbally expressed uncertainties. J Exp Psychol Hum Percept Perform 14:281–294
Budescu DV, Karelitz TM, Wallsten TS (2003) Predicting the directionality of probability words from their membership functions. J Behav Decis Mak 16:159–180
Chesley GR (1985) Interpretation of uncertainty expressions. Contemp Account Res 2:179–199
Clarke VA, Ruffin CL, Hill DJ, Beamen AL (1992) Ratings of orally presented verbal expressions of probability by a heterogeneous sample. J Appl Soc Psychol 22:638–656
Dhami MK, Wallsten TS (2005) Interpersonal comparison of subjective probabilities: toward translating linguistic probabilities. Mem Cognit 33:1057–1068
Erev I, Cohen BL (1990) Verbal versus numerical probabilities: Efficiency, biases, and the preference paradox. Organ Behav Hum Decis Process 45:1–18
Erev I, Wallsten TS, Neal M (1991) Vagueness, ambiguity, and the cost of mutual understanding. Psychol Sci 2:231–324
illenbaum S, Wallsten TS, Cohen BL, Cox JA (1991) Some effects of vocabulary and communication task on the understanding and use of vague probability expressions. Am J Psychol 104:35–60
Fischer K, Jungermann H (1996) Rarely occurring headaches and rarely occurring blindness: is rarely = rarely? The meaning of verbal frequentistic labels in specific medical contexts. J Behav Decis Mak 9:153–172
Fischhoff B (1994) What forecasts (seem to) mean. Int J Forecast 10:387–403
Frederick S (2005) Cognitive reflection and decision making. J Econ Perspect 19(4):25–42
Fox CR, Irwin JR (1998) The role of context in the communication of uncertain beliefs. Basic Appl Soc Psychol 20:57–70
Hamm RM (1991) Selection of verbal probabilities: a solution for some problems of verbal probability expression. Organ Behav Hum Decis Process 48:193–223
Heath Y, Gifford R (2006) Free-market ideology and environmental degradation: the case of belief in global climate change. Environ Behav 38:48–71
Ha-Duong M, Swart R, Bernstein L, Petersen R (2007) Uncertainty management in the IPCC: agreeing to disagree. Glob Environ Chang 17:8–11
Intergovernmental Panel on Climate Change (2001) Summary for policymakers: contribution of working group I to the third assessment report of the intergovernmental panel on climate change. Retrieved July 2009 from http://www.grida.no/climate/ipcc_tar/wg1/pdf/WG1_TAR-FRONT.pdf
Intergovernmental Panel on Climate Change (2005) Guidance notes for lead authors of the IPCC fourth assessment report on addressing uncertainties. Retrieved February 2011 from http://www.ipcc.ch/meetings/ar4-workshops-express-meetings/uncertainty-guidance-note.pdf
Intergovernmental Panel on Climate Change (2007) Summary for policymakers: contribution of working group I to the fourth assessment report of the intergovernmental panel on climate change. Retrieved February 2007 from http://www.ipcc.ch/pdf/assessment-report/ar4/wg1/ar4-wg1-spm.pdf
Johnson EM (1973) Numerical encoding of qualitative expressions of uncertainty (Technical paper No. 250). US Army Research Institute for the Behavioral & Social Sciences
Karelitz TM, Budescu DV (2004) You say "Probable" and I say "Likely": improving interpersonal communication with verbal probability phrases. J Exp Psychol Appl 10:25–41
Klayman J, Ha Y (1987) Confirmation, disconfirmation and information in hypothesis testing. Psychol Rev 94:211–228
Kong A, Barnett GO, Mosteller F, Youtz C (1986) How medical professionals evaluate expressions of probability. N Engl J Med 315:740–744
Kunda Z (1990) The case for motivated reasoning. Psychol Bull 180:480–498
Leiserowitz A, Maibach E, Roser-Renouf C, Hmielowski JD (2011) Politics & global warming: democrats, republicans, independents, and the tea party. Yale University and George Mason University. New Haven, CT: Yale Project on Climate Change Communication. Retrieved from http://environment.yale.edu/climate/files/PoliticsGlobalWarming2011.pdf
Leiserowitz A, Smith N (2010) Knowledge of climate change across global warming's six Americas. Yale University. New Haven, CT: Yale Project on Climate Change Communication
Leiserowitz A, Smith N, Marlon JR (2010) Americans' knowledge of climate change. Yale University. New Haven, CT: Yale Project on Climate Change Communication. Retrieved from http://environment.yale.edu/climate/files/ClimateChangeKnowledge2010.pdf
Lipkus IM (2007) Communicating the uncertainty of harms and benefits of medical interventions. Med Decis Making 27:681–695
Mapes REA (1979) Verbal and numerical estimates of probability terms. J Gen Intern Med 6:237
Mastrandea MD, Field CB, Stocker TF, Edenhofer O, Rbi KL, Frame DJ, Held H, Kreogler E, Mach KJ, Matschoss PR, Plattner GK, Yohee GW, Zwiers FW (2010) Guidance note for lead authors of the IPCC

fifth assessment report on consistent treatment of uncertainties. Intergovermental Panel on Climate Change (IPCC). (Available at http://www.ipcc.ch)
Merz JF, Druzdzel MJ, Mazur DJ (1991) Verbal expressions of probability in informed consent litigation. J Med Decis Making 11:273–281
Mosteller F, Youtz C (1990) Quantifying probabilistic expressions. Stat Sci 5:2–16
Mullet E, Rivet I (1991) Comprehension of verbal probability expressions in children and adolescents. Lang Comm 11:217–225
Nakao M, Axelrod S (1983) Numbers are better than words: verbal specifications of frequency have no place in medicine. Am J Med 74:1061
Nickerson RS (1998) Confirmation bias: a ubiquitous phenomenon in many guises. Rev Gen Psychol 2:175–220
Nobelprize.org (2008) The Nobel peace prize 2007. Retrieved November 2007 from http://nobelprize.org/nobel_prizes/peace/laureates/2007/press.html
Olson MJ, Budescu DV (1997) Patterns of preference for numerical and verbal probabilities. J Behav Decis Mak 10:117–131
Patt AG, Schrag DP (2003) Using specific language to describe risk and probability. Clim Chang 61:17–30. doi:10.1023/A:1026314523443
Pepper S, Prytulak LS (1974) Sometimes frequently means seldom: context effects in the interpretation of quantitative expressions. J Res Personal 8:95–101
Peters E, Västfjäll D, Slovic P, Mertz CK, Mazzocco K, Dickert S (2006) Numeracy and decision making. Psychol Sci 17:407–413
Pew Research Center for for the People and the Press (2009) Scientific achievements less prominent than a decade ago: public praise science; Scientists fault public, media. Retrieved Novermber 2009 from http://people-press.org/reports/pdf/528.pdf
Piercey MD (2009) Motivated reasoning and verbal vs. numerical probability assessment: evidence from an accounting context. Organ Behav Hum Decis Process 108:330–341
Pyszcynski T, Greenberg J (1987) Toward an integration of cognitive and motivational perspectives on social inferences: a biased hypothesis testing model. In: Berkowitz L (ed) Advances in experimental social psychology. Academic Press, New York, pp 297–340
Reagan R, Mosteller F, Youtz C (1989) Quantitative meanings of verbal probability expressions. J Appl Psychol 74:433–442
Reyna VF, Brainerd CJ (2007) The importance of mathematics in health and human judgment: numeracy, risk communication, and medical decision making. Learn Individ Differ 17:147–159
Smithson M, Budescu DV, Broomell SB, Por H (2011) Never say "Not:" impact of negative wording in probability phrases on imprecise probability judgments. Presented at 7th International Symposium on Imprecise Probability: Theories and Applications, Innsbruck, Austria
Sterman JD (2011) Communicating climate change risks in a skeptical world. *Climatic Change.* In press
Sutherland HJ, Lockwood GA, Tritchler DL, Sem F (1991) Communicating probabilistic information to cancer patients: is there "noise" on the line? Soc Sci Med 32:725–731
Teigen KH, Brun W (1999) The directionality of verbal probability expressions: effects on decisions, predictions and probabilistic reasoning. Organ Behav Hum Decis Process 80:155–190
Wallsten TS, Budescu DV (1995) A review of human linguistic probability processing: general principles and empirical evidence. Knowl Eng Rev 10:43–62
Wallsten TS, Budescu DV, Zwick R, Kemp SM (1993) Preferences and reasons for communicating probabilistic information in numerical or verbal terms. Bull Psychon Soc 31:135–138
Wallsten TS, Fillenbaum S, Cox JA (1986) Base rate effects on the interpretations of probability and frequency expressions. J Mem Lang 25:571–587
Weber EU, Hilton DJ (1990) Contextual effects in the interpretations of probability words: perceived base rate and severity of events. J Exp Psychol Hum Percept Perform 16:781–789
Witteman C, Renooij S (2003) Evaluation of a verbal–numerical probability scale. Int J Approx Reason 33:117–131
Witteman C, Renooij S, Koele PM (2007) Medicine in words and numbers: a cross-sectional survey comparing probability assessment scales. BMC Med Informat Decis Making 7
Zwick R, Wallsten TS (1989) Combining stochastic uncertainty and linguistic inexactness: theory and experimental evaluation of four fuzzy probability models. Int J Man Mach Stud 30:69–111