



Climate scientists' wide prediction intervals may be more likely but are perceived to be less certain

Article

Accepted Version

Løhre, E., Juanchich, M., Sirota, M., Teigen, K. H. and Shepherd, T. G. (2019) Climate scientists' wide prediction intervals may be more likely but are perceived to be less certain. *Weather, Climate and Society*, 11 (3). pp. 565-575. ISSN 1948-8327 doi: <https://doi.org/10.1175/WCAS-D-18-0136.1> Available at <http://centaur.reading.ac.uk/83185/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1175/WCAS-D-18-0136.1>

Publisher: American Meteorological Society

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

Climate scientists' wide prediction intervals are more likely but perceived to be less certain

Erik Løhre^{1,2}, Marie Juanchich³, Miroslav Sirota³, Karl Halvor Teigen^{1,4}, and Theodore G. Shepherd⁵

¹Simula Research Laboratory, Oslo, Norway

²Inland Norway University of Applied Sciences, Lillehammer, Norway

³Department of Psychology, University of Essex, UK

⁴Department of Psychology, University of Oslo, Norway

⁵Department of Meteorology, University of Reading, UK

Short title: Wide intervals are perceived as uncertain

Corresponding author:

Erik Løhre

Inland Norway University of Applied Sciences

Postboks 400, 2418 Elverum, Norway

Telephone: +47 41 42 01 35

E-mail: erik.lohre@gmail.com

26 **Abstract** (149 words)

27 The use of interval forecasts allows climate scientists to issue predictions with high levels of
28 certainty even for areas fraught with uncertainty, since wide intervals are objectively more
29 likely to capture the truth than narrow intervals. However, wide intervals are also less
30 informative about what the outcome will be than narrow intervals, implying a lack of
31 knowledge or subjective uncertainty in the forecaster. In six experiments, we investigate how
32 lay people perceive the (un)certainty associated with wide and narrow interval forecasts, and
33 find that the preference for accuracy (seeing wide intervals as “objectively” certain) vs.
34 informativeness (seeing wide intervals as indicating “subjective” uncertainty) is influenced by
35 contextual cues (e.g., question formulation). Most importantly, we find that people more
36 commonly and intuitively associate wide intervals with uncertainty than with certainty. Our
37 research thus challenges the wisdom of using wide intervals to construct statements of high
38 certainty in climate change reports.

39

40 Keywords: uncertainty, intervals, IPCC, climate change, communication

41

42 **1. Introduction**

43
44 The knowledge of general principles governing the climate system is sufficient to
45 make strong qualitative predictions about climate change. For instance, the Intergovernmental
46 Panel on Climate Change (IPCC) leaves little room for doubt when concluding that
47 “continued emissions of greenhouse gases will cause further warming and changes in all
48 components of the climate system” (IPCC 2013). In contrast, it is not possible to make precise
49 quantitative predictions of exactly how the climate will change, even under a given forcing
50 scenario (such conditional predictions are typically called projections). Thus, climate
51 scientists generally issue predictions in the form of interval (range) forecasts (e.g., 0.3 to
52 1.7°C temperature rise¹, 0.26 to 0.55 m sea level rise) rather than point forecasts (e.g., 1.0°C
53 temperature rise). Interval estimates allow a tradeoff between forecast precision and forecast
54 certainty, or what Yaniv and Foster (1995) has described as a tradeoff between
55 informativeness and accuracy. If a high degree of certainty (accuracy) is desired, one can
56 forecast a wide interval (the rate of sea level rise [during the 21st century] will *very likely*
57 exceed that observed during 1971 to 2010 [meaning more than a 20 cm rise]). This is
58 commonly done in the IPCC reports when summary statements of high certainty are sought.
59 Alternatively, if a high level of precision (informativeness) is desired, one can forecast a
60 narrower interval with a lower degree of certainty (it is *likely* the sea level will rise between
61 26 and 55 cm).

62 While a large body of research shows that people often misunderstand the verbal
63 probability expressions (e.g., “very likely”, “unlikely”) used by the IPCC (Budescu et al.
64 2009; Budescu et al. 2012; Budescu et al. 2014; Harris and Corner 2011; Harris et al. 2017;
65 Harris et al. 2013; Ho et al. 2015; Juanchich and Sirota 2017), few studies have examined

¹ All examples are taken from IPCC, 2013: Summary for policymakers. *Climate change 2013: The physical science basis. Contribution of Working Group I to the fifth assessment report of the Intergovernmental Panel on Climate Change*, T. F. Stocker, and Coauthors, Eds., Cambridge University Press.

66 how lay people respond to the use of intervals to communicate degrees of (un)certainty in the
67 climate change domain (Dieckmann et al. 2015; Dieckmann et al. 2017; Joslyn and LeClerc
68 2016; Løhre and Teigen 2017). We argue and demonstrate in this paper that the relationship
69 between interval width (i.e., forecast precision) and certainty is ambiguous: a wide interval
70 (an imprecise forecast) is “accurate” in the sense that it has a high probability of capturing the
71 actual outcome, but its width also signals greater uncertainty about what the outcome will be,
72 in comparison to a narrow interval (a more precise and hence more informative forecast). This
73 ambiguity makes it important for forecasters to know whether lay people see wide intervals as
74 more (or less) certain than narrow ones, and which of these two perspectives on intervals is
75 more frequent and more intuitively appealing.

76 The two perspectives on the relationship between interval width and certainty may
77 rely on two forms of certainty (Fox and Ülkümen 2011; Hacking 1975; Kahneman and
78 Tversky 1982). On the one hand, certainty refers to our state of knowledge or belief. Such
79 internal or subjective certainty is often expressed by statements where the subject is a sentient
80 being (“I am 90% certain”), and using subjective terms like being confident, or sure (Fox and
81 Ülkümen 2017; Ülkümen et al. 2016). But certainty can also be used in an external, more
82 objective sense, reflecting variability, predictability and randomness in the outside world.
83 Degrees of certainty are in these contexts often embedded in statements with an impersonal
84 subject (“it is 90% certain”), and are used synonymously with degrees of probability,
85 likelihood, or chance (Juanchich et al. 2017; Løhre and Teigen 2016).

86 With interval predictions, a wider interval allows for a greater degree of objective
87 certainty (more hits and fewer misses). Even if the exact number of hits vs. misses can be
88 assessed only retrospectively, after the outcomes are known, this general relationship can be
89 claimed prospectively on purely logical grounds. Subjective certainty, however, might not
90 increase with interval width. In fact, people may see wide intervals as cueing *uncertainty* and

91 lack of knowledge, for two reasons. First, more knowledge about a topic enables one to be
92 more precise in one's statements about it (Yaniv and Foster 1997). Second, conversational
93 norms suggest that people seek to maximize informativeness in communication (Grice 1975).
94 The prediction "The temperature in Oslo will be between -35 and +35°C tomorrow" is true,
95 with close to 100% certainty, but is also far too vague to be useful for someone preparing for
96 a visit. A forecaster with higher subjective confidence may make a more precise, informative
97 prediction ("The temperature at noon will be between 15 and 18°C"), which can be seen as
98 conveying more certain expectations about tomorrow's weather.

99 Thus, different concepts of certainty might lead to different views on the implications
100 of wide vs. narrow interval predictions. Those who find a wide interval to be more certain, by
101 being more likely to include the true (actual) values, will in this paper be referred to as
102 showing a *preference for accuracy*. In contrast, those who consider a wide interval to be less
103 certain, by being less informative and expressing lower confidence about expected outcomes,
104 display a *preference for informativeness*.

105 Previous research has found support for both types of preference (or "mindsets"). In
106 line with the informativeness mindset, lay people expect experts to give narrower interval
107 estimates than novices (McKenzie et al. 2008). Recipients of information prefer precise
108 statements (Du et al. 2011; Jørgensen 2016), with narrow intervals occasionally preferred
109 over wide intervals even when the wide interval includes the correct answer while the narrow
110 interval does not (McKenzie and Amin 2002; Yaniv and Foster 1995). Teigen (1990) found
111 that people placed more confidence in precise statements than in vague statements, but also
112 that people chose the more precise statement when asked which statement they would be
113 more skeptical about. Participants in a recent study received high and low probability
114 forecasts made by climate change experts, and completed the forecasts by filling in
115 corresponding intervals (Løhre and Teigen 2017). Some associated high probabilities with

116 wide intervals, but many did the opposite and assigned *narrow* intervals to high probabilities.
117 Similar results were obtained when people were given wide and narrow interval forecasts, and
118 asked to fill in missing probability values. Some participants assumed wide intervals were
119 more probable, whereas others felt they were *less* probable than narrow intervals.

120 These studies leave open several important questions that we address in the present
121 paper. (1) Is one “mindset” more prevalent than the other? (2) Can contextual and linguistic
122 cues, which are known to change the way people think about probabilities (Løhre and Teigen
123 2016; Nisbett et al. 1983; Reeves and Lockhart 1993; Ülkümen et al. 2016), also influence
124 people’s views on the relationship between interval width and certainty? These two questions
125 were investigated in Experiments 1-5, where we manipulated the focus of a question about
126 certainty. We predicted that a question about which of two intervals is “more certain to be
127 correct”, would promote reflections about objective certainty, accuracy and the probability of
128 hits and misses, and should accordingly be answered in favour of the wide interval. On the
129 other hand, a question about which interval “conveys more certainty”, would make thoughts
130 about informational value and subjective certainty more salient, and induce people to find
131 wide intervals to imply *less* certainty than narrow ones. (3) A third issue is which mindset
132 people find more intuitive. Experiment 6 investigated people’s lay theories about interval
133 width and probability, and asked people to rate how intuitively appealing two statements
134 compatible with the two mindsets were.

135

136

<Insert Table 1 about here>

137

138 **2. Experiments 1-5: Effects of question type on the perception of wide vs. narrow** 139 **intervals**

140 **a. Participants**

141 The participants in these experiments (total $N = 923$, see Table 1) were university
142 students from the UK and Norway who volunteered to participate or who received course
143 credits for participation, and Amazon MTurk workers from the US who were paid to
144 complete the questionnaires. Both of these types of convenience samples are typical in
145 psychology experiments, and are often reasonably similar to community samples (Goodman
146 et al. 2013; Paolacci et al. 2010). For the purpose of the current studies, namely to investigate
147 subjective perceptions of interval forecasts of climate change, we would expect that
148 participants from these samples should be at least as well-equipped (if not better) to interpret
149 the information as more representative samples.

150 **b. Materials and procedure**

151 In all experiments, the participants received interval forecasts of sea level rise and
152 temperature rise by the end of the century from two different teams of climate scientists. One
153 team issued a forecast with a wide interval (e.g., “The temperature will increase between 1.1°
154 Celsius and 6.4° Celsius”), while the other team gave a forecast with a narrower interval (e.g.,
155 “The temperature will increase between 2.2° Celsius and 5.4° Celsius”). The participants were
156 asked, in three to four different conditions in the different experiments, to choose which
157 prediction “conveys more uncertainty [certainty]” or which prediction “is more likely [certain,
158 uncertain] to be correct”. These questions were formulated to focus on informativeness or on
159 accuracy, respectively. An overview of the questions used in the different experiments is
160 provided in Table 2, and more detailed descriptions of the procedure for each experiment is
161 provided below. The full description of the scenarios, as well as separate statistical analysis of
162 each experiment, can be found in the Supplementary materials (in the Results section, only
163 the overall results are described). Several of the experiments also investigated secondary
164 hypotheses, which are briefly described below, while more detailed descriptions and analyses
165 are provided in the Supplementary materials.

166

167

<Insert Table 2 about here>

168

169 1) MATERIALS AND PROCEDURE VARIATIONS IN EXPERIMENTS 1-5

170

171

172

173

174

175

176

177

178

179

In Experiment 1, we manipulated question type and reasons for variability in a 2 x 2 within-subject design. Participants completed a daily survey for 14 days. On the third day, the participants received questions about which interval “is most likely to be correct” and on day 6 which interval “conveys most uncertainty”. The same questions were repeated on days 9 and 11, but here, participants also received an explanation for the variability in the expert forecasts. The variability was explained by referring to temperature rise “in different countries” and sea level rise “in different parts of the world”. On day 14 participants rated their belief in climate change by answering four questions taken from Heath and Gifford (2006). For each scenario (temperature and sea level rise), participants could choose one of the two predictions or rate them as equal.

180

181

182

183

184

Participants in Experiment 2 received the same questions as in Experiment 1, but this was a 2 x 2 design with question type and reason for variability varied between subjects. Hence, participants in different groups received questions either about which interval “conveys most uncertainty” or which interval “is most likely to be correct”, and either received an explanation for the variability in estimates or did not receive such an explanation.

185

186

187

188

189

190

In Experiment 3, we attempted to control for some potential confounding factors in Experiments 1 and 2. Beside their focus on informativeness or accuracy, the questions used in the first two experiments differed in several respects. First, the term “uncertainty” was used in the informativeness-focus condition and the term “likely” was used in the accuracy-focus condition. These terms were assumed to be associated with different sources of uncertainty, with “uncertainty” being an internal/epistemic term, and “likely” an external/aleatory term

191 (Ülkümen et al. 2016). Second, the two terms differ in their *directionality* (Teigen and Brun
192 1995, 1999). While “uncertain” has a negative directionality (i.e., it points towards the
193 possibility that an outcome might not occur), “likely” has a positive directionality (i.e., it
194 points towards the possibility that an outcome might occur). To better control for the source
195 of uncertainty and directionality of the verbal probabilities used in the question, we used the
196 two terms “uncertain(ty)” and “certain(ty)”, which are usually considered as reflecting
197 epistemic uncertainty (Fox and Ülkümen 2011; Teigen and Løhre 2017; Ülkümen et al.
198 2016). The word stem was hence kept constant, while directionality and question type varied
199 between-subjects, with different groups of participants receiving the question about which
200 prediction “conveys more [un]certainty” and which prediction is “more [un]certain to be
201 correct”.

202 In Experiment 4, we removed the (arguably incorrect) “equal” option, so the
203 participants chose between the wide and the narrow interval in each condition. Participants
204 read the same temperature rise and sea level rise vignettes as in previous experiments in one
205 of three conditions: uncertainty conveyed, certainty conveyed, and certain to be correct.

206 In Experiment 5, we added a third prediction that featured a narrower interval to each
207 vignette, for two reasons: first, to highlight even more strongly that the teams differ in width
208 of prediction intervals; and second, since the intervals in previous experiments were both
209 quite wide, to include a very narrow interval that suggests high precision, but might be “too
210 good to be true”. Participants read the sea level and temperature rise scenarios and for each
211 selected one of the three forecasts as the one that conveyed more certainty, conveyed more
212 uncertainty or was more certain to be correct, in three between-subjects conditions.

213 2) SECONDARY HYPOTHESES

214 In addition to investigating the prevalence of the informativeness and accuracy
215 mindsets and their associations with different kinds of questions, Experiments 1-5 also

216 addressed some additional hypotheses. In Experiments 1 and 2, we investigated whether the
217 accuracy mindset would be seen as more appropriate (i.e., wide intervals associated with
218 certainty) in contexts where interval width could be related to variability. Predictions
219 concerning a class of multiple outcomes might induce more distributional (“outside view”)
220 thinking, with wide intervals reflecting external variability, in contrast to predictions of a
221 singular outcome, where wide intervals are more easily taken to reflect the forecaster’s
222 ignorance (Kahneman and Tversky 1982; Kahneman and Lovallo 1993; Nisbett et al. 1983;
223 Reeves and Lockhart 1993). Hence, participants in different conditions in Experiments 1
224 (within-subjects) and 2 (between-subjects) were told that the intervals described temperature
225 rise “in different countries” and sea level rise “in different parts of the world”, while no
226 explanation for the variability in the estimate was given in the other conditions.

227 In Experiment 3, we investigated whether perceptions of expertise could be influenced
228 by question type, with the hypothesis that questions highlighting informativeness would lead
229 to a stronger preference for experts giving narrow interval forecasts, as compared to questions
230 highlighting accuracy. Therefore, after selecting the prediction that conveys more
231 (un)certainty/is more (un)certain to be correct, participants in Experiment 3 rated which team
232 seemed more trustworthy, seemed to have most knowledge (about temperature rise or sea
233 level rise), seemed to have the best models (for predicting temperature rise or sea level rise),
234 and which team seemed to be most competent. These ratings were done on scales from 1
235 (definitely the team with the wide interval) to 5 (definitely the team with the narrow interval).

236 Experiment 4 investigated factors that might explain people’s preference for narrow
237 intervals: their fluency and the perceived expertise of the speaker. Previous research has
238 found that statements that are more fluent (i.e., easier to process), for example due to
239 repetition or to heightened visibility, are judged as more truthful than less fluent statements
240 (Arkes et al. 1989; Reber and Schwarz 1999). We expected that predictions with narrower

241 intervals might be easier to process than predictions with wider intervals, and that this
242 heightened fluency could be a reason why people prefer narrow intervals. Narrow intervals
243 might also be preferred due to the association between precision and expertise. Hence,
244 participants in Experiment 4 rated the fluency of the predictions featuring a narrow and a
245 wide interval, as well as the perceived expertise of the teams (see Supplementary materials for
246 more details about the rating scales).

247 For exploratory purposes, we included in Experiment 5 three measures of individual
248 differences that might be related to the degree of perception of wide intervals as more
249 uncertain and narrow intervals as more certain. Specifically, strong climate change beliefs
250 could explain a preference for wide intervals as certain, since wide intervals can incorporate
251 more extreme climate change values. In addition, people who are more numerate, and people
252 who are able to understand the probability of occurrence of more than one event (i.e., people
253 who correctly assess that the probability of one of two events is greater than the probability of
254 occurrence of each of those events), might be better able to appraise that a wider interval
255 means a greater likelihood to be correct. Hence, we included a climate change belief scale
256 (Heath and Gifford 2006), a numeracy scale (Lipkus et al. 2001), and a disjunction task
257 (adapted from Costello 2009).

258 **c. Results**

259 1) EFFECTS OF QUESTION FOCUS

260 Participants in Experiments 1-5 received wide and narrow interval forecasts of sea
261 level rise and temperature rise from two different (fictional) teams of climate scientists, and
262 indicated which interval *conveyed more (un)certainty* (question focused on informativeness)
263 or was *more likely [(un)certain] to be correct* (question focused on accuracy).

264

265 <Insert Figures 1, 2, and 3 about here>

266

267 Question focus strongly influenced certainty judgments (Figures 1 and 2). Participants
268 largely chose the wide interval as the one that conveyed more uncertainty, and indicated that
269 the narrow interval conveyed more certainty. Responses to questions about which interval
270 was more likely or more certain to be correct were mixed: some experiments showed a small
271 preference for the wide interval, while narrow and wide intervals were seen as equally certain
272 in other experiments.

273 Figure 3 summarizes the overall results (for all experiments with three response
274 options, i.e., all experiments except Experiment 4), with responses coded according to
275 whether wide intervals are seen as more certain (consistent with the accuracy mindset),
276 narrow intervals are seen as more certain (consistent with the informativeness mindset), or
277 both intervals are seen as equally likely. Analysis of Experiments 2, 3 and 5, where question
278 focus was varied between-subjects and three response alternatives (wide more certain, narrow
279 more certain, equal/"medium" interval more certain) were provided, showed a clear effect of
280 question focus, $\chi^2(2, N=1080) = 213.373, p < .001$. While wide intervals were clearly
281 associated with uncertainty after informativeness-focused questions, more participants
282 associated wide intervals with certainty after accuracy-focused questions. However, even for
283 questions about correctness, where wide intervals should logically be chosen as more certain,
284 only about 40% of the participants did so.

285 2) RESULTS FOR SECONDARY HYPOTHESES

286 In Experiments 1 and 2, we investigated whether giving people an explanation for
287 variability, for instance by telling them that the forecasts concerned sea level rise "in different
288 parts of the world", would facilitate the accuracy mindset (i.e., would make more people
289 associate wide intervals with certainty). However, this hint about variability did not affect
290 participants' interval choice in either Experiment 1 ($p = .150$) or Experiment 2 ($p = .303$).

291 We further examined whether the accuracy and informativeness mindsets led to
292 different inferences about the forecaster. Participants in Experiment 3 rated whether they
293 found teams giving wide or narrow interval forecasts to have more expertise, on scales from 1
294 (definitely the team with the wide interval) to 5 (definitely the team with the narrow interval).
295 The average of the ratings of the experts across scenarios (i.e., an average of the four
296 questions per scenario) were slightly higher in the “conveys more”-conditions ($M = 3.50$, SD
297 $= .73$) than in the “to be correct”-conditions ($M = 3.29$, $SD = .87$), and this difference was
298 significant, $F(1,234) = 3.991$, $p = .047$, $\eta^2_p = .017$. In other words, the team with narrow
299 intervals was rated more positively after informativeness-focused questions, indicating that
300 making one or the other mindset salient can influence how well both the prediction and the
301 communicator is received.

302 Experiment 4 investigated whether people find narrow intervals easier to process (i.e.,
303 more fluent) and more related to expertise than wide intervals. As predicted, participants
304 judged the narrow interval as being easier to process and as reflecting more expertise than the
305 wide interval (see Supplementary materials for more details about these findings).

306 Finally, in Experiment 5 we set out to investigate individual differences that might be
307 related to the preference for informativeness vs. accuracy. Specifically, we asked participants
308 about their climate change beliefs, and gave them a test measuring numeracy, and a test
309 measuring their understanding of disjunctive probabilities. However, there were no clear
310 correlation patterns between interval choice and any of these three measures across groups,
311 and the experiment did not have enough power to detect differences within each condition.

312

313 **3. Experiment 6: Is it more intuitive to associate wide intervals with uncertainty**
314 **than with certainty?**

315 Experiments 1-5 demonstrated that different question focus promotes different views
316 about the relationship between certainty and interval width. However, the fact that only about
317 40% endorsed wide intervals as “more certain to be correct”, indicates that it is more common
318 to associate wide intervals with (subjective) uncertainty than with (objective) certainty. This
319 raises the possibility that the lay view about the relationship between interval width and
320 certainty is more in line with the informativeness mindset than with the accuracy mindset.

321 In support of this idea, research on confidence intervals has repeatedly shown that
322 people produce intervals that are too narrow for the assigned degree of certainty (Moore et al.
323 2016). This consistent overprecision (Moore and Healy 2008) is very hard to eliminate and
324 suggests that the preference for informativeness may be a dominant intuitive response.
325 Studies showing that recipients of information in general prefer narrow intervals illustrate a
326 similar point (Du et al. 2011; Jørgensen 2016; McKenzie and Amin 2002; Yaniv and Foster
327 1995), as does the preliminary finding that people with higher numeracy can (sometimes)
328 better appreciate the trade-off between precision and certainty than those with lower
329 numeracy (Løhre and Teigen 2017). Hence, we ran Experiment 6 to test the hypothesis of an
330 intuitive preference for informativeness among lay people.

331 **a. Materials and procedure**

332 The opening paragraph of the survey in Experiment 6 explained that climate scientists
333 sometimes use intervals when giving their predictions of future outcomes, and presented two
334 predictions concerning the expected sea level rise in the Oslo fjord. One of the predictions
335 contained a wide interval (minimum 20 and maximum 60 cm sea level rise) and the other
336 prediction contained a narrow interval (minimum 30 and maximum 50 cm sea level rise).
337 Participants (students at the University of Oslo, N = 105, see Table 1) were randomly
338 assigned to either the wide condition, where it was pointed out that one prediction is wider

339 than the other, or to the narrow condition, where it was pointed out that one prediction is
340 narrower than the other.

341 The text then explained that there are two different ways that one can think about the
342 relationship between interval width and uncertainty, using the following formulation in the
343 wide condition:

344 “ – On the one hand, WIDE intervals indicate that it is MORE UNCERTAIN what the
345 outcome will be (the sea level could rise by anything from 20 to 60 cm, compared to 30 to
346 50 cm for the narrow interval)
347 - On the other hand, it is MORE CERTAIN that projections using WIDE intervals will be
348 correct (the forecast is correct if the sea level rises by anything from 20 to 60 cm,
349 compared to 30 to 50 cm for the narrow interval)”

350 In other words, the accuracy mindset (seeing the wide interval as more certain to be
351 correct) and the informativeness mindset (seeing the wide interval as indicating that it is more
352 uncertain what the outcome will be) were explained to the participants. In the narrow
353 condition, the text explained that narrow intervals could be seen as indicating that it is more
354 certain what the outcome will be, or that it is more uncertain that predictions using narrow
355 intervals will be correct. The order of the statements was counterbalanced in both conditions.

356 After reading the description of the different ways of thinking about intervals and
357 uncertainty, participants were asked to rate how intuitive, natural, appealing, logical, and
358 complicated they found the two ways of thinking, on scales from 1 (not intuitive/natural etc.
359 at all) to 7 (very intuitive/natural etc.). Next, the participants were given tests of numeracy
360 (Cokely et al. 2012; Schwartz et al. 1997) and cognitive reflection (Frederick 2005) to see
361 whether individual differences in these abilities were related to a preference for
362 informativeness or accuracy. Finally, participants were asked if they had already seen or
363 responded to the cognitive reflection test online or in other experiments.

364

365

<Insert Figures 4 and 5 about here>

366

367

b. Results

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

Figures 4 and 5 display the ratings of the different mindsets for both wide and narrow intervals, and show that the view that wide intervals convey uncertainty was judged as more intuitive, natural, appealing, logical, and less complicated than the view that wide intervals are more certain to be correct. For simplicity we refer to this combination of attributes as more “intuitively appealing”. We also computed an average difference score to measure the degree to which one “mindset” was judged as more intuitively appealing than the other, by taking the “wide = uncertain” and “narrow = certain” ratings, which are in line with the informativeness mindset, and subtracting the corresponding “wide = certain” and “narrow = uncertain” ratings, which are in line with the accuracy mindset.² Thus, positive difference scores indicate that the informativeness mindset is seen as more intuitively appealing than the accuracy mindset. The average difference score for the five items (Cronbach’s $\alpha = .74$) did not differ between conditions, $F(1,103) = .144, p = .706, \eta^2_p = .001$. More interestingly, the average difference score across conditions was positive, $M = .42, SD = 1.32$, and differed significantly from 0, $t(104) = 3.290, p = .001, 95\% CI [.17, .68]$. Hence, participants overall judged the informativeness mindset as more intuitively appealing than the accuracy mindset.

There was no significant correlation with the average difference score for either the cognitive reflection test ($r = .01, p = .958$) or numeracy ($r = .09, p = .355$). However, people with higher cognitive reflection and numeracy perceived *both* mindsets as more intuitive, as shown by positive correlations between CRT and the informativeness ($r = .20, p = .040$) and accuracy mindsets ($r = .21, p = .037$), and between numeracy and the informativeness ($r =$

² The only exception was for the ratings of how complicated the participants found the two ways of thinking. Here the “wide = uncertain” ratings were subtracted from the “wide = certain” ratings, and the “narrow = certain” ratings were subtracted from the “narrow = uncertain” ratings.

388 .24, $p = .014$) and accuracy mindsets ($r = .14$, $p = .161$). Hence, higher scores on these
389 measures indicate a tendency to find it intuitive to use intervals to express both certainty and
390 uncertainty.

391

392 **4. General Discussion**

393 The experiments reported in this paper fill a gap in the literature about climate change
394 communication (Moser 2010; Pidgeon and Fischhoff 2011) by investigating lay perceptions
395 of the relationship between interval width (forecast precision) and certainty. We found
396 evidence of two alternative ways of thinking. Overall, independent of question focus, 45% of
397 our participants³ perceived narrow intervals as giving more certain knowledge about what the
398 outcome will be, in line with what we have called a *preference for informativeness*; while
399 26% of the participants perceived that wide intervals have a higher certainty of capturing the
400 true value, displaying a *preference for accuracy*. These two opposite “mindsets” can be made
401 more or less salient by drawing attention to different types of uncertainty. Questions about
402 which interval conveys more (un)certainty (i.e., focusing more on subjective uncertainty) led
403 to a consistent preference for informativeness, while questions about which interval is more
404 certain/likely to be correct (i.e., focusing more on objective certainty) led to a response
405 pattern more in line with the accuracy mindset.

406 Questions focused on informativeness led to a clearer response pattern (wide intervals
407 seen as uncertain and narrow ones as certain) than did questions focused on accuracy. It is
408 somewhat puzzling that people were so divided in their answers to the question about which
409 interval is more likely/certain to be correct. Logically, wider intervals are objectively more
410 likely to capture the outcome value that will occur, as they cover both central (likely) and
411 more peripheral (unlikely) values. Our results indicate that (perhaps for good reasons) people

³ These percentages are based on all experiments with three response alternatives (wide more certain, narrow more certain, equal), i.e., Experiments 1, 2, 3, and 5.

412 would like to know more precisely what the expected values are, and hence find it more
413 intuitive to adopt the informativeness than the accuracy mindset, as shown in Experiment 6.
414 Although the generalizability of the results should be investigated in non-western samples, we
415 find it noteworthy that they are replicated in two different languages (Norwegian vs. English),
416 in three different countries (Norway, UK, USA), and with both student and MTurk samples.
417 Note also that our participants should be more educated and arguably more knowledgeable
418 about these topics than more representative samples. Hence, one might expect an even
419 stronger preference for informativeness in a more representative sample.

420 These results have important theoretical implications, particularly for the literature on
421 overprecision (Moore et al. 2016). The intuitive preference for informativeness means that
422 wide intervals are usually associated with uncertainty, and as a result, people may not
423 understand or agree that they should widen their intervals to increase their certainty. This can
424 be said to strengthen the conversational norms/informativeness account of overprecision
425 (Kaesler et al. 2016; Yaniv and Foster 1995, 1997).

426 Climate scientists may choose to give wide intervals in order to present predictions
427 with high certainty. Yet, our results show that wide intervals are a stronger signal of
428 (subjective) uncertainty than of (objective) certainty, and the use of wide intervals may
429 therefore undermine trust in climate scientists and their predictions. Although language that
430 accentuates the accuracy mindset may make wide intervals more acceptable to the public (see
431 Experiment 3), our results suggest that many recipients will still prefer narrow intervals, as
432 suggested by 25% of the participants given accuracy-focused questions in our experiments
433 (see Figure 3). Note, however, that in the current experiments, the participants only received
434 intervals, and were asked about their perceptions of (un)certainty. In statements from the
435 IPCC, intervals are often accompanied by verbal or numerical probability statements (e.g.,
436 “During the last interglacial period, the Greenland ice sheet *very likely* contributed between

437 1.4 and 4.3 m to higher global mean sea level”) (IPCC 2013). A recent study showed that
438 explicitly mentioning the high certainty of wide intervals can counteract the tendency of lay
439 people to see such intervals as uncertain, with most people stating that a wide interval with
440 90% probability was more certain than a narrow interval with 50% probability (Teigen et al.
441 2018).

442 Nevertheless, the current evidence gives reason to be skeptical about the use of wide
443 intervals to achieve high certainty in statements about climate change. However, presenting a
444 precise interval along with a statement about the low certainty of such an interval is arguably
445 not a much better option. One compromise solution would be to provide two intervals rather
446 than one: a narrow (informative) interval paired with a wide (confident) interval, to satisfy
447 both camps of readers. The drawback is that presenting two intervals simultaneously adds
448 complexity to the communication of an already complex topic. Using graphical
449 representations could be useful to simultaneously communicate informativeness and accuracy
450 in a relatively simple way (Spiegelhalter et al. 2011). In any case, communicators should be
451 aware that the current practice of claiming to be very certain about a very wide interval will to
452 many readers sound like a contradiction in terms, which might damage rather than strengthen
453 the public’s belief in climate science.

454

455 **Acknowledgements**

456 This research was supported by grant no.235585/E10 from The Research Council of
457 Norway. The authors report no conflicts of interest. All datafiles and materials are available
458 on OSF (https://osf.io/h95cj/?view_only=3ddfde9204f3497eaebf42aa0c091ade).

459 **References**

- 460
- 461 Arkes, H. R., C. Hackett, and L. Boehm, 1989: The generality of the relation between
- 462 familiarity and judged validity. *Journal of Behavioral Decision Making*, **2**, 81-94.
- 463 Budescu, D. V., S. Broomell, and H. H. Por, 2009: Improving communication of uncertainty
- 464 in the reports of the Intergovernmental Panel on Climate Change. *Psychological Science*, **20**,
- 465 299-308.
- 466 Budescu, D. V., H. H. Por, and S. B. Broomell, 2012: Effective communication of uncertainty
- 467 in the IPCC reports. *Climatic Change*, **113**, 181-200.
- 468 Budescu, D. V., H.-H. Por, S. B. Broomell, and M. Smithson, 2014: The interpretation of
- 469 IPCC probabilistic statements around the world. *Nature Climate Change*, **4**, 508-512.
- 470 Cokely, E. T., M. Galesic, E. Schulz, S. Ghazal, and R. Garcia-Retamero, 2012: Measuring
- 471 risk literacy: The Berlin Numeracy Test. *Judgment and Decision Making*, **7**, 25-47.
- 472 Costello, F. J., 2009: Fallacies in probability judgments for conjunctions and disjunctions of
- 473 everyday events. *Journal of Behavioral Decision Making*, **22**, 235-251.
- 474 Dieckmann, N. F., E. Peters, and R. Gregory, 2015: At home on the range? Lay
- 475 interpretations of numerical uncertainty ranges. *Risk Anal*, **35**, 1281-1295.
- 476 Dieckmann, N. F., R. Gregory, E. Peters, and R. Hartman, 2017: Seeing what you want to
- 477 see: How imprecise uncertainty ranges enhance motivated reasoning. *Risk Anal*, **37**, 471-486.
- 478 Du, N., D. V. Budescu, M. K. Shelly, and T. C. Omer, 2011: The appeal of vague financial
- 479 forecasts. *Organizational Behavior and Human Decision Processes*, **114**, 179-189.
- 480 Fox, C. R., and G. Ülkümen, 2011: Distinguishing two dimensions of uncertainty.
- 481 *Perspectives on Thinking, Judging, and Decision Making*, W. Brun, G. Keren, G. Kirkebøen,
- 482 and H. Montgomery, Eds., Universitetsforlaget, 21-35.
- 483 ———, 2017: Comment on Løhre & Teigen. *Thinking & Reasoning*, **23**, 483-491.

484 Frederick, S., 2005: Cognitive reflection and decision making. *Journal of Economic*
485 *Perspectives*, **19**, 25-42.

486 Goodman, J. K., C. E. Cryder, and A. Cheema, 2013: Data collection in a flat world: The
487 strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision*
488 *Making*, **26**, 213-224.

489 Grice, H. P., 1975: Logic and conversation. *Syntax and semantics, Vol. 3: Speech acts*, P.
490 Cole, and J. L. Morgan, Eds., Academic Press, 41-58.

491 Hacking, I., 1975: *The emergence of probability: A philosophical study of early ideas about*
492 *probability, induction and statistical inference*. Cambridge University Press.

493 Harris, A. J. L., and A. Corner, 2011: Communicating environmental risks: Clarifying the
494 severity effect in interpretations of verbal probability expressions. *Journal of Experimental*
495 *Psychology-Learning Memory and Cognition*, **37**, 1571-1578.

496 Harris, A. J. L., H. H. Por, and S. B. Broomell, 2017: Anchoring climate change
497 communications. *Climatic Change*, **140**, 387-398.

498 Harris, A. J. L., A. Corner, J. M. Xu, and X. F. Du, 2013: Lost in translation? Interpretations
499 of the probability phrases used by the Intergovernmental Panel on Climate Change in China
500 and the UK. *Climatic Change*, **121**, 415-425.

501 Heath, Y., and R. Gifford, 2006: Free-market ideology and environmental degradation - The
502 case of belief in global climate change. *Environment and Behavior*, **38**, 48-71.

503 Ho, E. H., D. V. Budescu, M. K. Dhami, and D. R. Mandel, 2015: Improving the
504 communication of uncertainty in climate science and intelligence analysis. *Behavioral*
505 *Science & Policy*, **1**, 43-55.

506 IPCC, 2013: Summary for policymakers. *Climate change 2013: The physical science basis.*
507 *Contribution of Working Group I to the fifth assessment report of the Intergovernmental*
508 *Panel on Climate Change*, T. F. Stocker, and Coauthors, Eds., Cambridge University Press.

509 Joslyn, S. L., and J. E. LeClerc, 2016: Climate projections and uncertainty communication.
510 *Topics in Cognitive Science*, **8**, 222-241.

511 Juanchich, M., and M. Sirota, 2017: How much will the sea level rise? Outcome selection and
512 subjective probability in climate change predictions. *Journal of Experimental Psychology:*
513 *Applied*, **23**, 386-402.

514 Juanchich, M., A. Gourdon-Kanhukamwe, and M. Sirota, 2017: «I am uncertain» vs «it is
515 uncertain». How linguistic markers of the uncertainty source affect uncertainty
516 communication. *Judgment and Decision Making*, **12**, 445-465.

517 Jørgensen, M., 2016: The use of precision of software development effort estimates to
518 communicate uncertainty. *Software Quality. The Future of Systems- and Software*
519 *Development*, D. Winkler, S. Biffl, and J. Bergsmann, Eds., Springer International Publishing,
520 156-168.

521 Kaesler, M., M. B. Welsh, and C. Semmler, 2016: Predicting overprecision in range
522 estimation. *Proceedings of the 38th Annual Meeting of the Cognitive Science Society*, Austin,
523 TX, Cognitive Science Society, 502-507.

524 Kahneman, D., and A. Tversky, 1982: Variants of uncertainty. *Cognition*, **11**, 143-157.

525 Kahneman, D., and D. Lovallo, 1993: Timid choices and bold forecasts: A cognitive
526 perspective on risk taking. *Management Science*, **39**, 17-31.

527 Lipkus, I. M., G. Samsa, and B. K. Rimer, 2001: General performance on a numeracy scale
528 among highly educated samples. *Med Decis Making*, **21**, 37-44.

529 Løhre, E., and K. H. Teigen, 2016: There is a 60% probability, but I am 70% certain:
530 Communicative consequences of external and internal expressions of uncertainty. *Thinking &*
531 *Reasoning*, **22**, 369-396.

532 ———, 2017: Probabilities associated with precise and vague forecasts. *Journal of Behavioral*
533 *Decision Making*, **30**, 1014-1026.

534 McKenzie, C. R. M., and M. B. Amin, 2002: When wrong predictions provide more support
535 than right ones. *Psychonomic Bulletin & Review*, **9**, 821-828.

536 McKenzie, C. R. M., M. J. Liersch, and I. Yaniv, 2008: Overconfidence in interval estimates:
537 What does expertise buy you? *Organizational Behavior and Human Decision Processes*, **107**,
538 179-191.

539 Moore, D. A., and P. J. Healy, 2008: The trouble with overconfidence. *Psychological Review*,
540 **115**, 502-517.

541 Moore, D. A., E. R. Tenney, and U. Haran, 2016: Overprecision in judgment. *Handbook of*
542 *Judgment and Decision Making*, G. Wu, and G. Keren, Eds., Wiley.

543 Moser, S. C., 2010: Communicating climate change: history, challenges, process and future
544 directions. *Wiley Interdisciplinary Reviews-Climate Change*, **1**, 31-53.

545 Nisbett, R. E., D. H. Krantz, C. Jepsen, and Z. Kunda, 1983: The use of statistical heuristics
546 in everyday inductive reasoning. *Psychological Review*, **90**, 339-363.

547 Paolacci, G., J. Chandler, and P. G. Ipeirotis, 2010: Running experiments on Amazon
548 Mechanical Turk. *Judgment and Decision Making*, **5**, 411-419.

549 Pidgeon, N., and B. Fischhoff, 2011: The role of social and decision sciences in
550 communicating uncertain climate risks. *Nature Climate Change*, **1**, 35-41.

551 Reber, R., and N. Schwarz, 1999: Effects of perceptual fluency on judgments of truth.
552 *Consciousness and Cognition*, **8**, 338-342.

553 Reeves, T., and R. S. Lockhart, 1993: Distributional versus singular approaches to probability
554 and errors in probabilistic reasoning. *Journal of Experimental Psychology-General*, **122**, 207-
555 226.

556 Schwartz, L. M., S. Woloshin, W. C. Black, and H. G. Welch, 1997: The role of numeracy in
557 understanding the benefit of screening mammography. *Annals of Internal Medicine*, **127**, 966-
558 972.

559 Spiegelhalter, D., M. Pearson, and I. Short, 2011: Visualizing uncertainty about the future.
560 *Science*, **333**, 1393-1400.

561 Teigen, K. H., 1990: To be convincing or to be right: A question of preciseness. *Lines of*
562 *Thinking*, K. J. Gilhooly, M. T. G. Keane, R. H. Logie, and G. Erdos, Eds., Wiley.

563 Teigen, K. H., and W. Brun, 1995: Yes, but it is uncertain: Direction and communicative
564 intention of verbal probabilistic terms. *Acta Psychologica*, **88**, 233-258.

565 ———, 1999: The directionality of verbal probability expressions: Effects on decisions,
566 predictions, and probabilistic reasoning. *Organizational Behavior and Human Decision*
567 *Processes*, **80**, 155-190.

568 Teigen, K. H., and E. Løhre, 2017: Expressing (un)certainty in no uncertain terms: Reply to
569 Fox and Ülkümen. *Thinking & Reasoning*, **23**, 492-496.

570 Teigen, K. H., E. Løhre, and S. M. Hohle, 2018: The boundary effect: Perceived post hoc
571 accuracy of prediction intervals. *Judgment and Decision Making*, **13**, 309-321.

572 Yaniv, I., and D. P. Foster, 1995: Graininess of judgment under uncertainty: An accuracy-
573 informativeness trade-off. *Journal of Experimental Psychology: General*, **124**, 424-432.

574 ———, 1997: Precision and accuracy of judgmental estimation. *Journal of Behavioral Decision*
575 *Making*, **10**, 21-32.

576 Ülkümen, G., C. R. Fox, and B. F. Malle, 2016: Two dimensions of subjective uncertainty:
577 Clues from natural language. *Journal of Experimental Psychology: General*, **145**, 1280-1297.

578

579
580
581
582

Tables

Table 1. Demographics for the samples used in the different experiments.

Experiment no.	n	Sample	Mean age (SD)	Female	Male
1	81	University of Essex students	24.0 (6.5)	80.2%	19.8%
2	201	Amazon Mechanical Turk	37.9 (12.0)	51.7%	48.3%
3	238	Amazon Mechanical Turk	37.7 (11.2)	47.9%	52.1%
4	302	Amazon Mechanical Turk	34.6 (10.4)	44.4%	55.6%
5	101	University of Essex, snowball sampling	28.0 (13.1)	36.6%	62.4%
6	105	University of Oslo students	23.1 (4.9)	76.2%	23.8%

583
584

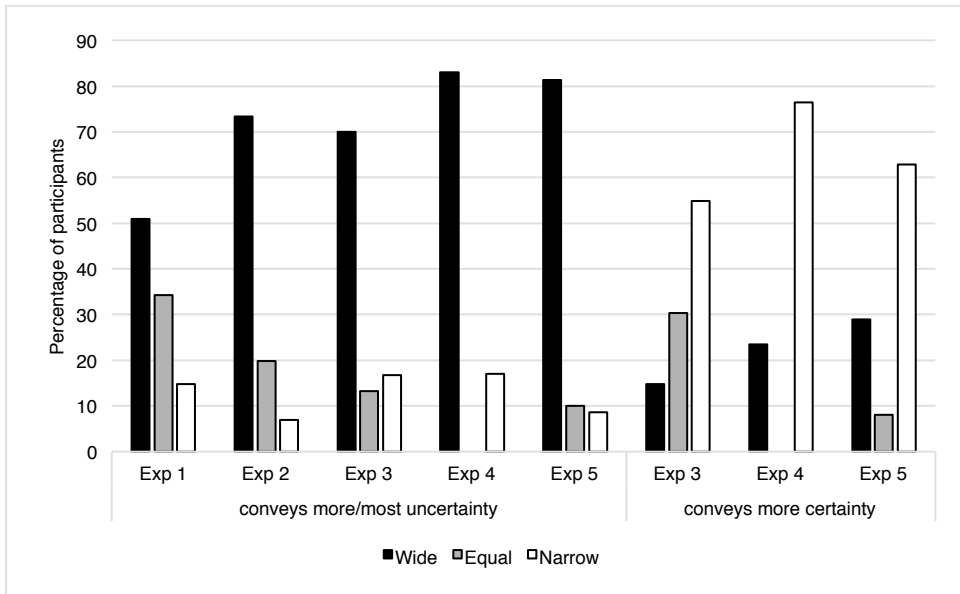
585 Table 2. Overview of questions, response options and design used in the different experiments
 586 regarding interval predictions of climate change outcomes.
 587

Experiment no.	Question(s)/ statements focused on informativeness: “Which interval conveys...”	Question(s)/statements focused on accuracy: “Which interval is...”	Response options	Design
1	“...most uncertainty”	“... most likely to be correct”	Wide, narrow, equal	Within-subjects
2	“... most uncertainty”	“... most likely to be correct”	Wide, narrow, equal	Between-subjects
3	“... more uncertainty” “... more certainty”	“... more certain to be correct” “... more uncertain to be correct”	Wide, narrow, equal	Between-subjects
4	“... more uncertainty” “... more certainty”	“... more certain to be correct”	Wide, narrow	Between-subjects
5	“... more uncertainty” “... more certainty”	“... more certain to be correct”	Wide, “medium”, narrow	Between-subjects
6	“Wide intervals indicate that it is more uncertain what the outcome will be” “Narrow intervals indicate that it is more certain what the outcome will be”	“It is more certain that projections using wide intervals will be correct” “It is more uncertain that projections using narrow intervals will be correct”	Ratings of the intuitive appeal of both statements	Within-subjects

588

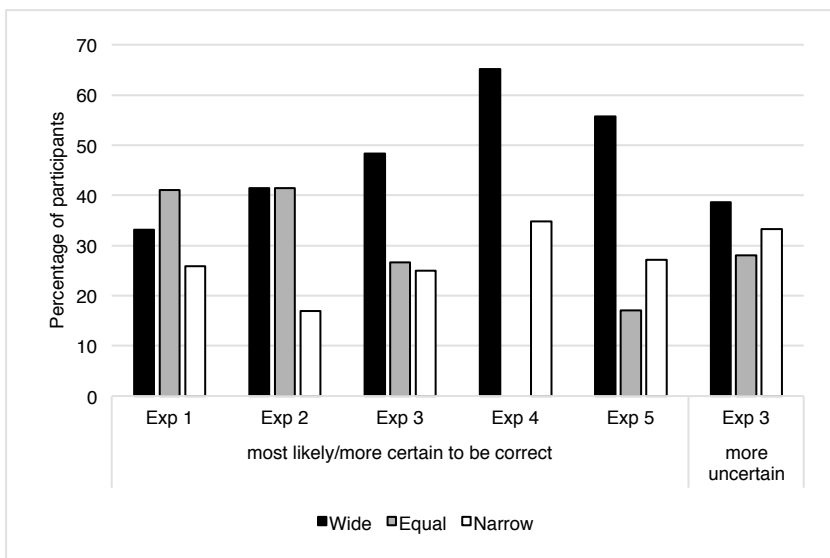
589
590

Figures



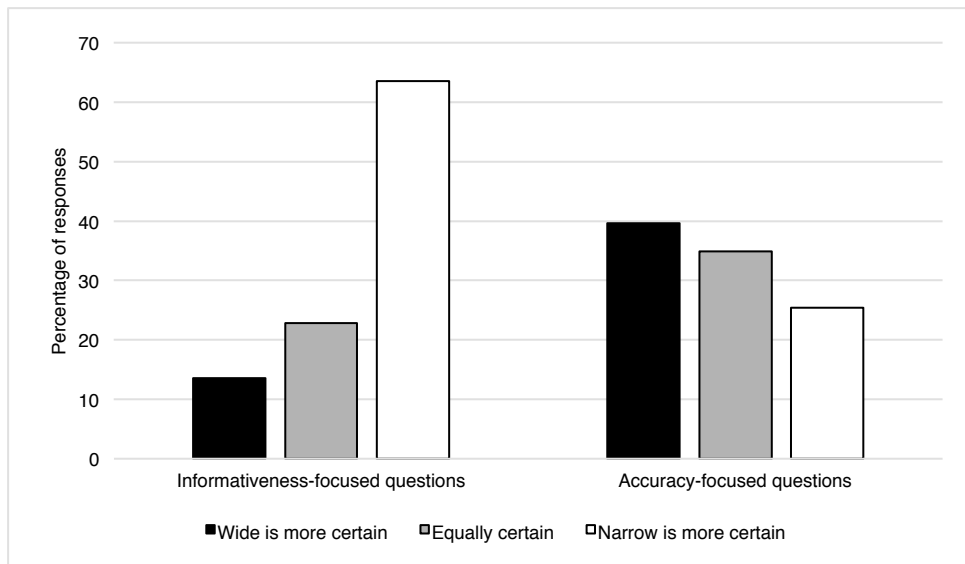
591
592
593
594
595

Figure 1. Choices of which interval conveys more certainty and uncertainty.



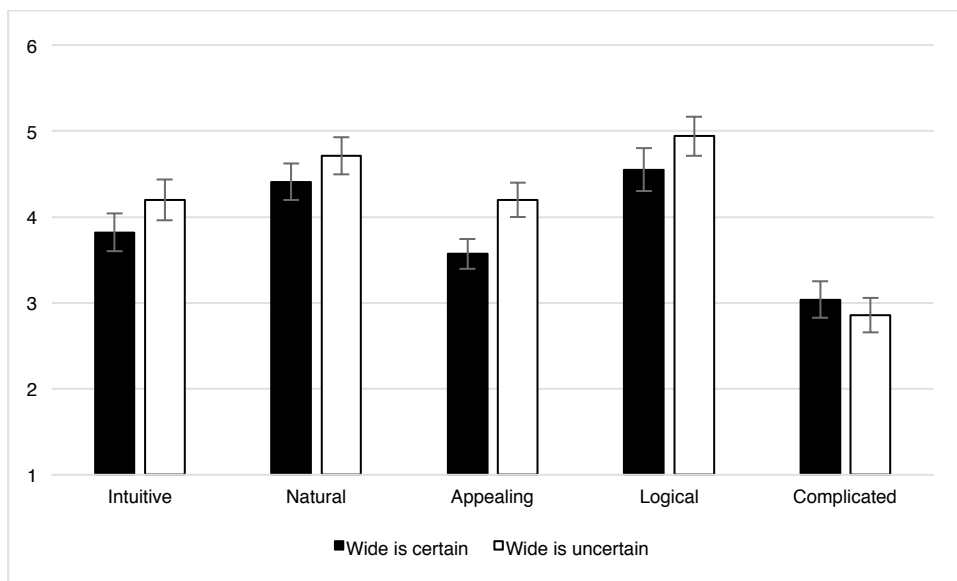
596
597
598
599
600
601

Figure 2. Choices of which interval is more certain/likely and more uncertain to be correct.



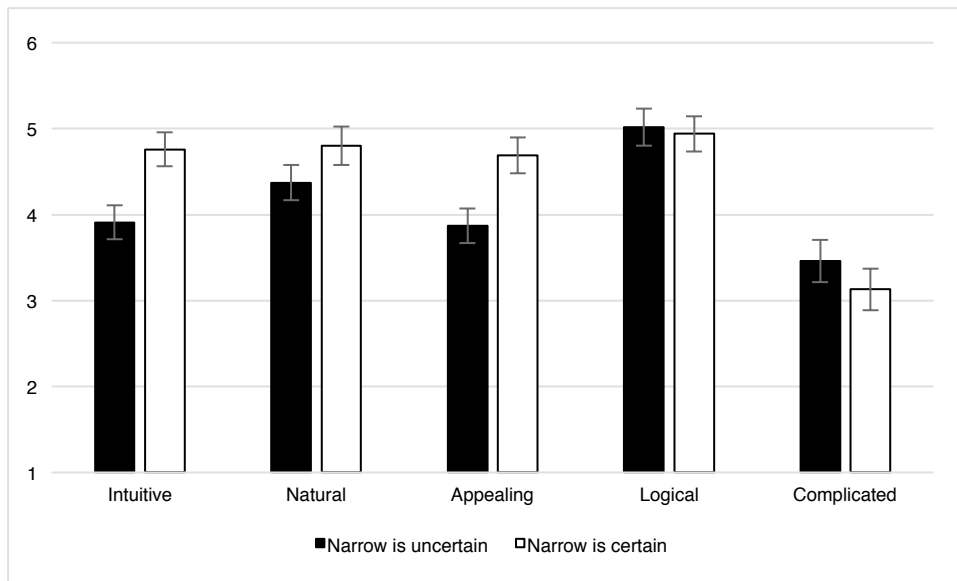
602
603
604
605
606
607
608

Figure 3. Overall preference for wide vs. narrow intervals as “more certain” for all experiments with three response options (Experiments 1, 2, 3, and 5).



609
610
611
612
613
614
615

Figure 4. Mean perceptions of two ways of thinking about wide intervals (wide is certain vs. wide is uncertain) in Experiment 6, error bars +/- 1 SEM.



616
 617
 618
 619
 620

Figure 5. Mean perceptions of two ways of thinking about narrow intervals (narrow is uncertain vs. narrow is certain) in Experiment 6, error bars +/- 1 SEM.