

LETTER • OPEN ACCESS

## When IPCC graphs can foster or bias understanding: evidence among decision-makers from governmental and non-governmental institutions

To cite this article: Helen Fischer *et al* 2020 *Environ. Res. Lett.* **15** 114041

View the [article online](#) for updates and enhancements.

# Environmental Research Letters



## LETTER

### OPEN ACCESS

RECEIVED  
28 February 2020

REVISED  
10 June 2020

ACCEPTED FOR PUBLICATION  
28 September 2020

PUBLISHED  
16 November 2020

Original content from this work may be used under the terms of the [Creative Commons Attribution 4.0 licence](#).

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



# When IPCC graphs can foster or bias understanding: evidence among decision-makers from governmental and non-governmental institutions

Helen Fischer<sup>1,7</sup> , Karlijn L van den Broek<sup>2,3</sup> , Kilian Ramisch<sup>4,5</sup> and Yasmina Okan<sup>6</sup> 

<sup>1</sup> Stockholm Resilience Centre, Stockholm University, Stockholm, Sweden

<sup>2</sup> Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, The Netherlands

<sup>3</sup> Research Centre for Environmental Economics, Heidelberg University, Heidelberg, Germany

<sup>4</sup> Institute of Psychology, Heidelberg University, Heidelberg, Germany

<sup>5</sup> Department for Economics & Policy, Technical University Munich, Munich, Germany

<sup>6</sup> Centre for Decision Research, Leeds University Business School, University of Leeds, Leeds, United Kingdom

<sup>7</sup> Author to whom any correspondence should be addressed.

E-mail: [helen.fischer@su.se](mailto:helen.fischer@su.se)

**Keywords:** climate change, IPCC graphs, objective and subjective comprehension, metacognition

Supplementary material for this article is available [online](#)

## Abstract

To develop effective climate change policy, decision-makers need to have the best possible understanding of the available climate science. The IPCC Assessment Reports therefore aim to lay the foundation for informed political decision-making by providing policy-relevant information. But how successful are IPCC reports at communicating key findings? Although IPCC reports display key information in graphs, the interpretation of such graphs has received little attention. Here we provide an empirical evaluation of IPCC graph comprehension among IPCC target audience (N = 110), (political) decision-makers from climate-related (non-)governmental organizations from 54 countries, and a comparative sample of German junior diplomats, representing future international decision-makers (N = 33). We assess comprehension of current climate change risk visualizations using two IPCC graphs, one that employs principles of intuitive design, and one that violates principles of intuitive design. Results showed that (i) while a minority of IPCC target audience misinterpreted the intuitive graph, (ii) the majority of participants systematically misinterpreted the counter-intuitive graph, drawing the opposite conclusion from what was meant to be conveyed by the graph, despite (iii) having high confidence in the accuracy of their interpretation. Since misinterpretation of IPCC graphs does not allow for optimal use of the scientific information for policy-making, the results emphasize the importance of IPCC graphs that follow the principles of intuitive design.

## 1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) was founded by the World Meteorological Organisation and the United Nations Environment Programme to provide policymakers with regular scientific assessments on climate change, and its risks. One of the main principles of the IPCC Assessment Reports is to lay the foundation for informed political decision-making by providing policy-relevant information. But how successful are IPCC reports at communicating key findings? Within IPCC reports, key findings and major results are often displayed

in graphs. Despite the importance of these graphs, research examining how people interpret climate data visualizations is limited, and has largely focused on either objective or subjective comprehension and preferences for different designs (McMahon *et al* 2015, 2016; Taylor *et al* 2015). However, it is largely unknown how well objective and subjective understanding align: Do viewers have insight into the degree of their understanding of IPCC graphs? Can they correctly indicate which graphs they do, and do not, understand, and does that vary with the design of the graph? There is substantial evidence that graphs that are not well-designed often

lead to systematic misunderstandings and even sub-optimal decisions (e.g., Okan *et al* 2012, 2016), while difficulties in understanding can be significantly reduced by depicting risk using simple, well-designed, graphical displays (e.g. Garcia-Retamero and Cokely 2013). Here we examined objective and subjective understanding of IPCC graphs in a sample of IPCC target audience, political decision-makers from climate-related governmental organizations and staff from climate-related non-governmental organizations.

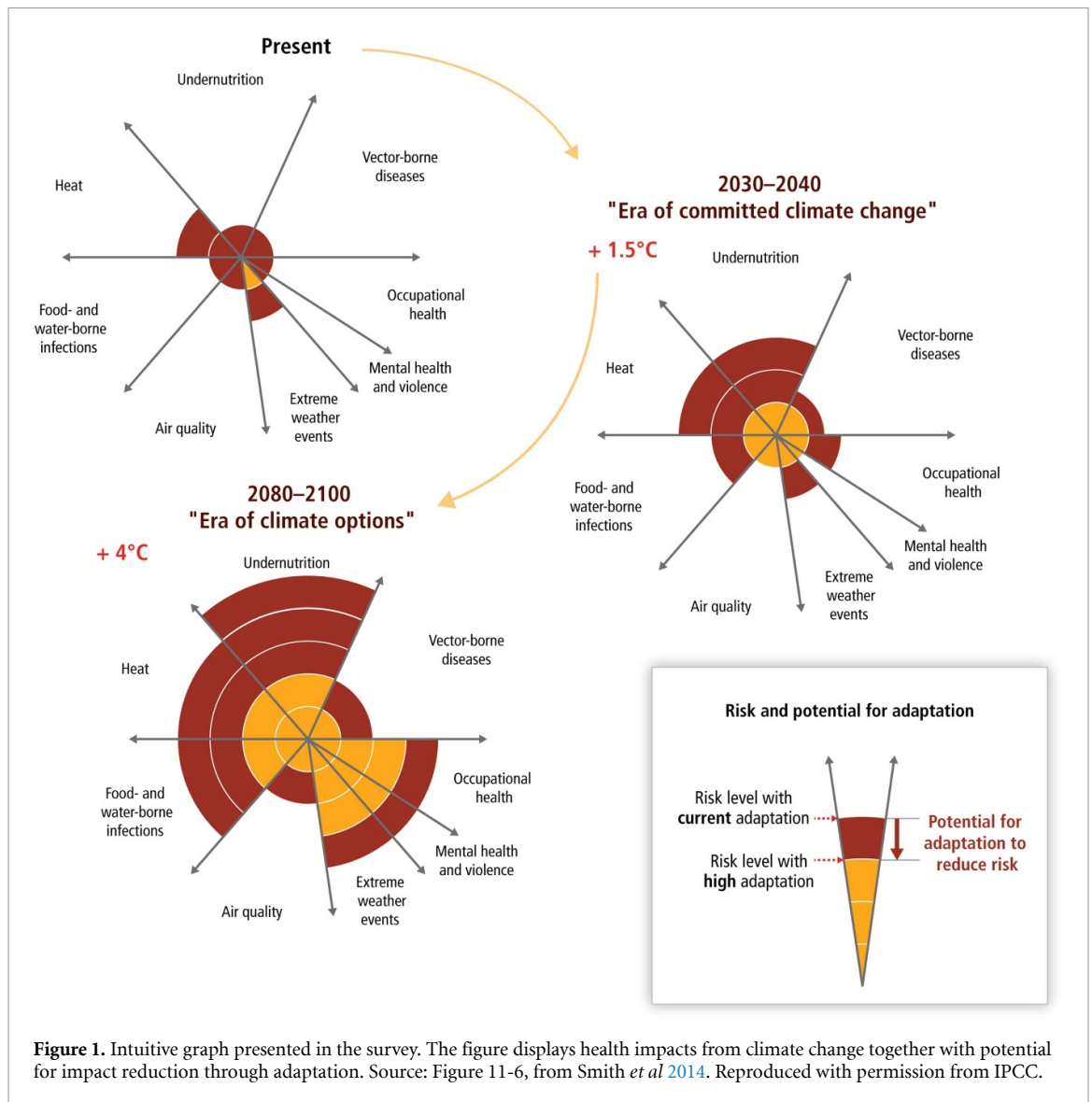
Graphs can have important cognitive and motivational advantages, such as reducing the cognitive effort required to understand risks or information processing time (Smerecnik *et al* 2010, Brewer *et al* 2012). Graphs can also reveal data patterns that may otherwise go undetected and evoke automatic mathematical operations (Lipkus 2007). However, the effectiveness of graphs to improve understanding depends crucially on how they are designed. An important design principle to support understanding is that spatial and conventional features in graphs should convey the same meaning (Okan *et al* 2012, 2016). Spatial features include bars of different heights or lines following a trend, whereas conventional features are graph elements linked to arbitrary graph conventions, such as green denoting gains, and red denoting losses. When spatial features are incongruent with conventional features (e.g. higher bars represent lower values), people tend to misinterpret graphs as they often fail to detect such incongruencies. Instead, people often rely on ‘spatial-to-conceptual’ mappings such as ‘high equals more’ or ‘steeper equals faster’ to interpret graphs, particularly if they have low graph literacy—the ability to understand graphically presented information (Okan *et al* 2012, 2016).

Another important principle for effective design is to keep graphs as simple as possible, ideally focusing on one main message per graph (Garcia-Retamero and Cokely 2017, Kause *et al* 2020); or only information that is essential for its intended purpose (Kosslyn 2006, Harold *et al* 2019). Graph complexity may also hinder viewers’ insight into their objective comprehension (Fischer *et al* 2018). However, whether this varies with graph congruency is still unclear. Here we examined people’s interpretations of a more complex IPCC graph containing a conflict between spatial and conventional features (where higher values indicated lower values of climate change impact), and a simpler one without such a conflict. These two graphs will be referred to as ‘counter-intuitive’ and ‘intuitive’, respectively.

While objective comprehension of the information conveyed in the IPCC reports clearly represents a necessary condition for it to be implemented in actual policy, psychology research also shows that subjective comprehension—individual confidence in one’s comprehension—can be decision-relevant

(Jackson and Kleitman 2014). In fact, people typically have to rely on subjective comprehension to estimate their objective comprehension since subjective comprehension can be assessed internally, whereas objective comprehension can only be assessed externally (e.g. through objective tests, or feedback). Detrimental political consequences can arise from differences in interpretations of complex information between decision-makers (van den Broek 2018), but also from unwarranted confidence (Robinson and Marino 2015). For example, public servants whose subjective comprehension of climate change was higher than their objective comprehension tended to endorse more risky policy choices, above and beyond attitudes and political views (Liu *et al* 2017). Furthermore, among incumbent members of national parliament, subjective overconfidence in the likelihood of their re-election (but not objective likelihood) predicted risk taking (Sheffer and Loewen 2019). Subjective confidence has also been found to be particularly ill-calibrated for politicized science such as climate change, compared to non-politicized science such as biology and physics (Fischer *et al* 2019). On the other hand, higher confidence in climate change knowledge predicts higher climate change beliefs (estimated riskiness and anthropogenicity), above and beyond knowledge (Fischer and Said 2020). Similarly sufficiently high confidence in knowledge is required to put that knowledge into practice, also when controlling for differences in knowledge (Parker *et al* 2012). For example, healthcare workers with higher vaccination confidence (estimated benefits of vaccination) were more likely to recommend vaccination (Karlsson *et al* 2019).

Here, we evaluate objective and subjective IPCC graph comprehension among the main IPCC target audience, political decision-makers from climate-related governmental, and non-governmental organizations. We also recruited future decision-makers, junior diplomats from the German Federal Foreign Office as a comparison group of ‘the next generation’ of national representatives. Graphs on human health risks (figures 1 and 2) were selected from the Health chapter of the latest IPCC report (Smith *et al* 2014) because human health risks are among the most severe risks caused by climate change (Watts *et al* 2018). We used multiple-choice questions to assess objective graph comprehension, and asked participants to indicate their confidence in their answers to assess subjective comprehension. This allowed us to investigate how well subjective and objective understanding align: Do decision-makers realize whether they do, or do not understand IPCC graphs? Does such insight vary with the design of the graph? And: is subjective understanding correlated with objective understanding such that participants with higher subjective comprehension are, in fact, more accurate? The correlation between subjective



and objective comprehension is relevant because it allows one to estimate the reliability of informants, and because more confident informants appear more persuasive (Schwardmann and Van der Weele 2019; Smith *et al* 2017; Solda *et al* 2019). Hence policy-makers who overestimate their understanding may appear particularly convincing to others.

## 2. Method

### 2.1. Participants

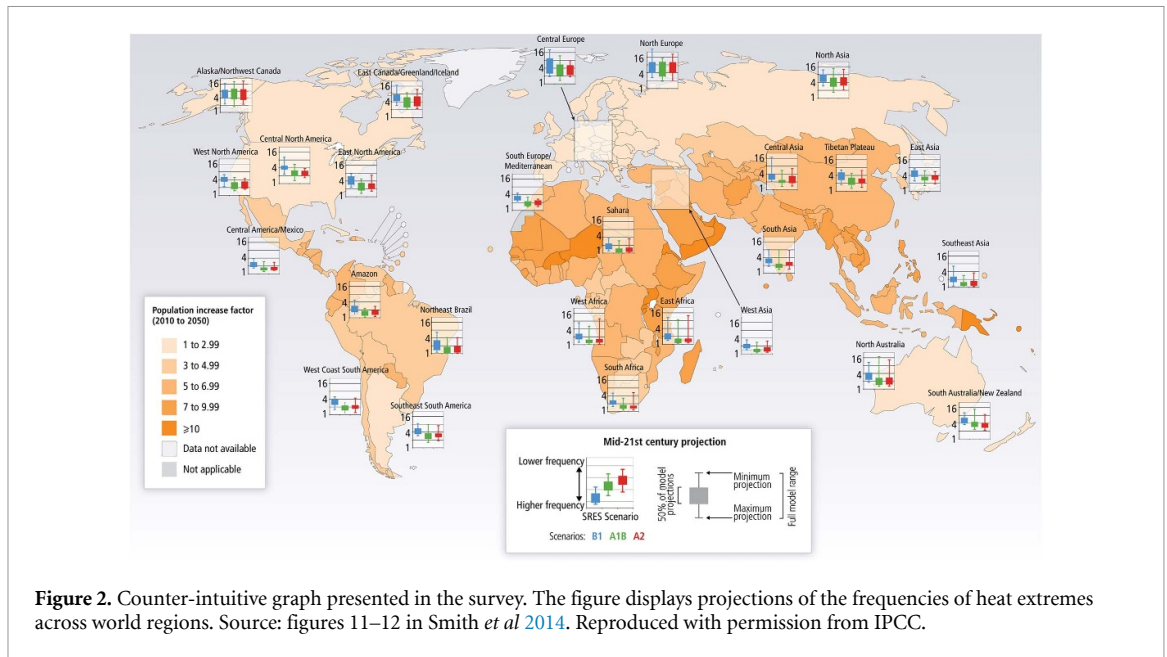
This study included two samples: IPCC target audience and German junior diplomats. The IPCC target audience included a sample of climate change decision-makers ( $N = 110$ ),  $n = 67$  expert political decision-makers from climate-related governmental organizations (consisting of 24 public servants working for Environmental Protection Agencies, EPAs; and 43 governmental policymakers working for ministries as well as governmental, and inter-governmental institutions,), and  $n = 43$  experts from climate-related non-governmental

organizations from various sectors such as environmental protection, urban planning, energy business or economic development. The sample represented a total of 54 countries (see appendix for figure A1 depicting all countries, see table 1 for other demographics).

The comparative sample of German junior diplomats consisted of future diplomats for the German Federal Foreign Office (*Auswärtiges Amt*) in Berlin, Germany ( $N = 33$ ). All junior diplomats had German citizenship.

### 2.2. Recruitment

For the IPCC target audience, invitations to take part in the survey were sent via email. The contact information was extracted from the UNFCCC website which lists the admitted parties and observer organizations of the United Nations Climate Change conference, the Conference of the Parties (COP). Contact details for EPA staff were extracted from the respective websites. In total, 1036 invitation emails were sent, of which 110



**Figure 2.** Counter-intuitive graph presented in the survey. The figure displays projections of the frequencies of heat extremes across world regions. Source: figures 11–12 in Smith et al 2014. Reproduced with permission from IPCC.

**Table 1.** Sample composition.

	IPCC target audience, n = 110	Junior Diplomats, n = 33
Age	Mean (SD): 46.9 (11.7)	Mean (SD): 31.5 (4)
Gender	Female: 44 (40.0 %)	Female: 12 (36.3%)
Education	High school: 5 (4.5 %) Bachelor: 12 (10.9%) Master: 55 (50.0%) PhD: 38 (34.5%)	High school: 0 Bachelor: 0 Master: 27 (81.8%) PhD: 6 (18.2%)
Organization	Climate-related governmental institutions: 43 (39.1%); Environmental Protection Agencies: 24 (21.8%) Climate-related non-governmental institutions: 43 (39.1%);	German Federal Foreign Office: 33 (100%)

participants completed the survey, yielding a completion rate of 11%. For the junior diplomat group, 47 junior diplomats of the German Federal Foreign Office were asked in person to participate in the study, and were sent the invitation link via email. Of those, 43 started the survey and 33 finished it, yielding a completion rate of 70%. Data collection was conducted in September 2017 (IPCC target audience except for EPA staff; junior Diplomats), and June to September 2018 (EPA staff).

### 3. Materials

#### 3.1. Prior and posterior beliefs

Participants indicated their belief in the anthropogenic nature of climate change (‘How much is climate change caused by humans?’). They also indicated their beliefs about specific aspects related to each graph, both before and after seeing the graphs. Specifically, they rated (1) the impact of climate change on human health (‘How much can climate change affect human health?’) and (2) the changes in the number of heat extremes until 2050 (‘Is the number of heat extremes going to increase globally

until the year 2050?’). Beliefs were rated using a slider (0 = *not at all*, 100 = *very much*). We recorded beliefs before and after inspection of the graphs to assess the degree to which IPCC visuals influence decisions-makers’ relevant beliefs.

#### 3.2. IPCC graphs

The graphs were selected from chapter 11, *Human health: impacts, adaptation and co-benefits*, from the recent IPCC report (Smith et al 2014). In the intuitive graph, higher values indicated higher levels of climate change impact; in the counter-intuitive graph, higher values indicated lower levels of climate change impact. Content-wise, the intuitive graph displayed the relative health impact from climate change and the potential for impact reduction for eight health-related sectors, and in light of three different temperature projections. The counter-intuitive graph displayed a world map showing the urban population increase factor until 2050 and bar graphs indicating mid-21st century projection of the frequency of extreme daily temperature for each region, with lower numbers indicating more frequent events.



### 3.3. Objective graph comprehension

We developed multiple-choice questions assessing comprehension of each graph. The comprehension questions were constructed such that (i) the main messages of the graphs were assessed; and (ii) prior knowledge on climate change or health were not required. Figure captions were displayed to ensure that displays were fully representative of the IPCC report.

The item corresponding to the intuitive graph assessed participants' comprehension of the potential risk reduction ('What does the graph tell you about the + 4 °C warmer world?'). Participants indicated for which facet the potential for risk reduction was greatest according to the graph (a) *undernutrition*, (b) *extreme weather events OR mental health and violence OR occupational health* (c) *vector-borne diseases* or (d) *heat OR food- and water-borne infections*). Option (a) was coded as correct (1) and all other responses were coded as incorrect (0). The item corresponding to the counter-intuitive graph assessed participants' comprehension of the extreme weather projections depicted ('Which of the following regions is expected to have the highest frequency in maximum daily temperatures that would have occurred only once in 20 years in the late 20th century according to scenario A2 (red bar)?'). The response options were (a) *North Asia*, (b) *Central America/Mexico*, (c) *North Europe*, (d) *South Asia*. Option (b) was coded as correct (1) and all other responses were coded as incorrect (0). We also inspected the distributions of incorrect responses, to examine whether participants relied on the mapping *higher = more/lower = less*. A reliance on this mapping would lead viewers to assume that higher values in the graph indicated more climate change impact, and incorrectly infer that (c) (North Europe) is the correct response.

For the sub-sample of EPA staff, we included an additional question with respect to the counter-intuitive graph to further assess whether participants made use of a *higher = more/lower = less*-mapping. Participants were asked to indicate the region that is expected to have the *lowest* frequency in heat extremes with the same response options as the previous item. We would expect that participants relying on the mapping would in this case infer that option (b) (Central America/Mexico) is the correct answer, due to the lowest values displayed for this region in the graph.

### 3.4. Subjective graph comprehension

After each objective comprehension question, participants were asked to judge the certainty with which they answered the comprehension question correctly ('How certain are you that your answer is correct?'), which participants rated on a four-point scale ('25%: just guessing' to '100%: completely certain'). This four-point scale was chosen such that the subjective comprehension scale aligns naturally with objective

comprehension (e.g. if participants are only guessing, they have a 25% chance of being correct), and because full-range scales (0%–100% certainty) were shown to yield worse calibrated confidence judgments (Weber and Brewer 2003).

### 3.5. Satisfaction with the graphs

Satisfaction with the graphs was assessed asking 'How satisfied are you with the graphs overall?' (0: *not at all*, 100: *completely*).<sup>1</sup>

## 4. Procedure

The survey was conducted in English for all participants. The order of the questions in the survey is given in table 2. Immediately below the display of each graph, participants were given the objective comprehension question, followed by the subjective comprehension question while the graph was still displayed. That is, participants had the option to look at the graphs while answering both questions.

## 5. Results

We report Bayesian analyses alongside frequentist analyses, where appropriate. We report  $BF_{10}$  throughout that express the probability of the data given  $H_1$  relative to  $H_0$ . For Bayes Factors, values  $> 1$  indicate increasing strength of evidence for  $H_1$ , values  $< 1$  indicate increasing evidence for  $H_0$ . As a guideline, values 1–3 (0.3–1) are seen as anecdotal, 3–10 (0.33–0.1) as substantial, and 10–30 (0.1–0.03) as strong evidence for the  $H_1$  ( $H_0$ ; Jeffreys 1998). All McNemar's  $X^2$  tests are continuity corrected.

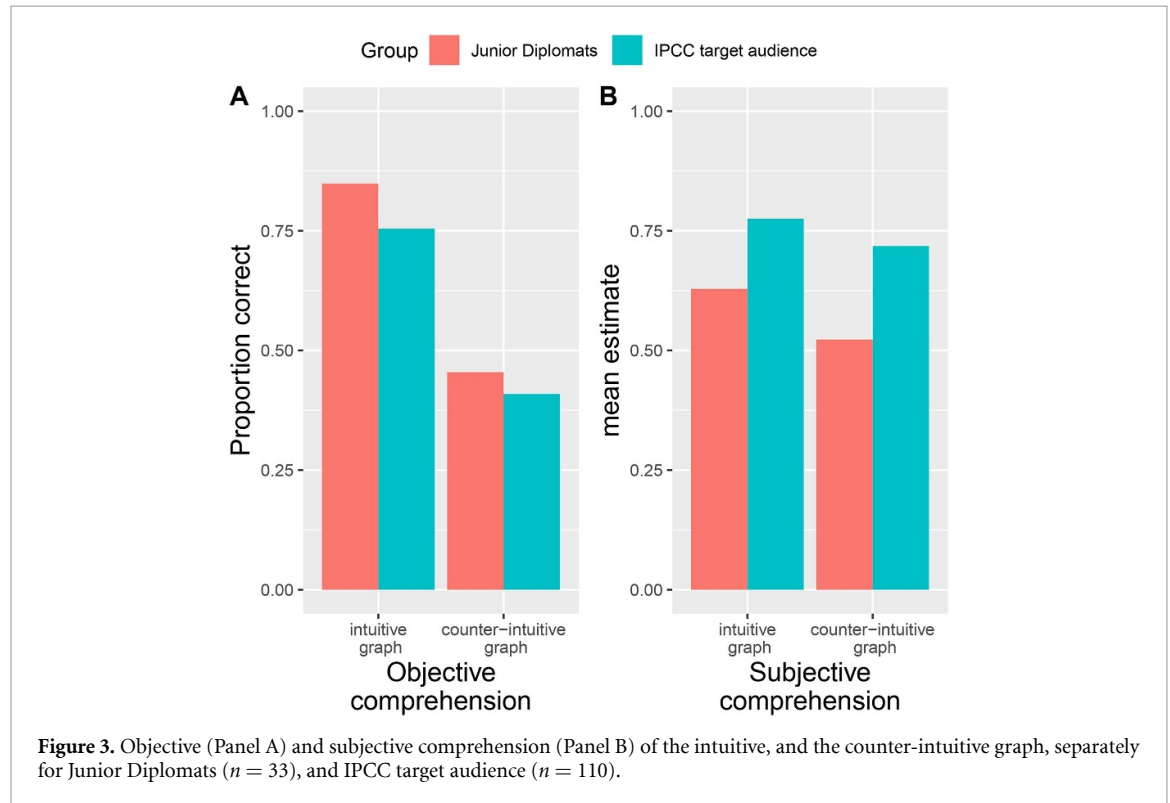
### 5.1. Objective graph comprehension

Of the IPCC target audience sample, 76% gave the correct answer to the intuitive graph. In contrast, only 41% of the sample gave the correct response to counter-intuitive graph, representing a drop in accuracy rate by  $> 30\%$ , McNemar's  $X^2(1) = 25.3$ ,  $p < .001$  (figure 3). Similarly, 85% of junior diplomats gave the correct answer to the intuitive graph, but only 46% gave the correct answer to the counter-intuitive graph, McNemar's  $X^2(1) = 11.1$ ,  $p < .001$ . Differences in proportions correct between both groups were negligible,  $X^2(1) = 0.81$ ,  $p = .37$ ,  $BF = 0.5$  and  $X^2(1) = 0.07$ ,  $p = .79$ ,  $BF = 0.38$ , for the intuitive and counter-intuitive graph, respectively. The most frequently selected response gives an indication as to *why* IPCC target audience provided incorrect responses: a total of  $n = 59$  (41%) of all participants answered that the region most affected by heat extremes was *North Europe* (as opposed to Central

<sup>1</sup>Numeracy and graph literacy were assessed exploratively, but using single items rather than full scales. Results are therefore not reported here.

Table 2. Procedure of the survey.

Section	Description
1	Welcome page
2	Perceived anthropogenic nature of climate change; prior beliefs
3	Objective comprehension intuitive graph: Health risks and potential for adaptation
4	Objective comprehension counter-intuitive graph: Frequency of heat extremes
5	Posterior beliefs
6	Satisfaction with the graphs
7	Demographics



America/Mexico), which had the highest values in the graph but is actually the region that is *least* affected.

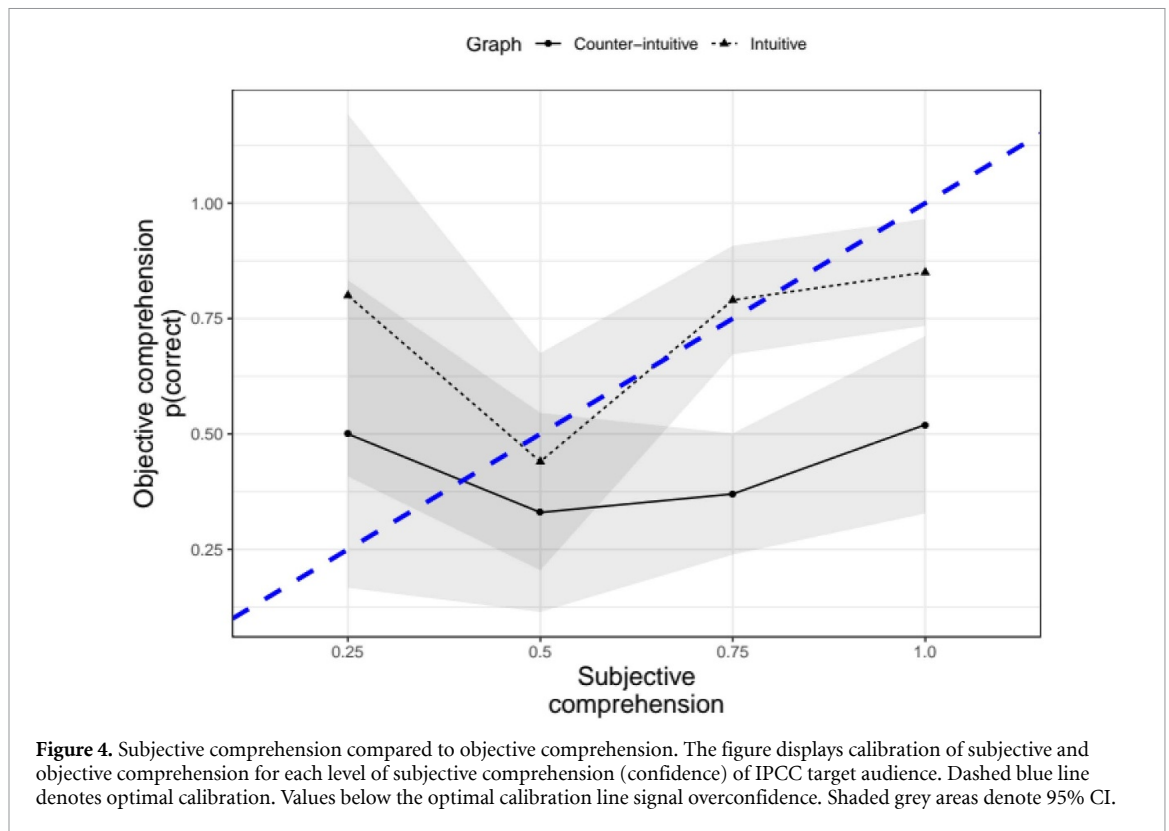
The second comprehension question for the counter-intuitive graph included for the sub-sample of EPA staff allowed us to investigate the potential use of a *higher = more-mapping*. Indeed results showed that accuracy correlated strongly between both questions for the counter-intuitive graph,  $\phi(22) = .81$ ,  $p < .001$ . Specifically, most IPCC target audience who answered ‘Northern Europe’ to question 1, also tended to answer ‘Central America/Mexico’ to question 2, suggesting that they did in fact employ a *higher = more-mapping* leading to these highly systematic errors.

## 5.2. Alignment of subjective and objective graph comprehension

Descriptive results of subjective comprehension per graph are given in figure 3. We used two indices to quantify the extent to which participants could judge whether they had correctly interpreted the graph, that is, the degree to which participants’ confidence is justified by their objective comprehension: (I)

*Correlation* between objective and subjective comprehension across participants which quantifies the extent to which more confident participants are more accurate; and (ii) *Calibration* which quantifies objective comprehension as a function of subjective comprehension, and is optimal when objective and subjective comprehension align (Weber and Brewer 2003). Calibration is hence relevant to assess the extent to which each level of subjective comprehension (e.g. pure guessing or 100% certainty) is predictive of objective comprehension.

(i) *Correlation*. Subjective comprehension was related to objective comprehension across IPCC target audience for the intuitive graph,  $r(108) = .22$ , 95% CI [.04, .39],  $p = .02$ , BF = 3.24. Importantly, however, subjective comprehension was unrelated to objective comprehension for the counter-intuitive graph,  $r(108) = .056$ , 95% CI [−.13, .24],  $p = .55$ , BF = 0.26. Among junior diplomats, correlation results provide anecdotal evidence for a positive relationship between subjective and objective comprehension for the intuitive graph,  $r(31) = .36$ , 95%CI [.02, .62],  $p = .04$ , BF = 2.3, and the counter-intuitive



graph,  $r(31) = .31$ , 95%CI  $[-.04, .59]$ ,  $p = .07$ ,  $BF = 1.5$ .

(ii) *Calibration.* Figure 4 shows objective comprehension for each level of subjective comprehension for the IPCC target audience. Calibration curves could not be estimated reliably for Junior Diplomats due to an insufficient number of participants in subjective\*objective comprehension cells. Results show that calibration curves generally followed optimal calibration (the diagonal) more closely for the intuitive compared to the counter-intuitive graph in that the confidence intervals of actual calibration of the intuitive graph entailed the optimal calibration at more levels of subjective comprehension compared to the counter-intuitive graph. This was especially so at the upper end of the confidence scale, that is, where IPCC target audience were almost or fully certain to have understood the graphs, which was indeed the case for the intuitive, but not the counter-intuitive graph. One interesting exception is the lower end of the confidence scale, that is, where participants believed to be guessing. Here, the subjective confidence overlaps with optimal calibration for the counter-intuitive graph only.

Figure 4 also shows that for the counter-intuitive graph, viewers were vastly overconfident of their understanding for all judgments made with medium to high confidence (75% or 100% certainty). This contrasts with results for the intuitive graph, where viewers were overconfident in their understanding of judgments made with 100% certainty only, and to a considerably lesser extent.

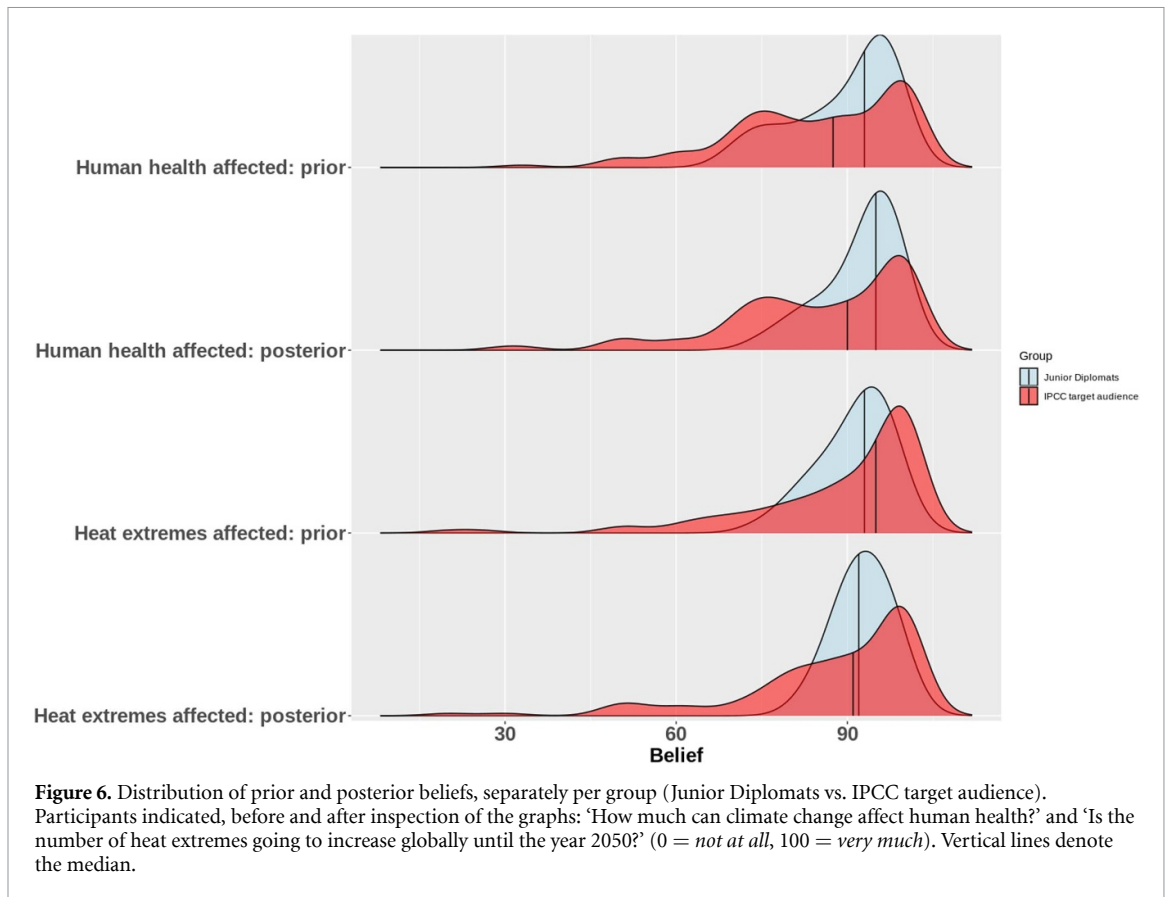
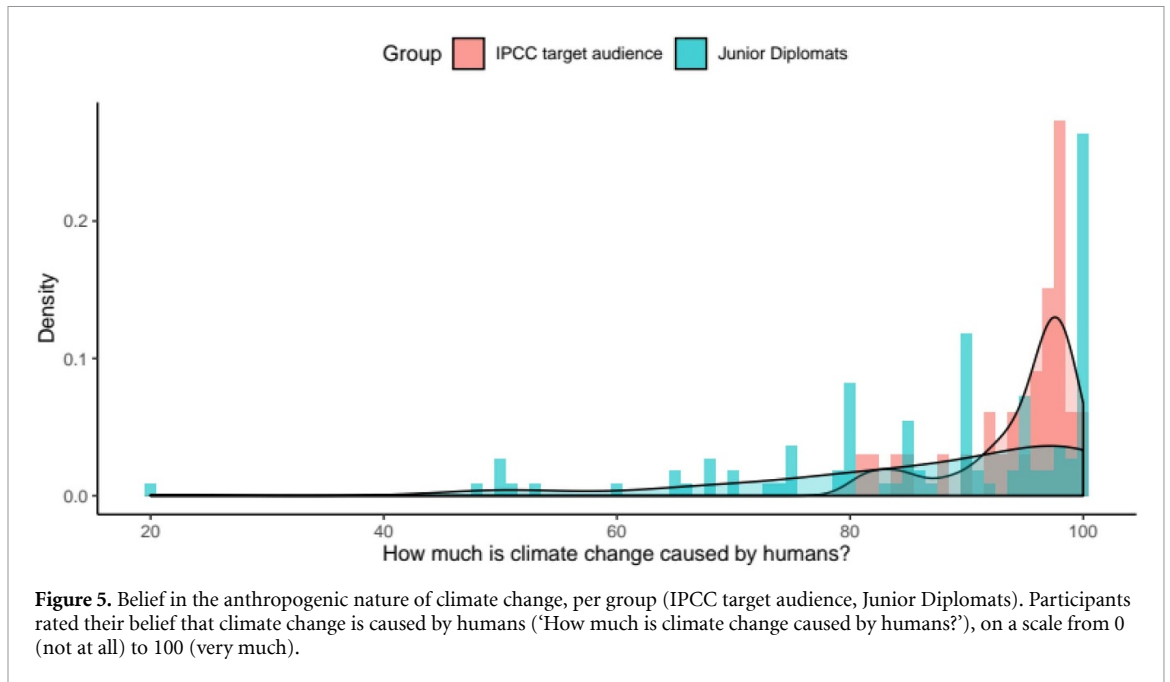
### 5.3. Belief change through inspection of IPCC graphs?

The distribution of the perceived anthropogenic nature of climate change was clearly left-skewed for both IPCC target audience, and junior diplomats. Beliefs about the anthropogenicity of climate change were, however, also varied, particularly for IPCC target audience (figure 5). IPCC target audience and Junior Diplomats did not appreciably update their prior beliefs regarding how much climate change can affect human health after inspection of the intuitive graph, nor their prior beliefs regarding the frequency of heat extremes after inspection of the counter-intuitive graph (figure 6). We additionally tested whether participants' prior beliefs were related to objective graph comprehension, as prior research has suggested that motivated interpretation can affect the accuracy of interpreting evidence as a function of the belief-congruency of the evidence (e.g. Kahan *et al* 2017). Results of these analyses are provided in the supplementary material (available online at <https://stacks.iop.org/ERL/15/114041/mmedia>).

### 5.4. Satisfaction with the graphs

IPCC target audience rated their overall satisfaction with the graphs with an average of  $M = 46.7$ , 95% CI  $[41.3, 52.1]$  on a 0 to 100 scale (Junior Diplomats:  $M = 31.3$ , 95% CI  $[25.8, 36.8]$ ). The distribution of satisfaction ratings (figure 7), however, shows a large variability that spanned the entire 0–100 scale. Interestingly, exploratory analyses showed



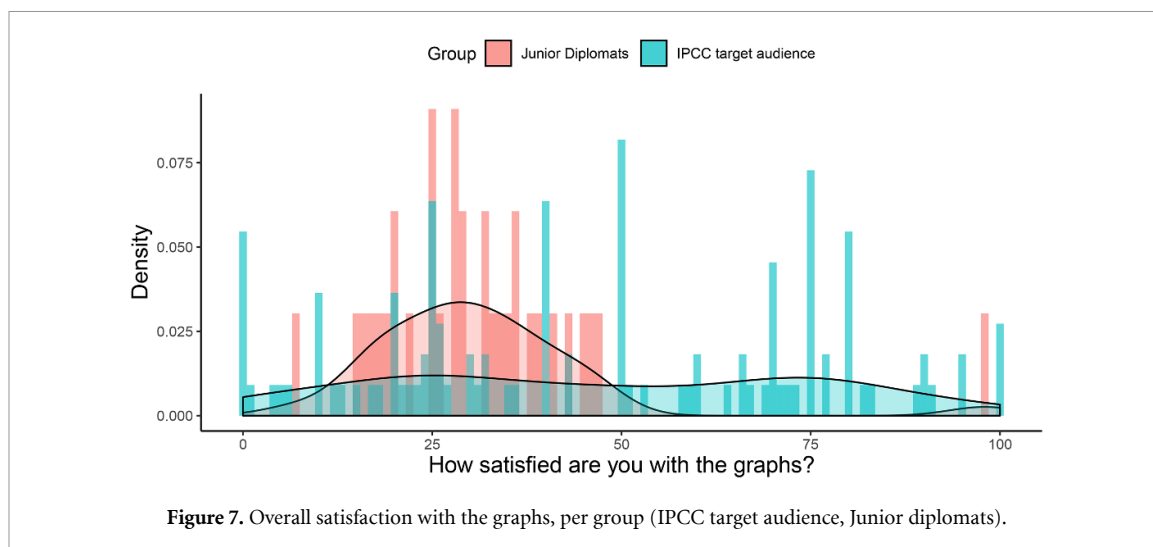


that a good predictor for graph satisfaction among IPCC target audience was subjective comprehension,  $r(108) = .24$ , 95% CI [.06, .41],  $p = .01$ , BF = 5.0.

**5.5. Discussion**

This study provides an empirical evaluation of the comprehension of IPCC graphs among IPCC target audience, decision-makers from climate-related

governmental organizations and climate-related non-governmental organizations, as well as a comparative sample of future diplomats. Results showed that while a majority of IPCC target audience correctly interpreted the graph that employed an intuitive design, less than half (47%) did so for the graph with the counter-intuitive design. As a worrying consequence of the counter-intuitive design, IPCC target



audience often drew the opposite conclusion from what was meant to be conveyed by the graph. This shows that the design of IPCC graphs can systematically bias interpretation among decision-makers.

Notably, the IPCC target audience seemed largely unaware of their misinterpretation of the counter-intuitive graph. Specifically, IPCC audience who answered with medium to high confidence were vastly overconfident of their objective understanding of the graph. In fact, even among those who were 100% certain that their interpretation was correct, only 50% interpreted the graph accurately (compared to approx. 80% for the intuitive graph). Importantly, a lack of awareness of the misinterpretations means that decision-makers may not be compelled to seek out information that will rectify their misunderstanding (Porto and Xiao 2016), and may even be prone to taking more risky policy choices (Simon and Houghton 2003, Ditttrich *et al* 2005, Robinson and Marino 2015). Moreover, political decision-makers who are overconfident in their understanding might—ironically—appear particularly convincing to others (Schwardmann and Van der Weele 2019, Solda *et al* 2019).

The intuitive graph, in contrast, yielded not only higher objective comprehension, but also subjective comprehension that was more indicative of objective comprehension. As revealed by the calibration plot, this was particularly the case at the upper end of the confidence scale, that is, where viewers were almost or completely certain to have understood the graph. This was indeed the case for the intuitive, but not the counter-intuitive graph. Counter-intuitive IPCC graphs hence pose a double hazard: First, by preventing policymakers from distilling the correct scientific information, which warrants political decision-making that is not based on the available evidence; and second by preventing decision-makers from realizing their lack of understanding. Both perils can be prevented by employing principles of intuitive design in the creation of informative graphs.

IPCC target audience who felt more certain of their interpretation were more likely to be satisfied with the graphs, suggesting that a feeling of understanding of the information was associated with being satisfied with its communication. This is in line with previous research on graphical risk communication that reported positive associations between perceived understanding of graphs and user evaluations (Okan *et al* 2020), and suggests that subjective comprehension may play a key role in the satisfaction of IPCC graphs among its target audience.

These results are particularly interesting in light of recent findings showing how IPCC authors themselves can reliably indicate which graphs viewers will find difficult to comprehend (Harold *et al* 2019). Although the two studies cannot be directly compared since they included different samples and graphs, the contrasting findings suggest that IPCC authors may be better at judging their audience's subjective graph comprehension than the audience themselves.

The pattern of results was broadly similar across Junior Diplomats and the IPCC target group, in that subjective understanding tended to be more aligned with objective understanding for the intuitive, compared to the counter-intuitive graph meaning that actual calibration of the intuitive graph entailed the optimal calibration at more levels of subjective comprehension compared to the counter-intuitive graph. These results are noteworthy given substantial demographic differences between groups: Junior Diplomats tended to be younger, more female, and have higher education levels. Tentatively, the similarity in results between both groups suggests that better-calibrated subjective understanding of intuitive IPCC graphs might be a general pattern that is generalizable to different groups of viewers.

Decision-makers' misinterpretation of IPCC graphs may have dire consequences for climate change policy. Specifically, the design of IPCC

graphs may lead to suboptimal use of the scientific information available. Fortunately, misinterpretations can be greatly reduced when risks are communicated using simple, well-designed graphical displays (Garcia-Retamero and Cokely 2017). The present findings therefore highlight a need for more intuitive design of IPCC graphs that follow evidence-based principles of effective graph design. Specifically, our findings emphasize the detrimental consequences of violating the general principle that spatial features (e.g. heights of bars) should convey the same meaning as conventional features such as legends or numerical labels on the scale (Okan *et al* 2012, 2016).

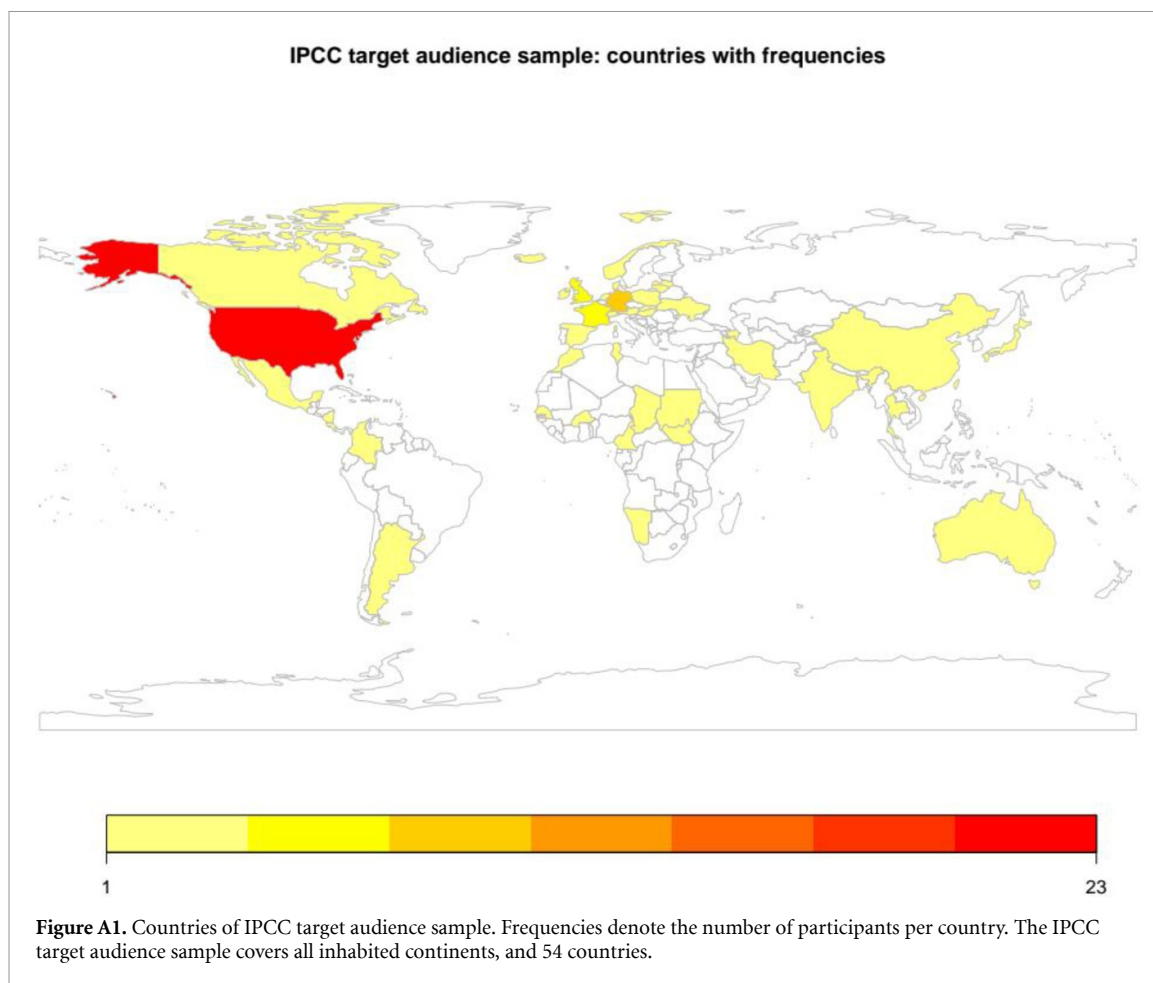
Taken together, our findings suggest that graph design should reflect congruency with regards to spatial-to-conceptual mappings, to ensure that accurate interpretations can be reached without extensive elaboration. Our findings also support the idea that high visual complexity should be avoided where possible. Although our study design does not allow for determining which specific aspects of complexity contributed to misinterpretation of the counter-intuitive graph (e.g. multiple variables, multiple data points), complex visualizations tend to be associated with slower and less accurate responses (Hegarty *et al* 2012, Padilla *et al* 2018). Reducing the visual complexity of graphs depicting climate data may be challenging as a certain level of detail may be needed to maintain scientific rigor and nuance. However, strategies have been outlined in the literature considering insights from the cognitive sciences, such as breaking down the data into different visual ‘chunks’ or considering whether information that is not essential for interpretation could be provided in text or a separate figure (Harold *et al* 2016). Other evidence-based strategies to promote understanding of visualizations include using sufficiently large sizes for relevant graph features and reducing the spatial distance between the visual pattern and the captions or legends (Harold *et al* 2016, Kause *et al* 2020). This can help to direct viewers’ attention to key information provided in textual elements and increase the likelihood that they will process and integrate such information. These strategies can be particularly helpful for individuals with lower graph literacy, who are less prone to attend to such information in counterintuitive graphs (Okan *et al* 2016). Additionally, an important general principle for effective science communication is to test the recipient’s understanding of the communication material (Bruine de Bruin and Bostrom 2013). This may help avoid unintended interpretations such as the ones documented here.

The findings of this study need to be carefully interpreted in light of several limitations. First, our sample of IPCC target audience was not a representative (or random) sample of the population of interest. Rather, the fairly low response rate among

current political decision-makers likely implies a self-selection bias. This self-selection probably caused an overestimation of IPCC graph comprehension since the current sample may be (1) more motivated to contribute to science (Jun *et al* 2017), (2) better at interpreting climate change information, or at least, (3) may have estimated the required time and effort to comprehend the graphs to be lower than policy makers that did not respond. Therefore, difficulties with interpretation, and outright misinterpretations may be more prevalent in the entire population of political decision-makers. Another limitation is that we did not assess participants’ graph literacy—the ability to understand graphically presented information, or numeracy—the ability to work with basic numerical concepts, using full scales. Numeracy might be relevant for comprehending IPCC graphs, because (i) numeracy may moderate any existing relationship between prior beliefs and accuracy of understanding (Kahan *et al* 2017), and because (ii) IPCC graphs tend to make extensive use of numerical information (Ame- lung *et al* 2016) that sometimes requires transformations. Furthermore, (iii) low graph literacy is associated with stronger reliance on spatial features such as heights of bars (Okan *et al* 2012, 2016), and future research could examine whether this is the case for IPCC graphs as well.

An important avenue for future research lies in experimental variation of the extent to which principles of intuitive design are employed in one and the same graph. Since original IPCC graphs were selected in the present study, relevant differences (other than the use or violation of principles of intuitive design) exist between both graphs. Specifically, although care was taken to select graphs from one chapter only, the graphs still differ in the type and amount of information they display, and how they display it. Particularly, the counter-intuitive graph uses non-linear axes, which has been shown to affect both objective and subjective comprehension (Fischer *et al* 2018), and lacks clear axis labels that indicate what exactly the displayed numbers represent. Building on existing insights from the cognitive science of graph comprehension (e.g. Kosslyn 2006), as well as our findings that suggest that EPA staff may have relied on spatial-to-conceptual mappings, future experimental research could help estimate in how far objective and subjective understanding of IPCC graphs can be improved through principles of intuitive design.

To conclude, climate change risk displays in IPCC reports can bias both objective and subjective understanding of the graphs, leading to systematically inaccurate conclusions, and ill-calibrated subjective comprehension. Specifically, the present results show that IPCC target audience tended to falsely believe they understood a graph the majority had misinterpreted.



Since inaccurate interpretation of key findings conveyed in IPCC graphs can have far-reaching implications for climate change policy, these findings signal a critical need for more intuitive design of IPCC graphs.

### Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: <https://osf.io/tzkbq/>. Data will be available from 24 June 2020.

### Appendix A1 Countries of origin of IPCC target audience

#### ORCID iDs

Helen Fischer

<https://orcid.org/0000-0001-5037-0390>

Karlijn L van den Broek

<https://orcid.org/0000-0002-0933-1194>

Yasmina Okan

<https://orcid.org/0000-0001-7963-1363>

#### References

Amelung D, Fischer H, Kruse L and Sauerborn R 2016 Defogging climate change communication: How cognitive research can

promote effective climate communication *Front. Psychol.* **7** 1340

Brewer N T, Gilkey M B, Lillie S E, Hesse B W and Sheridan S L 2012 Tables or bar graphs? Presenting test results in electronic medical records *Med. Decis. Mak.* **32** 545–53

Bruine de Bruin W B and Bostrom A 2013 Assessing what to address in science communication *Proc. Natl Acad. Sci.* **110** 14062–8

Dittrich D A, Güth W and Maciejovsky B 2005 Overconfidence in investment decisions: an experimental approach *Eur. J. Finance* **11** 471–91

Fischer H, Amelung D and Said N 2019 The accuracy of German citizens' confidence in their climate change knowledge *Nat. Clim. Change* 1–5

Fischer H and Said N 2020 Importance of domain-specific metacognition for explaining beliefs about politicized science: the case of climate change (<http://doi.org/10.31234/osf.io/fd6gy>)

Fischer H, Schütte S, Depoux A, Amelung D and Sauerborn R 2018 How Well Do COP22 Attendees Understand Graphs on Climate Change Health Impacts from the Fifth IPCC Assessment Report? *Int. J. Environ. Res. Public Health* **15** 875

Garcia-Retamero R and Cokely E T 2013 Communicating health risks with visual aids *Curr. Dir. Psychol. Sci.* **22** 392–9

Garcia-Retamero R and Cokely E T 2017 Designing visual aids that promote risk literacy: a systematic review of health research and evidence-based design heuristics *Hum. Factors* **59** 582–627

Harold J, Lorenzoni I, Shipley T F and Coventry K R 2016 Cognitive and psychological science insights to improve climate change data visualization *Nat. Clim. Change* **6** 1080–9

- Harold J, Lorenzoni I, Shipley T F and Coventry K R 2019 Communication of IPCC visuals: IPCC authors' views and assessments of visual complexity *Clim. Change* 1–16
- Hegarty M, Smallman H S and Stull A T 2012 Choosing and using geospatial displays: effects of design on performance and metacognition *J. Exp. Psychol. Appl.* **18** 1
- Jeffreys H 1998 *The Theory of Probability* (Oxford: Oxford University Press)
- Jun E, Hsieh G and Reinecke K 2017 Types of motivation affect study selection, attention, and dropouts in online experiments *Proc. ACM Hum. Comput. Interact.* **1** 56
- Kahan D M, Peters E, Dawson E C and Slovic P 2017 Motivated numeracy and enlightened self-government *Behav. Public Policy* **1** 54–86
- Karlsson L C, Lewandowsky S, Antfolk J, Salo P, Lindfelt M, Oksanen T, Kivimäki M and Soveri A 2019 The association between vaccination confidence, vaccination behavior, and willingness to recommend vaccines among Finnish healthcare workers *PLoS One* **14** e0224330
- Kause A, Bruine de Bruin W, Fung F, Taylor A and Lowe J 2020 Visualizations of projected rainfall change in the United Kingdom: an interview study about user perceptions *Sustainability* **12** 2955
- Kosslyn 2006 *Graph Design for the Eye and Mind* (Oxford: Oxford University Press)
- Lipkus I M 2007 Numeric, verbal, and visual formats of conveying health risks: suggested best practices and future recommendations *Med. Decis. Mak.* **27** 696–713
- Liu X, Stoutenborough J and Vedlitz A 2017 Bureaucratic expertise, overconfidence, and policy choice *Governance* **30** 705–25
- Mcmahon R, Stauffacher M and Knutti R 2015 The unseen uncertainties in climate change: reviewing comprehension of an IPCC scenario graph *Clim. Change* **133** 141–54
- Mcmahon R, Stauffacher M and Knutti R 2016 The scientific veneer of IPCC visuals *Clim. Change* **138** 369–81
- Okan Y, Galesic M and Garcia-Retamero R 2016 How people with low and high graph literacy process health graphs: evidence from eye-tracking *J. Behav. Decis. Mak.* **29** 271–94
- Okan Y, Garcia-Retamero R, Galesic M and Cokely E T 2012 When higher bars are not larger quantities: On individual differences in the use of spatial-to-conceptual mappings in graph comprehension *Spat. Cogn. Comput.* **12** 195–218
- Okan Y, Stone E, Parillo J, Bruine de Bruin W and Parker A 2020 Probability size matters: the effect of foreground-only vs. foreground+background graphs on risk aversion diminishes with larger probabilities *Risk Anal.* **40** 771–88
- Padilla L M, Creem-Regehr S H, Hegarty M and Stefanucci J K 2018 Decision making with visualizations: A cognitive framework across disciplines *Cogn. Res.: Princ. Implic.* **3** 29
- Parker A M, De Bruin W B, Yoong J and Willis R 2012 Inappropriate confidence and retirement planning: four studies with a national sample *J. Behav. Decis. Mak.* **25** 382–9
- Porto N and Xiao J J 2016 Financial literacy overconfidence and financial advice seeking *J. Financ. Serv. Prof.* **70**
- Robinson A T and Marino L D 2015 Overconfidence and risk perceptions: do they really matter for venture creation decisions? *Int. Entrepreneurship Manag. J.* **11** 149–68
- Schwardmann P and Van der Weele J 2019 Deception and self-deception *Nature human behaviour* **3** 1055–61
- Sheffer L and Loewen P 2019 Electoral confidence, overconfidence, and risky behavior: evidence from a study with elected politicians *Polit. Behav.* **41** 31–51
- Simon M and Houghton S M 2003 The relationship between overconfidence and the introduction of risky products: evidence from a field study *Acad. Manage. J.* **46** 139–49
- Smerecnik C M, Mesters I, Kessels L T, Ruiter R A, De Vries N K and De Vries H 2010 Understanding the positive effects of graphical risk information on comprehension: measuring attention directed to written, tabular, and graphical risk information *Risk Anal.* **30** 1387–98
- Smith K R, Woodward A, Campbell-Lendrum D, Chadee D D, Honda Y, Liu Q, Olwoch J M, Revich B and Sauerborn R 2014 Human health: impacts, adaptation, and co-benefits *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* ed C B Field et al (Cambridge: Cambridge University Press)
- Solda A, Ke C, Page L and Von Hippel W 2019 Strategically delusional *Exp. Econ.* 1–28
- Taylor A L, Dessai S and Bruine de Bruin W B 2015 Communicating uncertainty in seasonal and interannual climate forecasts in Europe *Philos. Trans. R. Soc. A* **373** 20140454
- van den Broek K L 2018 Illuminating divergence in perceptions in natural resource management: A case for the investigation of the heterogeneity in mental models *J. Dyn. Decis. Mak.* **4** 2
- Watts N, Amann M, Ayeb-Karlsson S, Belesova K, Bouley T, Boykoff M, Byass P, Cai W, Campbell-Lendrum D and Chambers J 2018 The Lancet Countdown on health and climate change: from 25 years of inaction to a global transformation for public health *Lancet* **391** 581–630
- Weber N and Brewer N 2003 The effect of judgment type and confidence scale on confidence-accuracy calibration in face recognition *J. Appl. Psychol.* **88** 490