

The unseen uncertainties in climate change: reviewing comprehension of an IPCC scenario graph

Rosemarie McMahon¹ · Michael Stauffacher² ·
Reto Knutti¹

Received: 31 March 2015 / Accepted: 10 July 2015 / Published online: 25 July 2015
© Springer Science+Business Media Dordrecht 2015

Abstract The reports published by the Intergovernmental Panel on Climate Change (IPCC) are comprehensive assessments of the scientific knowledge and uncertainties surrounding climate projections. They combine well-formed language with supporting graphical evidence and have the objective to inform policymakers. One of the most discussed and widely distributed visual in these reports is the graph, showing the global surface temperature evolution for the 21st century as simulated by climate models for various emission scenarios, which is part of the Summary for Policymakers (SPM) and the Working Group I contribution to the Fourth Assessment Report (AR4). It displays two types of uncertainties, namely the socio-economic scenarios and response uncertainty due to imperfect knowledge and models. Through 43 in-depth interviews this graph and caption was empirically tested with a sample of people analogous to the SPM target audience. It was found that novice readers were unable to identify the two different types of uncertainties in this graph without substantial guidance. Instead they saw a great deal of uncertainty but falsely attributed it to the climate model(s) and ignored the scenario uncertainties. Our findings demonstrate how the choice of display can directly impact a reader's perception of the scientific message. A failure to distinguish between these two types of uncertainties could lead to an overestimate of the response uncertainties, and an underestimation of socio-economic choices. We test this assumption and identify the difficulties non-technical audiences have with this graph and how this could inevitably impede its value as a decision support tool.

Electronic supplementary material The online version of this article (doi:10.1007/s10584-015-1473-4) contains supplementary material, which is available to authorized users.

✉ Rosemarie McMahon
rosemarie.mcmahon@env.ethz.ch
Michael Stauffacher
michael.stauffacher@env.ethz.ch
Reto Knutti
reto.knutti@env.ethz.ch

¹ Institute for Atmospheric and Climate Science, ETH Zurich, Zurich, Switzerland

² Institute for Environmental Decisions, ETH Zurich, Zurich, Switzerland

1 Introduction

The assessment by the Intergovernmental Panel on Climate Change (IPCC) is undoubtedly the most reliable and complete source of global climate knowledge with accuracy sustained through a highly exhaustive review and approval process. Providing governments with the relevant facts about climate change is highly complex and requires an “amalgamation of millions of gigabytes of data, as well as thousands of published findings” (IPCC 2013). Communicating how the climate will physically change is an essential component of the IPCC Summary for Policymakers (SPM) and the work that has been accomplished by Working Group I (WG I).

In general, scientists are not trained in communication and this is a recognised issue for climate communication (Pidgeon and Fischhoff 2011), and much of the guidance for scientists has focused on the communication of uncertainties, which is an inherent part of climate science (Moss 2011; Patt and Weber 2014; Rabinovich and Morton 2012). The IPCC has made a number of attempts to standardise the communication of uncertainties and the first formal effort to do that this was in the Third Assessment Report (TAR). A guidance note provided a common schema for representing and communicating uncertainty in both quantitative and qualitative terms (Jonassen and Pielke 2011). It offered a linguistic description for subjective confidence, which was later adapted by WG I and resulted in a descriptive scale for likelihood ranges. This was followed by a report from the U.S. Climate Change Science Program that helped to establish best practices for characterizing, communicating and incorporating uncertainty into climate change communication (CCSP 2009). Authors of the Fourth Assessment Report (AR4) received guidance on “how to report understanding, certainty and or confidence in the findings using common language” with emphasis on communication with policymakers (Mastrandrea et al. 2011). Today, climate scientists have successfully linked linguistic terms to quantitative statements about probability ranges in IPCC publications (Morgan and Mellon 2011).

Climate visualisation has unfortunately been largely neglected (Moser 2010) and moreover, Kandlikar et al. (2005), profess that many of the existing visuals inadequately communicate climate uncertainty. Yet visuals are commonly used to describe uncertainty in the IPCC report, but regrettably there exists limited literature on best practices for evaluating these graphics. It still remains unclear how visuals representing climate uncertainty are interpreted by non technical audiences (O’Neill and Smith 2014). Uncertainties are prevalent in SPM graphs such as, “Multi-Model Averages and Assessed Ranges for Surface Warming” (Fig. 1), which is part of the AR4 of WG I. This graph is routinely discussed in political and media circles and is an essential tool in the climate discourse (Schneider 2012). It represents uncertainties in the physical science but also the ambiguities in the socio-economic scenarios that drive these projections. In this paper this graph serves as a case study for the more general question “Can non-technical audiences distinguish between a climate response and scenario uncertainty?” We hypothesise that failure to distinguish between these two types of uncertainties could lead to a misjudgement about the overall magnitude of the uncertainties, thus resulting in an overestimate of the response uncertainties, and an underestimation of socio-economic choices. If this is the case, then this graph may not fulfil the “heuristic function of a decision making tool” (Schneider 2012, p. 189). We examine comprehension of this graph by capturing readers interpretation through in-depth interviews.

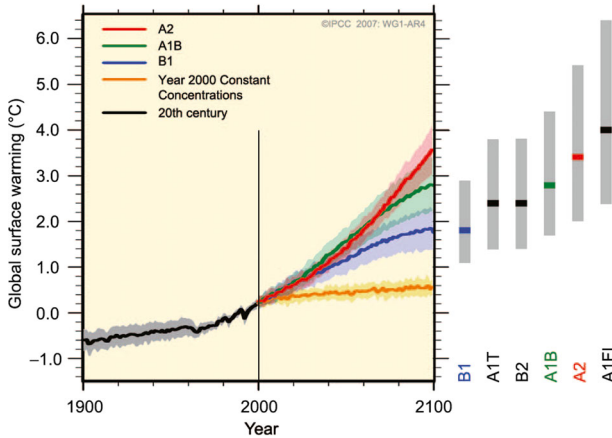


Fig. 1 Multi-model averages and assessed ranges for surface warming. *Solid lines* are multi-model global averages of surface warming (relative to 1980–1999) for the scenarios A2, A1B and B1, shown as continuations of the 20th century simulations. *Shading* denotes the ± 1 standard deviation range of individual model annual averages. The *orange line* is for the experiment where concentrations were held constant at year 2000 values. The *grey bars* at right indicate the best estimate (*solid line within each bar*) and the likely range assessed for the six SRES marker scenarios. The assessment of the best estimate and likely ranges in the grey bars includes the AOGCMs in the left part of the figure, as well as results from a hierarchy of independent models and observational constraints. Source (including caption): IPCC AR4 WG1 Figure SPM.5

1.1 Theory of visualisation

Many aspects of visualisation are well established, particularly how format, colour and aesthetics affect perception and cognition. The choice of visualisation however largely depends on the context and type of audience. Images that arouse an emotional response are known to influence first impressions and can affect usability and trust (Guzman and Schiller 2011). While visuals might be considered more accessible than other forms of communication they are however commonly misunderstood (Shah and Hoeffner 2002).

Graphs are known to improve the representation of quantitative information but often developed based on intuition, which is not the most reliable approach but common in many fields (Hegarty 2011). Moreover, graphs produced by experts often unintentionally exclude information that is critical to novice readers (Fischhoff 2013). Certain errors though are made during the interpretation of graphs, such as, viewing a graph as a picture instead of quantitative information (Glazer 2011). While graphs might seem more intuitive than other communication forms, the way in which people interpret a graph is based on a complex cognitive bottom up and top down process (Hegarty 2011). When a graph violates certain design principles, then readers will find it confusing and they will not be able to extract meaningful information (Kosslyn 1989). The bottom up process is an instinctive response to display features or, according to Hegarty (2011), the features that are salient in the display. The top down process involves the application of existing knowledge, which includes familiarity with graph schemas and prior knowledge of the topic.

Visualising uncertainty enables decision makers and others to deal with uncertainty information. As observed in the health risk sector, perception of uncertainty is heavily influenced by how it is displayed (Ancker et al. 2006; Garcia-Retamero et al. 2012; Hegarty et al. 2010). Much of the literature in visualisation is conducive to climate science and as Stephens et al.

(2012) demonstrate, the lessons learned from the communication of numerical weather predictions are applicable to climate communication.

1.2 Research questions

Our study examines the accuracy with which interpretation processes are performed and aims to distinguish the factors (graph salience, scientific literacy, content knowledge, and graph schema) that promote or hinder graph comprehension of this specific IPCC graph. It tackles a broad spectrum of questions such as; How is science visualising the different types of uncertainty?; Can non-technical audiences differentiate between the different types of uncertainties?; What are the main issues regarding the representation of climate uncertainties?; How do novice reader perceive uncertainty in the context of climate change?, and finally; How does a non-technical audience perceive climate science?

2 Methods

2.1 Sample selection

Recruitment was through a set of personalized emails (in total 129) sent to a convenient sample of interested scholars, governments, and non governmental intermediaries, who, according to Yohe and Oppenheimer (2011), represent the target audience of this graph. The response rate (33 %) was modest and presumably because of the broad interest in the topic of science communication, as paraphrased in the invite. In total 43 individuals participated in this study, with no incentives offered for participation. The sample was composed of four main groups; (1) decision makers (members of parliament, governmental administrators and affiliates of non-governmental organization) (2) communication experts (members of the media and graphic designers), (3) academics, and (4) climate scientists (expert). For the purpose of this study, all groups with the exception of climate scientists will be referred to as novice non-technical viewers but of course are experts in their respective fields, and some negotiators and government employees may well be experts in climate change as well. We acknowledge that the group selection may not fully represent the broader readership of the SPM and there exists other technical viewers, which have not been included in our sample but can be classed as experts.

2.2 Interview guide & materials

Discussions were guided by a set of predetermined semi-structured tasks established in accordance with the literature and covering all relevant independent (e.g., graph salience and prior knowledge) and dependent variables (e.g., comprehension of the displayed information). This conceptual framework captured all relevant interpretation levels known to influence graph perception and comprehension. The tasks were delivered in the following main conceptual blocks: (1) knowledge of probabilities and models, (2) interpretation and confidence in scientific information, (3) understanding of climate change and the IPCC, and lastly (3) recommendations and suggestions (see Online Resource 1). After the second block of questions were finalized, a coloured hardcopy of the graph was presented and only at this point did it become evident that the topic was about climate change. The graph was printed directly from

the IPCC website and included the main graph, header, title and caption but excluding the year of publication. No additional supporting text was provided during the interview.

Graph interpretation and comprehension was captured and categorised according to a conceptual framework (see Fig. 2a) that is based on existing literature (Canham and Hegarty 2010; Cleveland and McGill 1985; Hegarty 2011; Hegarty et al. 2010; Kosslyn 1989; Shah et al. 1999; Tufte 2001). This framework contains all variables that effect how a graph is deciphered and allows us to examine not only basic reading of a graph but also more importantly how the information displayed is conceptualized:

- The first level involves the bottom up process whereby syntactic perception is evaluated using a descriptive tasks (“What in your opinion is the main message of this graph?”) known to capture *graph salience* (Hegarty 2011).
- The second interpretation level is defined as *knowledge* and involves a top down cognitive process by which interpretation is affected by available information and in the context of this study this includes scientific literacy and content knowledge.
 - *Scientific literacy* relates to an understanding of concepts such as a model and probabilities and is not domain specific.
 - *Content knowledge* is a measure of climate science expertise. This was captured by a set of tasks that measured familiarity with the IPCC organization and its associated publications (“Are you familiar with the IPCC organization and if so where have you come across this organization?”; “What type of disciplines do you think currently work on the IPCC report?”).
- The third interpretation level is also top down process and evaluates *understanding of graph conventions and schemas* or simply how well the interviewed people can read a variable directly off a graph or read along an axis (“Looking at the green line what is the expected global mean temperature at the end of the century?”).
- These interpretation levels (graph salience, scientific literacy, content knowledge and graph schema) are known to affect *comprehension*, which in turn influences how information is conceptualized, which is the last stage. This is a vital stage as it helps an observer to form a pragmatic judgment of the information displayed, and these comprehension tasks measure how well a reader can add meaning to the information observed. Therefore testing whether or not a reader has fully comprehended the graph and conceptualized the information (“What do you perceive in this graph to be the pathway with the least impact for society?”, “Can you explain what this faded colour here represents and how would you use this information?”).

2.3 Interview procedures

Interviews were scheduled face to face and over a 5-month period up until the point of saturation.¹ Permission to record the interview was requested in advance of each discussion. All interviews were conducted in English and questions were not made available prior to the interview thus no preparation was required. When possible the interview was held at the

¹ “When the collection of new data does not shed any further light on the issue under investigation” (Mason 2010, p.1)

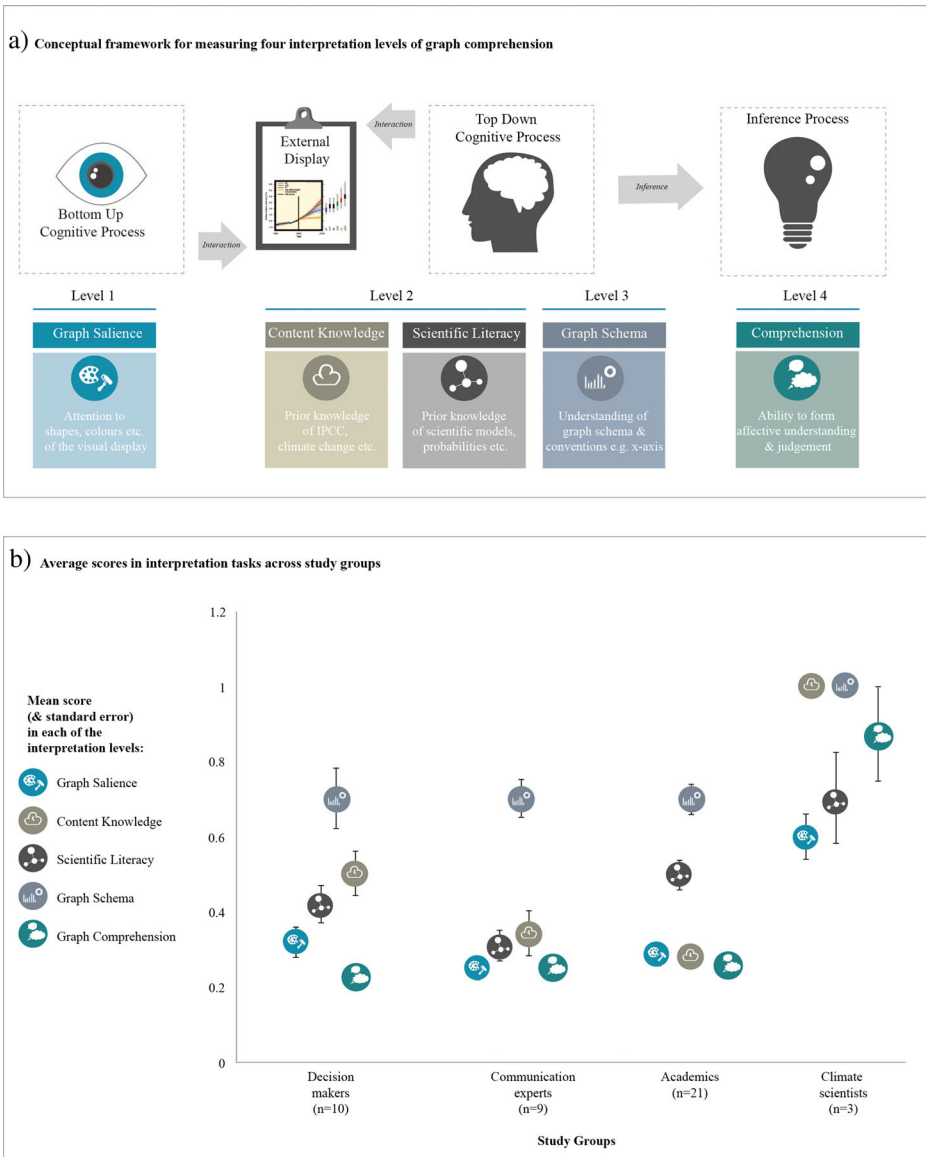


Fig. 2 **a** Conceptual framework outlining the bottom up and top down cognitive stages of graph comprehension that typically occurs when a reader interacts with an external visual display. This diagram is an adaption of a figure “Interaction with a Display” published by Hegarty 2011 (Fig. 4. pg. 453) **b** Quantitative scores for the response to tasks in graph salience, scientific literacy, content knowledge, graph schema and graph comprehension per group

working environment of the participant. Discussions lasted approximately 50 min. A full verbatim transcription was conducted to capture the full dialogue including expressions such as, “umms”, “am”, as well as pauses, changes in the tone of voice, stuttering, avoidance of a question or the repetition of a response. To ensure the transcript was free of error an

independent evaluator checked the quality. The full transcript was imported into a qualitative data analysis tool ² and each interview was coded according to the conceptual framework.

2.4 Content analysis of interview responses

Each interview was assessed qualitatively thus providing an appraisal for each readers overall comprehension. Relevant statements were highlighted as well as a full categorisation of the response to all tasks. This analysis was complemented with a subsequent quantitative assessment whereby responses were scored against a performance index, which measured Response Accuracy (RA) on an ordinal scale from one to five. A score of one represented an inadequate response, whereas five representing an accurate response. Each task had an associated anchoring text and RA score (see Online Resource 2). All participants were quantitatively measured on their ability to articulate a statement that matched this anchoring text and a score for RA was assigned. The index and assignment of scores was independently checked to ensure no biases were introduced.

2.5 Statistical analysis

Scores were aggregated and normalized across the independent and dependent variables. The independent variables were categorised in accordance with the results of a factor analysis and literature review. Differences in comprehension scores between groups were tested using a Kruskal Wallis test and the relationship between the variables was evaluated using a non-parametric Spearman's rank correlation coefficient test.

3 Results

The results are presented in two main sections; firstly the qualitative assessment of the rich descriptive responses, followed by a quantitative scoring of the responses. Details on the demographic profile of this sample are available in the Online Resource 3. Furthermore, additional interview responses are available in the Online Resource 4.

3.1 Qualitative assessment

3.1.1 Graph salience

This bottom up process is essential for graph comprehension as it determines how readers intuitively respond to a graph. Climate experts gave very elaborate descriptions of this graph, whereas novice readers offered weaker explanations. Expert descriptions went well beyond the visual cues available in the graph and they typically introduced information acquired through experience. Novice readers on the other hand, not having this background could only describe what they saw, which wasn't as informative, "*There is just one black line and then there are many coloured lines, and something grows and distributes I think*" [Doctoral in German Literature, 8].³ Novice readers saw one line that diverged at one point in time but the cause of this change was not clear, "*Developments in respect to*

² Qualitative data analysis tool, ATLAS for Mac OS X version, see product site <http://www.atlasti.com>

³ Interview statements are cataloged along with the participant's role and/or education level as well as an identification number; the latter can be cross-referenced with the Online Resource material.

some, I don't know, circumstances that will or will not happen" [NGO, *Degree in Geography*, 40]. All understood that temperatures were increasing with some margin of error, "*basically between 1900 and 2000 cooling down, I mean without knowing, what this is, it just kind of tells me that the warming is kind of at least increasing, although not entirely*" [Doctoral in Computer Science, 6]. From the descriptive accounts gathered, it was clear that this graph was not salient to novice readers and the intended message was distorted. These readers were distracted by less relevant information in the graph, such as the colour of the lines, which they believed was intentionally chosen "*red is obviously bad*" [Swiss MP, 42].

3.1.2 Scientific literacy

As anticipated, this highly educated sample was knowledgeable in scientific concepts such as models, probabilities, and uncertainty. Scientific literacy tended to be higher among academics particularly from natural science disciplines but in general, responses were less precise for more complex tasks. Attitudes to models very much depended on prior experience and some participants made incredulous comments, such as, "*I know you can make any model say anything about the same situation and the total opposite*" [Lobbyist, 5]. Others were sceptical about the reliability of a model; "*I mean I don't know how accurately we can predict the weather in all these parameters in the next 100 years*" [Doctoral in Neuroscience, 20]. Nevertheless, scientific literacy was high among all groups and the majority understood the conceptual difference between a scenario and a model although there were some imperfect responses, "*Well I would assume that the model takes into account the details of the scenario so they are not exactly the same thing but they are related. So different scenarios would give rise to different models or visa versa*" [Doctoral in Electrical Engineering, 33]. Curiously, many believed there was more uncertainty associated with a model than in a scenario but again there was a lot of indecision in this response.

3.1.3 Content knowledge

This sample had good general understanding of climate change but participants were not subject matter experts, which fulfilled our selection criteria. The vast majority were unfamiliar with the IPCC organization and this graph, again fulfilling our sample selection. Most of the participants were not climate experts and exhibited content knowledge that closely matched that of the SPM target audience. While responses were not as precise as the response of experts, it was still considered satisfactory. For those who had seen this graph before many understood both its origins and purpose. A few were even able to distinguish it from the earlier version in the third assessment report (AR3). Interestingly, the majority thought this graph was designed for a scientific community, "*I would expect to see this in a scientific article rather than for the general public*" [Doctoral in Astrophysics, 13]. While those who understood it was designed for policymakers thought it was more suitable for scientists and far too complex for politicians and the public, "*Scientists because nobody else understands it and nobody else has the patience*" [Swiss MP, 37]. The response to content tasks nicely distinguished our sample into those with high content knowledge ($n=3$) (climate experts) and those with no formal education in climate science ($n=40$).

3.1.4 Graph schema

In line with the high scientific literacy, this sample also exhibited substantial knowledge in graph conventions and were able to read direct measurements off this graph. The vertical axis

wasn't always easy to read as the units represented were not immediately evident, "*For the green line well I can read it, ah, its 2.5, oh (pause), so its almost 3 degrees*" [Graphic Designer, 21]; *So its slightly less than 3.0 so its 2.8 and 3.2 or something like that*" [Professor in Political Science, 22]. The overlapping colours also made the graph difficult to read, "*you always get an awful mix if you try to overlay four different colours*" [Game Designer, 30]. Overall, this sample performed well on graph schema tasks and understanding of graph conventions was high for all groups.

3.1.5 Graph comprehension

Graph comprehension tasks were a measure of how well a reader could differentiate between the two types of uncertainties represented in this scenario graph and conceptualise this information. With the exception of the three climate scientists, none of the readers were able to clearly distinguish between these two types of uncertainties. Even highly educated and analytical respondents were unable to precisely identify the portion of uncertainty associated with a scenario and what could be attributed to response uncertainty. Experts of course were quickly able to explain where the model uncertainty resided in this graph and the unknowns in social choices. The graph caption and legend offered no assistance to novice readers and somehow they deduced that each trend line was one single model, "*yes, I think every colour is a model*" [Governmental Administrator, 39]; "*I see here three models, the red, green and blue one*" [Professor in Agricultural Science, 14]. They became confused when asked to pick out the different scenarios in the graph and thought it might be the same thing as the model but knew intuitively this couldn't be correct, "*Well each line would be a model with its relevant background. Well yes a scenario (pause) it is not the same thing but what I referred to as a model might be a scenario*" [Doctoral in Geometry, 15].

Further comprehension tasks revealed that novice readers struggled to accurately explain what was meant by the transparent shading around the trend line, "*Am (pause), I don't know perhaps (pause) the maximum change or the mean, I don't know. It is hard to decide*" [Doctoral in Neuroscience, 20]. They understood it was "*some sort of error*" [Doctoral in Physics, 17] or "*an averaging*" [Scientific Journalist, 2], but were not clear what it actually represented. This was further accentuated by the fact that they had no idea how to interpret the grey bars on the right side of graph, "*I tried to figure out what these grey bars are but I couldn't so I just stopped*" [Doctoral in Computer Science, 6]. Likewise, the inconsistency between the size of the grey bars and the proportion of shading around the trend line confused novice readers, "*In here we have (pause), I guess (pause), they are not in the same height (pause), interesting (pause), and here it seems a kind of medium but I'm not really sure what they are telling me*" [Doctoral in Computer Science, 6]. All decision makers and communication experts believed that these grey bars should be removed altogether. They found it confusing and believed it carried no additional information. Contrary to this opinion, climate scientists thought the grey bars highly relevant for policy as it contained the full expert assessment. From the feedback received it was obvious that this novice audience was unable to appropriately judge the uncertainties displayed in this graph and were not able to conceptualise how socio-economic choices impact the global mean temperature over time ("*Hmm, Am (pause), no, it depends on what the goal is of this graph, so if the goal is to show what are the possible scenarios are with a certain model or different model for the global surface warming then I think this is reasonable adequate but if it is a goal of this graph to show how humans factors influence the global warming, then we need a different graph*" [Lobbyist, 5].

3.2 Quantitative assessment

Graph comprehension is characterised, as an ability to formulate a precise understanding of this graph and was determined from the response to nine graph comprehension tasks and the 16 tasks in graph salience, scientific literacy, and content knowledge. These responses were converted into quantitative scores, according to Section 2.4. An empirical analysis of the dependent variable confirm that graph comprehension was low ($M=0.29$, $SD=0.17$, $n=43$) across all groups with the exception of climate scientists who scored significantly higher (*Kruskal Wallis test*, $H(3)=10.99$ $p<0.05$) than all other readers. Aside of the climate scientists none of the other respondents were able to distinguish the two types of uncertainties represented in this graph, as shown in Fig. 2b.

A factor analysis and a reliability test of the independent variables support the grouping of tasks into the interpretation levels. An assessment of novice scores for all interpretation levels showed that graph salience was low ($M=.29$, $SD=.12$, $n=40$), whereas scores were slightly higher in scientific literacy ($M=0.45$, $SD=0.19$, $n=40$), content knowledge ($M=0.35$, $SD=0.19$, $n=40$), and graph schema ($M=0.76$, $SD=0.2$, $n=40$). Despite academics had a higher score in scientific literacy compared to decision makers and communication experts they still achieved a low score in graph comprehension. Likewise, decision makers with better content knowledge than academics and communication experts also performed poorly in graph comprehension. While the sample size limits what can be extrapolated from the statistical evidence, it would appear that graph salience might partly, but not fully explain why comprehension scores were low for this sample. Further details on the statistical evaluation are available in the Online Resource 5.

4 Discussion

Our findings suggest that the choice of graph conventions and schemas, while accepted in the field of climate science, are not conducive to novice readers as this audience was unable to identify the two types of uncertainties represented. This became evident in the appraisal of the interview responses and reinforced in the quantitative assessment. As outlined in the previous section, an inability to identify the scenario uncertainty in this graph meant that all the uncertainty shown was associated with the model(s) and scenario uncertainties were overlooked. This resulted in an overestimation of the uncertainties associated with the model(s) and a misguided perception that climate science is too uncertain to play any significant role in policy decisions.

During the interviews a number of anomalies were identified and coincide with the principle misunderstandings observed for this non-technical audience. The first irregularity was the widespread assumption that the lowest line on the graph represents the best outcome, as “*nobody wants to see global warming*” [*Graphic Designer, 21*]. Such cultural conventions were commonly held by all groups (with the exception of experts) and even among those who had prior exposure to this graph. Unfortunately, the lowest line is a counterfactual scenario or an idealized scenario only (Year 2000 constant concentration) and according to many experts, should never have been included in this graph with the other three economically plausible scenarios, whilst other experts believed that its inclusion makes the plot informative. Nevertheless, this heuristic association should be considered when designing future graphs, as concerned readers are likely to select the lowest line irrespective how egalitarian the choice of scenarios.

The second anomaly stems from a diverging perception of uncertainties between scientists and non-experts, which is a common theme in the literature (Patt and Weber 2014). When our sample discussed model uncertainties, participants considered the model to be inaccurate, “*full of error*” [Swiss MP, 1] whereas climate scientists believe it is an incomplete representation of reality. Likewise, opinions diverge in regard to scenario uncertainty; with climate scientists believing it was an uncontrollable unknown in the trajectory of climate change; whilst all others found it difficult to associate it with climate forecasts. From the perspective of a novice reader this graph is derived from a climate model and represented “*the technical aspects of climate change*” [Masters in Political Science, 18] and not the socio-economic choices. This psychological detachment is common in climate change and the corollary of this is novice readers tend to believe that all the uncertainties shown is related to the physical model errors.

The third anomaly was the common association that grey shading and transparency in the graph represents uncertainty, which aligns with the literature (Boukhelifa et al. 2012; Kinkeldey et al. 2014) and thus an appropriate representation technique. Nevertheless, these strong visual stimuli triggered our forth anomaly, which was the notion that “*there is a lot of scientific uncertainty*” [Government Administrator, 39] in this graph and the “*errors are huge*” [Doctoral in Physics, 19] which meant the “*data is meaningless right*” [Professor Mathematics, 7]. While the aim of the grey bars is to show the range of uncertainty across scenarios, it created the impression that the errors were vast and it would be “*impossible to model the climate with any level of accuracy*” [Swiss MP, 35]. This display of uncertainty caused viewers to see a lot of grey shading and relate it entirely to the physical model.

The last anomaly identified was the high confidence this audience had in the presented data and mainly because “*it looks scientific*” [Professor in Psychology, 34]. While scientific integrity is vital in the context of climate science it is no substitute for poor comprehension “*I have confidence in the data but no confidence in what I am reading*” [Doctoral in German Literature, 8]. This should be avoided and authors should not make figures overly complex just to seem scientific.

So what could have caused such poor understanding of this graph? On evaluation of the interpretation levels the manner in which uncertainty is displayed is not salient for non-technical audiences. Similar issues were identified by Stephens et al. (2012) who suggested that “*information needs to be presented at resolutions salient to a decision maker*”. Interpreting a graph is not straightforward and is a complex task even for an educated audience (Glazer 2011). This is because graphs follow a strict monosemic structure and those unfamiliar with such conventions struggle to comprehend the information displayed (Bertin 2010). Climate scientists applied graphical structures established in their field but such structures are not lucid to those outside this domain. An incomplete knowledge of the graphical structures used by climate experts prevented these novice readers from extracting the relevant information. Likewise, the abbreviations used in the legend further undermined the salience of this graph and decision makers persistently remarket this was confusing. Novice readers tended to be unaware of their own impaired judgement of the graph and much of their explanations were based on prior assumptions about climate change.

While our findings are insightful they can only be extrapolated to this study group and no assumptions can be made for policymakers in general. Additionally, open-ended questions are known to potentially introduce biases however this was counteracted through a conceptual framework and quality checks. Future studies should consider testing variations of this graph in a controlled experimental design in order to identify the design elements that hinder or promote better judgement.

5 Conclusion

Despite the efforts of scientists to display the two types of climate uncertainties in this graph, they failed to adequately communicate this intended scientific message to this study audience. While these findings elicit a sober prospect for this scenario graph, it does raise an important issue, the general lack of research on visual narratives in the SPM report. The inadvertent misrepresentation of uncertainty in this graph created the impression that the climate models were highly uncertain, thus making this the single most memorable aspect of this graph. Had the authors explicitly explained the “source, nature, and magnitude of the uncertainty” maybe such misunderstandings could have been avoided (Budescu et al. 2009) and the uncertainty correctly attributed. Instead, readers substituted the information in the figure with their own prior assumptions or read it in line with their prior assumptions (confirmation bias). Acting on these findings means that we can no longer rely on intuition to develop climate visuals but instead need to accept that other expertise are required and this could involve a co-authorship with experts in science communication, as has been done for the Frequently Asked Questions in the report. It is thus vital to acquire a clear understanding of the information requirements of decision makers (von Winterfeldt 2013) and to prioritise research in climate uncertainty communication (Fischhoff 2011), in particular the use of visual mediums. Future studies should also consider how best to balance the use of text and figures in the SPM and how this impacts usability. Furthermore, models and scenarios are used differently in different fields, so it is possible that prior assumptions differ, which would need to be examined in future IPCC reports.

A number of crucial lessons can be draw from our study, (1) stakeholders perspectives need to be considered in the development of a graph and it should be tested on a subset of this audience prior to publication, (2) design based on intuition needs to be replaced with evidence based design, (3) graph comprehension should be tested and retested, (4) graphs must convey the intended message, and lastly (5) a standalone graph should be immediately salient to the intended audience. The political framework surrounding the establishment of SPM visuals should not be ignored and is an important consideration. Radical changes to SPM visuals could result in unwanted political feedback but this shouldn't be a reason to suspend much needed research in climate visual communication. As noted in this paper, there exist considerable gaps in our understanding of how the visualisation of climate science aids decision-making. Improving our understanding based on empirical evidence will encourage graph authors to seek more effective approaches and thus exploits visual features that promote understanding of climate science and permit informed decision-making.

In conclusion we are unable to clearly tell whether the lack of graph comprehension is a more general issue, which would imply that such summaries would have to be done differently for novice readers to understand, or whether it is the different character of model uncertainty vs. societal choices that is hard to understand, in which case the presentation of this concepts in this particular figure would have to be rethought. We speculate it is the former, and plan to test alternative, dynamic and potentially interactive figures to further evaluate how best to communicate with novice readers.

Acknowledgments The Institute for Atmospheric and Climate Science and the Institute for Environmental Decisions at the ETH Zurich, Switzerland jointly supported the author of this paper.

References

- Ancker JS, Senathirajah Y, Kukafka R, Starren JB (2006) Design features of graphs in health risk communication: a systematic review. *J Am Med Inform Assoc* 13:608–618. doi:10.1197/Jamia.M2115
- Bertin J (2010) [Semiologie graphique. English] *Semiology of graphics: diagrams, networks, maps* / Jacques Bertin; translated by William J. Berg. 1 edn. Esri Press
- Boukhelifa N, Bezerianos A, Isenberg T, Fekete JD (2012) Evaluating sketchiness as a visual variable for the depiction of qualitative uncertainty. *IEEE Trans Vis Comp Graph* 18:2769–2778
- Budescu DV, Broomell S, Por HH (2009) Improving communication of uncertainty in the reports of the intergovernmental panel on climate change. *Psychol Sci* 20:299–308. doi:10.1111/J.1467-9280.2009.02284.X
- Canham M, Hegarty M (2010) Effects of knowledge and display design on comprehension of complex graphics. *Learn Instr* 20:155–166. doi:10.1016/J.Learninstruc.2009.02.014
- CCSP (2009) Best practice approaches for characterizing, communicating, and incorporating scientific uncertainty in decisionmaking. [M. Granger Morgan (Lead Author), Hadi Dowlatabadi, Max Henrion, David Keith, Robert Lempert, Sandra McBride, Mitchell Small, and Thomas Wilbanks (Contributing Authors)]
- Cleveland WS, McGill R (1985) Graphical perception and graphical methods for analyzing scientific-data. *Science* 229:828–833. doi:10.1126/Science.229.4716.828
- Fischhoff B (2011) Applying the science of communication to the communication of science. *Clim Chang* 108:701–705. doi:10.1007/S10584-011-0183-9
- Fischhoff B (2013) The sciences of science communication. *Proc Natl Acad Sci U S A* 110:14033–14039. doi:10.1073/Pnas.1213273110
- Garcia-Retamero R, Okan Y, Cokely ET (2012) Using visual aids to improve communication of risks about health: a review *Sci World J* doi: Artn 562637
- Glazer N (2011) Challenges with graph interpretation: a review of the literature. *Stud Sci Educ* 47:183–210. doi:10.1080/03057267.2011.605307
- Guzman E, Schiller J (2011) How does this look? In: C Stephanidis (Ed) *Desirability methods for evaluating visual design. Posters, Part I, HCI 2011, CCIS 173, Springer-Verlag Berlin Heidelberg*, pp 123–127
- Hegarty M (2011) The cognitive science of visual-spatial displays: implications for design. *Top Cogn Sci* 3:446–474. doi:10.1111/J.1756-8765.2011.01150.X
- Hegarty M, Canham MS, Fabrikant SI (2010) Thinking about the weather: how display salience and knowledge affect performance in a graphic inference task. *J Exp Psychol Learn Mem Cogn* 36:37–53. doi:10.1037/A0017683
- IPCC (2013) Working group 1 fact sheet. Intergovernmental Panel on Climate Change (IPCC), Geneva, Switzerland. http://www.climatechange2013.org/images/uploads/WG1AR5_FactSheet.pdf. Accessed 10 January 2015
- Jonassen R, Pielke R (2011) Improving conveyance of uncertainties in the findings of the IPCC. *Clim Chang* 108:745–753. doi:10.1007/S10584-011-0185-7
- Kandlikar M, Risbey J, Dessai S (2005) Representing and communicating deep uncertainty in climate-change assessments. *CR Geosci* 337:443–455. doi:10.1016/J.Crte.2004.10.010
- Kinkeldey C, MacEachren AM, Schiewe J (2014) How to assess visual communication of uncertainty? A systematic review of geospatial uncertainty visualisation user studies. *Cartogr J* 51:372–386. doi:10.1179/1743277414y.0000000099
- Kosslyn SM (1989) Understanding charts and graphs. *Appl Cogn Psychol* 3:185–226. doi:10.1002/Acp.2350030302
- Mason M (2010) Sample size and saturation in PhD studies using qualitative interviews [63 paragraphs] *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research* 11(3), Art. 8
- Mastrandrea MD et al (2011) The IPCC AR5 guidance note on consistent treatment of uncertainties: a common approach across the working groups. *Clim Chang* 108:675–691. doi:10.1007/S10584-011-0178-6
- Morgan MG, Mellon C (2011) Certainty, uncertainty, and climate change. *Clim Chang* 108:707–721. doi:10.1007/S10584-011-0184-8
- Moser SC (2010) Communicating climate change: history, challenges, process and future directions. *WIREs Clim Chang* 1:31–53. doi:10.1002/Wcc.011
- Moss RH (2011) Reducing doubt about uncertainty: guidance for IPCC's third assessment. *Clim Chang* 108:641–658. doi:10.1007/S10584-011-0182-X
- O'Neill SJ, Smith N (2014) Climate change and visual imagery. *WIREs Clim Chang* 5:73–87. doi:10.1002/Wcc.249
- Patt AG, Weber EU (2014) Perceptions and communication strategies for the many uncertainties relevant for climate policy. *WIREs Clim Chang* 5:219–232. doi:10.1002/Wcc.259

- Pidgeon N, Fischhoff B (2011) The role of social and decision sciences in communicating uncertain climate risks. *Nat Clim Chang* 1:35–41. doi:[10.1038/Nclimate1080](https://doi.org/10.1038/Nclimate1080)
- Rabinovich A, Morton TA (2012) Unquestioned answers or unanswered questions: beliefs about science guide responses to uncertainty in climate change risk communication. *Risk Anal* 32:992–1002. doi:[10.1111/J.1539-6924.2012.01771.X](https://doi.org/10.1111/J.1539-6924.2012.01771.X)
- Schneider B (2012) Climate model simulation visualization from a visual studies perspective. *WIREs Clim Chang* 3:185–193. doi:[10.1002/Wcc.162](https://doi.org/10.1002/Wcc.162)
- Shah P, Hoeffner J (2002) Review of graph comprehension research: implications for instruction. *Educ Psychol Rev* 14:47–69. doi:[10.1023/A:1013180410169](https://doi.org/10.1023/A:1013180410169)
- Shah P, Mayer RE, Hegarty M (1999) Graphs as aids to knowledge construction: signaling techniques for guiding the process of graph comprehension. *J Educ Psychol* 91:690–702. doi:[10.1037//0022-0663.91.4.690](https://doi.org/10.1037//0022-0663.91.4.690)
- Stephens EM, Edwards TL, Demeritt D (2012) Communicating probabilistic information from climate model ensembles—lessons from numerical weather prediction. *WIREs Clim Chang* 3:409–426. doi:[10.1002/Wcc.187](https://doi.org/10.1002/Wcc.187)
- Tufte ER (2001) *The visual display of quantitative information*. 2 edn. Graphics Press LLC
- von Winterfeldt D (2013) Bridging the gap between science and decision making. *Proc Natl Acad Sci U S A* 110:14055–14061. doi:[10.1073/Pnas.1213532110](https://doi.org/10.1073/Pnas.1213532110)
- Yohe G, Oppenheimer M (2011) Evaluation, characterization, and communication of uncertainty by the intergovernmental panel on climate change—an introductory essay. *Clim Chang* 108:629–639. doi:[10.1007/S10584-011-0176-8](https://doi.org/10.1007/S10584-011-0176-8)