# The potential power of experience in communications of expert consensus levels

Adam J. L. Harris[a]*, Oliver Sildmäe [a], Maarten Speekenbrink[a], and Ulrike Hahn[b]

[a]*Dept. of Experimental Psychology, University College London, London, UK;* [b]*Dept. of Psychological Sciences, Birkbeck College, University of London, UK.*

*Adam J. L. Harris, Dept. of Experimental Psychology, University College London, 26 Bedford Way, London, UK, WC1H 0AP. Email: adam.harris@ucl.ac.uk

Oliver Sildmäe, Dept. of Experimental Psychology, University College London, 26 Bedford Way, London, UK, WC1H 0AP. Email: oliver.sildmae@gmail.com

Ulrike Hahn, Dept. of Psychological Sciences, Birkbeck College, University of London, UK. Email: u.hahn@bbk.ac.uk

Word count: 8,022 (excluding abstract and figure captions)

# The potential power of experience in communications of expert consensus levels

Understanding the scientific consensus on anthropogenic climate change has been dubbed a 'gateway belief' to engaging people in sustainable behaviour. We consider the question of how the impact of a consensus communication can be maximised. Firstly, the credibility of the communicator should be maximised. One way of achieving this is to present the opinions of a sample of scientists directly to individuals. The decision making literature suggests that such a technique will confer an *additional* advantage over standard descriptions of consensus (e.g., "97% of scientists agree"). In decision making research, low probabilities tend to be overweighted when probabilities are described, but underweighted when probability information is experienced. Consequently, statements of high consensus may lead to an overweighting of the dissensus, a phenomenon that may be reversed were the consensus to be 'experienced.' We obtain some positive support for our proposal that consensus is best 'experienced' in one of two experiments. We suggest that the lack of stronger positive support could relate to ceiling effects for the topics studied and propose that investigation of effective methods for 'experiencing' the consensus is a fruitful area for future research.

Keywords: source credibility; decisions from experience; climate change; testimony

> ‘97% of climate scientists agree that humans are causing global warming’
> ‘Nine out of ten economists say Brexit will damage economy’
> ‘84% of medical scientists agree that regular aerobic exercise lowers diastolic blood pressure’ (Cook et al., 2013; Green, 2016; Whelton, Chin, Xin, & He, 2002).

The statements above represent a fundamental element of the scientific method, that of consensus building. As an incremental, cumulative practice, an emergent consensus becomes our best possible estimate of the true state of the world. Subsequently, decisions, policies, engineering breakthroughs,…, can all be made relative to this scientific knowledge. How, though, should this consensus best be communicated to the public and policy makers to ensure maximum effectiveness?

Expert consensus is relevant across almost any domain which requires specialist knowledge. Some recent research has, however, specifically highlighted the role of scientific consensus in relation to the topic of anthropogenic climate change (ACC). As a universal global concern, such a focus is understandable. Ding, Maibach, Zhao, Roser-Renouf and Leiserowitz (2011; see also, McCright, Dunlap, & Xiao, 2013; van der Linden, Leiserowitz, Feinberg, & Maibach, 2015) demonstrated, statistically, that correct perceptions of high levels of expert consensus about the reality of ACC predicted greater support for climate change mitigation policies through increasing beliefs that ACC is occurring. Subsequently, Lewandowsky, Gignac and Vaughan (2013; see also van der Linden, Leiserowitz, Feinberg, & Maibach, 2014) demonstrated this latter result experimentally. Specifically, participants informed that “97 out of 100 climate experts agree that global warming is a consequence of the burning of fossil fuels” (Lewandowsky, Gignac, & Vaughan, 2013, p. 403), subsequently reported higher estimates of the scientific consensus and a higher belief in the link between human CO2 emissions and climate change. Thus, highlighting the level of scientific consensus has been suggested as an approach to increase support for climate mitigation policies, being tagged a 'gateway belief' for the subsequent support of such actions (e.g., van der Linden et al., 2015).

The efficacy of consensus communications in affecting greater support for sustainable policies has not, however, gone unquestioned (see e.g., Cook, 2016; Cook & Jacobs, 2014, for reviews). As an example, Kahan (2015) argued that from 2003 to 2013, public option had not updated in favour of believing in ACC, despite the release of six studies quantifying high levels of scientific consensus in that time period, plus the release of 'An Inconvenient Truth.' Consequently, Kahan argues, lab-based demonstrations of the efficacy of consensus messaging must not be ecologically valid. This debate is not of central concern to the current manuscript, though we note that since 2010 there is evidence that public belief in ACC has increased (e.g., Gallup, 2017; Hamilton, 2016; the start of this pattern is also noticeable in Kahan's, 2015, Figure 8). What *is* of central concern to the current manuscript is that, in addition to demonstrating the effectiveness of messages about the scientific consensus, laboratory studies have also demonstrated that there is still room for improvement. Specifically, whilst estimated levels of scientific consensus increased an impressive 21%, from 67% to 88%, following receipt of the consensus message in Lewandowsky, Gignac, and Vaughan (2013), this 'posterior' degree of belief still remained some way short of the communicated figure of 97% (see Bolsen, Leeper, & Shapiro, 2014, for even more mixed results).

The current paper draws insights from formal work on source credibility, and descriptive research in human decision making to suggest methods for improving the effectiveness of the communication of expert consensus. Finally, we present two experiments as an initial test of the potential efficacy of these methods.

***Source credibility***

The statements presented at the top of this article represent pieces of testimony. That is, *someone* is stating that this proportion of experts agree on this fact. Arguably, most of what we believe we know stems from the testimony of others. As an example, most people's belief that North America is west of Europe is based on a map that another individual has drawn. Even our experience of travelling by plane provides little direct evidence as to the directional relationship between the continents. The current authors' (firm) belief that America does lie west of Europe is based primarily on a level of trust in map makers (and geography teachers).

As the example above suggests, source credibility is a key factor underlying how persuasive a piece of testimony should be (e.g., Bovens & Hartmann, 2003; Friedman, 1987; Schum, 1981). Moreover, Hahn, Harris and Corner (2009; see also Harris & Hahn, 2009; Harris, Hahn, Madsen, & Hsu, 2016) have demonstrated that people can incorporate considerations of source credibility into their evaluation of a communication in a normatively appropriate manner. It is thus important to give source credibility due attention in consensus communications.

One way in which a source might be unreliable is if they are deliberately deceitful. Perceptions of deceitfulness are at the heart of conspiracy claims such as 'the moon landings were faked to demonstrate the United States' superiority over the USSR'. The potential for deceit is an issue of genuine concern in an evaluation of testimony. However, considerations of source credibility are not only integral in instances of wilful deceit. A lack of competence should also influence the persuasiveness of a piece of testimony. As an extreme example, if your (honest – i.e., there is no misdirection involved) friend flips a coin under a cup and says "I think it's heads", you should not update your belief that the coin under the cup has landed on heads. Your friend in this instance has no privileged information – they are as likely to say "heads" if the coin has come up heads as if it has come up tails. Subsequently, their 'testimony' is non-diagnostic of the true state of the world. Less extreme examples are ubiquitous. As one example, multiple factors are known to impact the accuracy of eyewitness testimony, even where the eyewitness has no motivation to provide a biased testimony (for a review see Wells & Olson, 2003). The two components to source credibility (the more malevolent deceit and the less malevolent paucity of knowledge) have been termed, variously: veracity and sensitivity (Schum, 1981, 1994[1]); trustworthiness and expertise (Hahn, Oaksford, & Harris, 2012; Harris et al., 2016; Walton, 2008); helpfulness and knowledgeability (Shafto, Eaves, Navarro, & Perfors, 2012).

Concepts of source credibility are central to the evaluation of a communication of scientific consensus at two levels. The first level is the individual scientists themselves. There is a great difference in the convincingness of the agreement of 84% of medical scientists if they are known to be the 'crackpots' (low in trustworthiness and expertise), whereas the 16% who disagree with the statement are the respectable scientists high in both trustworthiness and expertise. Source credibility also becomes a relevant concept wherever an individual is not experiencing the level of consensus for themselves (e.g., by reading original journal articles or interviewing individual

---

[1] Schum additionally recognises the distinction between untruthful bias ('veracity') and 'truthful' bias ('objectivity'). Essentially, a source low in objectivity might report what they genuinely believe to be the true state of the world, but this belief has been influenced by their own bias, whilst low veracity represents a deliberate attempt to deceive.

scientists). As an illustration, throughout this manuscript we quote figures of scientific consensus. For you, the reader, therefore, not only is the credibility of the individual scientists within the consensus important, but also *our* credibility as accurate purveyors of information. In fact, as can be seen from citations accompanying the three examples at the outset of this article, we have not experienced the consensus levels ourselves, and consequently the credibility of Cook et al. (2013), Green (2016) and Whelton et al. (2002) become relevant. Figure 1 shows a Bayesian Network demonstrating the complexity of the inferences that must be made. In this network, the expertise and trustworthiness of the individual scientists are not represented explicitly due to space constraints, but they can be captured in the probabilistic relationships (the arrows) linking the true fact about ACC with what is reported by each expert (for an introduction to Bayesian Networks, see e.g., Korb & Nicholson, 2003; Pearl, 1988; on their application in this particular context see Cook & Lewandowsky, 2016; Hahn et al., 2016; for other examples see e.g., Fenton, Neil, & Lagnado, 2013; Harris et al., 2016; Jern et al., 2014; Kadane & Schum, 1996; Lagnado, 2011; Lagnado, Fenton, & Neil, 2013). Figure 1 demonstrates that the report of expert consensus on the topic of ACC presented at the outset of this article is 'third hand.' In addition to the potential for error (or bias) in the reports of each of the individual scientists (Experts 1-5), a reader might perceive Cook et al. as having a particular ideological bias[2], or Harris et al. (present paper) as too incompetent to accurately report on the report of Cook et al. All these factors are relevant to the question of how persuaded a reader *should* be by the scientific consensus presented at the outset of this article.

[2] Indeed, some potential readers (e.g., Legates, Soon, Briggs, Monckton of Brenchley, 2015) *have* perceived Cook et al. as having such a bias.

Following Hahn et al. (2016), in the present empirical work, we consider source credibility as it pertains to the *reporter* of the consensus (e.g., Cook et al. in Figure 1). Furthermore, although the two different components of source credibility (trustworthiness and expertise) appear conceptually important and have been demonstrated to be necessary to fully capture use of testimony (e.g., in children, Shafto et al., 2012), in the current paper we simplify the concept of source credibility and consider it as a single factor. Practically, it is however important to recognise that complete trust in the honesty (veracity) of a communicator, does not mean that they have not made honest mistakes (sensitivity).

### ***Lessons from decision making***

Having considered the formal question of what factors *should* influence the convincingness of a communication of expert consensus, we now turn to consider a potential lesson from descriptive research into decision making. Scholars of decision making have spent a considerable amount of research time and resources attempting to understand how people evaluate prospects such as 'A 90% chance of winning £10, otherwise nothing.' Indeed, gambles such as these have been deemed the 'fruit fly' of decision research (Lopes, 1983; see also Bateman, Dent, Peters, Slovic, & Starmer, 2007; Wulff, Mergenthaler-Canseco, & Hertwig, in press).

Research investigating evaluations of gambles like the one above has demonstrated that people's valuations (measured directly and through choice paradigms) are well described by two functions which are non-linear transformations of the probability described in the gamble (i.e., 90%) and the value described in the gamble (i.e., £10) (e.g., Kahneman & Tversky, 1979; Tversky & Kahneman, 1992; see e.g., Speekenbrink & Shanks, 2013, for a review). Specifically, and of central interest for the current paper, people's decision making proceeds as though they *under*weight high (though less than certain) probabilities, and *over*weight low probabilities (see also, e.g., Gonzalez & Wu, 1999). In the light of these findings, it is perhaps unsurprising that communication of high levels of scientific consensus has not had as much of an influence as might be hoped on people's overall perceptions of the consensus or decisions to engage in more sustainable behaviour. The 97% communicated in a statement such as the one at the start of this article is likely to be underweighted by recipients of the communication. Correspondingly, the 3% who (implicitly) are suggested not to agree with the consensus will be overweighted. More recent research into decisions from experience suggests a way in which the influence of the high consensus might be enhanced.

Barron and Erev (2003) shook up decision making research with their demonstration that a number of well-established and robust decision making phenomena did not hold (indeed, mostly were reversed) when the presentation of the gambles was altered. Specifically, rather than having the parameters of a given gamble described to them (e.g., "90% chance of winning £10, otherwise nothing"), participants learn the structure of the gamble through experience. For example, the gamble described above might be recreated by asking the participant to press a button multiple times (e.g., 100, 200, 400 times in the different experiments reported in Barron & Erev). On 90% of occasions, the button would yield £10, whilst on 10% of occasions it would yield nothing. Typically, in decisions from experience, participants would be required to choose between different buttons, with different payoffs.

The critical finding for the present paper is that when outcomes are experienced rather than described, the best fitting function capturing the influence of probability on decisions has the opposite properties to that observed for described outcomes. Namely, low probabilities are now underweighted and high probabilities are overweighted (e.g., Barron & Erev, 2003; Camilleri & Newell, 2011a; Fox & Hadar, 2006; Hau, Pleskac, &

Hertwig, 2010; Hau, Pleskac, Keifer, & Hertwig, 2008; Hertwig, Barron, Weber, & Erev, 2004; Rakow, Demes, & Newell, 2008; Ungemach, Chater, & Stewart, 2009; for an reviews see Hertwig & Erev, 2009; Wulff et al., in press). Although there are arguments that this so-called decision-experience ‘gap’ arises primarily from an under*experience* of the rare outcome, either in the whole, or in recently experienced samples (Camilleri & Newell, 2011b; Fox & Hadar, 2006; Rakow et al., 2008; see also Jarvstad, Hahn, Rushton, & Warren, 2013), the descriptive phenomenon remains. Typically, when gambles are experienced rather than described, people make choices as though they *over*weight high probabilities and *under*weight low probabilities. Thus, if we wish to maximise the influence of high scientific consensuses on people’s subsequent decisions (e.g., regarding sustainable behaviour), the decision making literature suggests that presenting people with the consensus in an experiential format will be more effective than merely describing it to them.

***How might we improve the communication of scientific consensus?***

In our experiments we focus on the importance of source credibility and the format of consensus communication. The literature reviewed above provides two clear suggestions for improving consensus communications: 1) Present the information from a reputable (trustworthy and expert) source; 2) Have people experience the consensus, rather than merely describing it to them. It is worth noting that the latter suggestion stems both from the lessons learnt from decision making research *and* from formal work on source credibility. In the case of the latter, pure experience of the scientists’ consensus (i.e., personal interviews with all the scientists), removes two ‘layers’ from Figure 1 (‘Cook et al.’ and ‘Harris et al.’). Consequently, there are fewer, potentially fallible, steps between the level of consensus of the scientists (zero – this is experienced directly) and the objective truth about ACC (one). Thus, we predict that the effect of a communication of expert consensus will be greater the more trustworthy and expert the source describing the consensus is, but it will be greatest where participants experience the consensus.

In the following, we present two experiments to test these predictions, noting that the manipulation of experience is somewhat impoverished. Nevertheless, we hope that these experiments will encourage future researchers to further test the effectiveness of *experiencing* expert consensus, given the strong theoretical support for its effectiveness.

## Experiment 1

### *Method*

#### *Participants*

After excluding four incomplete entries, 161 British nationals (75 female, 1 preferred not to say; aged 18-61 [median = 33]) were recruited through the online crowdsourcing platform Prolific Academic (www.prolific.ac). The participants were each paid £0.60. Ethical approval for both experiments in this paper was granted from the Departmental Ethics Chair for Speech, Hearing and Phonetic Sciences (University College London).

#### *Design*

Participants were presented with three out of nine possible communications, each one presenting consensus information about one of three topics. The first independent variable was *Topic* – participants were presented with consensus information on each of following three topics: climate change, politics (Brexit) and medicine. The second independent variable was *Format* – consensus information pertinent to the relevant topic was presented either by a reliable source or an unreliable source, or 'experienced.'

Each participant saw one (different) format for each topic, and each topic was presented only once.

The dependent variables were estimates of consensus levels (*perceived consensus* [out of 100]) and *beliefs* about the claims towards a topic, measured on a 5-point scale from "strongly disagree" [1] to "strongly agree" [5]. These were collected both before (prior) and after (posterior) participants received consensus information about each of the three topics.

*Materials*

Across all three communication formats (Experience, Reliable source and Unreliable source), the statements that experts agreed with were identical. For climate change, the relevant statement was 'Humans are causing global warming'; for the political topic, it was 'Leaving the European Union and the single market (Brexit) would be bad for the UK economy'; for the medical topic it was 'Regular aerobic exercise lowers blood pressure.'

The level of consensus presented within each topic was taken from an existing reputable source, with external validity prioritised over the internal validity associated with keeping the level of consensus communicated constant. For climate change, Cook et al.'s (2013) figure of 97% of climate scientists agreeing that humans are causing global warming was used; for the political topic (Brexit), the Ipsos MORI (2016) survey figure of 88% of economists agreeing that leaving the EU would be bad for the UK economy was used; and for the medical topic, Whelton et al.'s (2002) figure of 84% of medical scientists agreeing that regular aerobic exercise lowers diastolic blood pressure was used (the technical term 'diastolic' was not included in the experiment).

In the Experience condition, participants were presented with ten sets of ten fictional experts' silhouettes, each expert either agreeing or disagreeing with one of the three aforementioned statements. The level of agreement/disagreement was identical to the consensus reported by the reliable and unreliable sources. Although the order of the silhouettes within each set was fixed, their initial order had been randomised, and the order in which the ten sets were displayed was also randomised. To personify the fictional experts, they were each identified with pseudo-random initials (such that there were not a high number of uncommon initials; see Figure 2 for an example set from the climate change condition).

As the reliable sources, fictional anonymous professors of renowned UK universities (Oxford, Cambridge, LSE) were presented. For example, a participant presented with a reliable medical source would have read the following:

> Carefully consider the following report:
>
> ‘84% of medical scientists agree that aerobic exercise lowers blood pressure.’
>
> Source: Review by a professor of Medical Sciences, University of Cambridge

To determine the unreliable sources, a pilot study was run via Prolific Academic. For each of the three topics, 40 participants were presented with the appropriate aforementioned consensus information, and presented with a list of nine sources expected to vary in perceived trustworthiness and expertise. Participants were asked to rate the sources’ trustworthiness and expertise on a scale of 0-100. The following unreliable sources were subsequently chosen for the main experiment: a political columnist in a politically left-leaning publication for climate change; a political columnist in a politically right-leaning publication for Brexit; and a medical columnist in a politically right-leaning publication for the medical topic. For each of these three sources, the balance of bias (i.e. lack of trustworthiness) and expertise was the same, namely they were all rated as more untrustworthy than they were unknowledgeable.

Although for the climate change topic, a political columnist in a politically right-leaning publication was, similarly to the Brexit topic, rated as being the least trustworthy, we chose the left-leaning columnist instead. Because it is generally politically right-leaning individuals who deny climate change (e.g., Hornsey, Harris, Bain, & Fielding, 2016; McCright & Dunlap, 2011), a left-leaning columnist might be perceived as more likely to advance ‘left-wing propaganda’ advocating the reality of ACC.

Before being presented with any consensus information, participants completed a prior beliefs and attitudes questionnaire assessing: a) their beliefs about the claims relevant to the consensus information (such as “Humans are causing global warming” for the climate change topic, requiring an answer ranging from strongly disagree to strongly agree); b) how many experts out of every 100 they believed agreed with these same statements; c) to what degree they agreed with the statement “I am in favour of Britain leaving the European Union”; d) five questions about support for the free-market system (Heath & Gifford, 2006). These latter questions required the participant to indicate their belief in the free market system on a 5-point scale from ‘strongly disagree’ to ‘strongly agree’. Following Lewandowsky, Oberauer, & Gignac (2013), one question (‘I support the free-market system but not at the expense of environmental quality’) was left out, the authors finding that this question loaded on a second factor by itself. Lewandowsky, Gignac, and Vaughan (2013), and Cook and Lewandowsky (2016; see also van der Linden, Leiserowitz, & Maibach, in press) found that the effect of free market ideology on climate change scepticism was almost eliminated when providing Australian participants with ‘socially-normative’ information about the scientific consensus on climate change. For U.S.-based participants, however, the difference between free-market supporters and free-market opponents in climate change skepticism was enhanced by a consensus communication (Cook & Lewandowsky, 2016; but see van der Linden et al., 2015, for opposite results with party political affiliation, rather than free market beliefs). Free market beliefs (FMBs) were thus assessed to check for the robustness of the effects of displaying consensus information to participants. We additionally assessed attitudes towards Britain leaving the European Union to determine the scope of influence of the consensus communication.

*Procedure*

The experiment was run through the online data collection platform Qualtrics.com. After receiving instructions and giving consent to participate, participants completed the prior beliefs and attitudes questionnaire. Next, they were presented with consensus information on all three topics. After each presentation of consensus information, participants were asked to provide some demographic data before being asked to again indicate their beliefs about statements relevant to the consensus information just presented (the questions were identical to the first two questions in the prior beliefs and attitudes questionnaire outlined above). Finally, after proceeding through these questions for all three topics, the participants were again asked to indicate their agreement with the statement “I am in favour of Britain leaving the European Union”.

***Results***

In both Experiments 1 and 2, we analyse perceived consensus levels separately from beliefs about the claims. For each dependent variable, we use a linear mixed effects analysis as an alternative to a repeated-measures ANOVA, which would have been inappropriate as every participant did not complete every combination of Format and Topic. The models includes fixed effects for Format, Topic, and Prior/Posterior, as well as all two-way and three-way interactions. To account for the repeated-measures, the models included participant-specific random intercepts. All models were estimated with the lme4 package (Bates, Maechler, Bolker, & Walker, 2016) for the R software environment (R Core Team, 2017), and test results are based on the Satterthwaite approximation to the degrees of freedom, obtained with the R package lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017).

*Perceived consensus levels*

The linear mixed-effects model yielded no effects of, nor interactions involving, Format. Of most relevance, the Format x Prior/Posterior interaction was not significant, $F(2, 784.20) = 1.10$, $p = .34$. Figure 3 shows mean estimates of the consensus across the three topics. The first thing to note is that, whilst participants initially underestimated the level of consensus for Climate Change and Brexit, this was less apparent for the Medical topic. Subsequently, there was less of an increase in estimated levels of consensus for the Medical topic, as confirmed by a Topic x Prior/Posterior interaction, $F(2, 784.20) = 19.19$, $p < .001$. Despite this, there was still an overall main effect of the consensus communication (Prior/Posterior), $F(1, 784.72) = 65.25$, $p < .001$. Other than the unsurprising main effect of Topic, $F(2, 784.27) = 37.73$, $p < .001$, no other significant effects were observed (all *Fs* < 1).

*Beliefs about the claims*

Given the lack of effects on perceived consensus levels, it would be surprising to observe an effect on people’s beliefs about the claims. Inspection of Figure 4 indeed suggests, once again, that there is no effect of communication format in these data. A linear mixed-effects model confirmed this, yielding no effects of, nor interactions involving, Format. Again, critically, the Format x Prior/Posterior interaction was not significant ($F < 1$). Aside from main effects of Prior/Posterior, $F(1, 786.85) = 11.70$, $p < .001$, and Topic, $F(2, 786.52) = 35.92$, $p < .001$, there were no other significant effects (all *p*s > .16).

There was a significant interaction between FMBs (mean support was 2.5 [1-5 scale]) and Prior/Posterior in predicting beliefs about ACC, Beta = 0.12, $p = .04$. There was a significant relationship between beliefs about ACC and FMB before receiving the consensus communication, $r(159) = -.32$, $p < .001$, but not afterwards, $r(159) = -.12$, $p = .13$. These results are in line with those observed in Lewandowsky, Gignac and Vaughan (2013), van der Linden et al. (in press), and Cook and Lewandowsky's (2016) Australian participants, and counter to those in the latter's U.S. participants.

Nineteen of the 157 respondents who reported their attitude to Brexit both before and after the consensus communications provided different ratings at the two time points. Fourteen of these became less favourable towards Brexit, whilst 5 became more favourable.

### *Discussion*

The results of Experiment 1 were somewhat disappointing. Despite the clear theoretical justification for the predictions that the Format manipulation would affect the persuasiveness of the consensus communication, there was no evidence to support this. The effect of the consensus communication (main effect of Prior/Posterior) on participants' reported beliefs about the claims presented replicates other findings in the literature (Lewandowsky, Gignac, & Vaughan, 2013; van der Linden et al., 2014). Communicating information about actual levels of scientific consensus influences both perceptions of that consensus, as well as beliefs about the claims about which there is a consensus.

Despite the lack of support obtained in Experiment 1 for the main prediction that communication format would influence perceptions of scientific consensus levels, and associated beliefs, the basic design of the experiment was repeated in Experiment 2. Whilst Experiment 1 recruited UK participants via Prolific Academic, Experiment 2 recruited U.S. participants via Amazon Mechanical Turk. Whilst about 90% of people in the UK believe that climate change is happening (Capstick, Demski, Sposato, Pidgeon, Spence, & Corner, 2015), the number in the US is only about 70% (Leiserowitz, Maibach, Roser-Renouf, Feinberg, & Rosenthal, 2014). We thus anticipated greater variance in prior beliefs about ACC in this sample, potentially increasing the power of our manipulation. In addition, we (nearly) doubled the sample size.

## Experiment 2

### *Method*

#### *Participants*

Excluding two participants who reported problems viewing the silhouettes in the Experience condition, 300 U.S. nationals completed the experiment (103 females, 197 males; aged 18-65 [median = 31]). They were recruited through Amazon Mechanical Turk and each paid $0.50.

#### *Design, Materials, and Procedure*

The materials and procedure were, for the most part, identical to Experiment 1. Due to the Brexit vote being in the past by the time Experiment 2 was conducted, and the U.S. population perhaps not being overly involved or interested in the issue in the first place, the politics topic was changed to the issue of whether ‘undocumented immigrants currently living in the US should have a way to legally stay in the US, if certain legal requirements, for example the application for citizenship or permanent residence, are met’. The relevant consensus statistic (72% of Americans agree with the statement) was taken from the Pew Research Center (2015) survey. Due to this consensus representing the opinions of regular Americans, the 72% here represents social consensus rather than expert consensus. The political topic was therefore of less primary research interest, and mainly served to provide a third level of Topic to facilitate the presentation of three levels of Format. The questions about the participants' Brexit voting intentions were also removed.

Secondly, because in Experiment 1 the participants' prior consensus belief ratings for the medical topic were on average higher than the reported 'correct' consensus statistic, this number was raised from the previous 84% to 96% in Experiment 2, in order that prior consensus beliefs were mostly lower than the reported statistics, facilitating more meaningful comparisons between the climate change and medical topics. Participants were notified of this change in the debrief.

Finally, the university for all the reliable sources was changed to Harvard University to ensure relevance for U.S. participants.

### ***Results***

#### *Perceived consensus levels*

In the linear mixed-effects analysis, the critical Format x Prior/Posterior interaction was significant, $F(2, 1483.71) = 6.58$, $p < .001$. Figure 5 shows mean estimates of the consensus across the three topics. In each of the three topics, the greatest difference between prior and posterior estimates was in the Experience condition, as predicted. This is confirmed by the individual contrasts comprising the interaction, which show that the Prior/Posterior effect was greater in the Experience condition versus the Reliable and Unreliable conditions, $t(1483.70) = 3.63$, $p < .001$, whilst there was no significant difference in the Prior/Posterior effect between the Reliable and Unreliable conditions, $t(1483.70) = 0.14$, $p = .89$. In addition, the Format x Prior/Posterior interaction was not moderated by Topic, $F(4, 1483.71) = 1.30$, $p = .27$. Once again, overall, there was a significant main effect of the consensus communication (Prior/Posterior), $F(1, 1483.84) = 865.22$, $p < .001$, as well as uninteresting effects of Topic, $F(2, 1483.84) = 865.22$, $p < .001$, and a Topic x Prior/Posterior interaction, $F(2, 1483.72) = 12.51$, $p < .001$.

*Beliefs about the claims*

Having observed an effect of communication format on perceived consensus levels, a downstream effect might also be observed on people's beliefs about the claims. Figure 6 shows that, numerically, the pattern of results is in the predicted direction, with the greatest change in belief observed in the Experience condition across all three topics. There was no evidence, however, that this was a reliable result, with the Format x Prior/Posterior interaction non-significant, $F(2, 1484.07) = 1.54$, $p = .21$. The only significant effects were the main effects of Prior/Posterior, $F(1, 1484.07) = 19.99$, $p < .001$, and Topic, $F(2, 1484.17) = 161.31$, $p < .001$. No other significant effects were observed (all *p*s > .13).

In Experiment 2, beliefs about ACC were not predicted by an interaction between FMB (mean support 2.6/5) and Prior/Posterior, Beta = 0.03, $p = .37$, but FMB did predict beliefs both before, $r(298) = -.58$, $p < .001$, and after, $r(298) = -.54$, $p < .001$, consensus communications.

***Discussion***

Once again, an overall effect of consensus communication on reported beliefs was observed. Of more relevance to the current paper, and more encouragingly than Experiment 1, Experiment 2 showed the greatest change in perceived consensus levels was brought about in the Experience condition. This provides the first direct evidence that such a presentation might be beneficial in expert consensus communications. That the effect did not filter through to influence reported beliefs about the claims presented, despite an effect of consensus communication on said beliefs, might be a result of the small effect size on perceived consensus levels, which was not able to influence responses on a 5-point response scale.

## General Discussion

We have set out conditions that we expect to facilitate the communication of expert consensus information, specifically highlighting the role of 'experiential' communications from reliable experts. Two experiments provide mixed support for the prediction that such communications will lead to greater change in perceived consensus levels and related beliefs. Experiment 1 provided no support for the prediction in a UK sample, whilst Experiment 2 (U.S. sample) observed a small, but reliable, effect on perceived consensus levels, but no effect on beliefs about the statements being made.

There are a number of possible reasons for the mixed support for our hypotheses, and space constraints prevent us from considering them all here. Of course, the first is that the hypotheses were wrong. Given the justification for the hypotheses from different areas of psychology, however, we are reluctant to arrive at that conclusion too swiftly. We believe that a search for appropriate means for conveying consensus information 'experientially' is a fruitful one. We therefore primarily focus on two alternative, non-exclusive, reasons for the mixed nature of our data. One is that our operationalisation of 'experience' was insufficient. The other relates to the potential for ceiling effects.

Support for the potential for ceiling effects operating in these data is evident from a comparison of the two experiments, focussing on the climate change scenario, since all elements of that scenario were common across the two experiments. In both experiments, participants were informed that the scientific consensus was 97%. Averaging across all experimental conditions, participants' prior estimates of consensus were 80% in Experiment 1 and 79% in Experiment 2[3]. Thus, the potential for updating beliefs is the same across experiments. The U.S. participants in Experiment 2 typically updated their estimates less on receipt of the consensus communication. If one uses the Unreliable condition as a benchmark (since this was the condition in which we expected participants to update least), U.S. participants in Experiment 2, on average, only revised their consensus estimates to 83%, whilst U.K. participants in Experiment 1 revised their estimates to 92%. The difference between 92% and 97% leaves little room for the other Format conditions to exert additional influence, which provides a parsimonious explanation for why an effect was observed in Experiment 2, but not in Experiment 1[4]. Similarly, even in the U.S., mean ratings of belief in human caused climate change were 4.1 (out of 5) at the start of the experiment, and 4.3 after receiving an Unreliable consensus communication. Thus, once more, the skew of responses towards the top of the scale likely contributed to the lack of an effect on belief ratings.

Ceiling effects might therefore have compromised the power of the experiments presented. This is, however, something of a conjecture, and future research should seek to test the influence of experience in consensus communications using messages where:

---

[3] These estimates are higher than usually observed; we did not recruit nationally representative samples.

[4] Cook and Lewandowsky (2016) noted different responses of Australian and U.S. participants to consensus communications, especially noting that U.S. (but not Australian) free market endorsers showed reduced trust in scientists following a consensus communication. Whilst we argue that ceiling effects offer a parsimonious account for the differences observed between our Experiments 1 and 2, we cannot rule out alternative cultural explanations.

a) participants greatly underestimate the expert consensus; and b) a standard consensus communication results in new estimates that are still considerably below the actual expert consensus. We suggest that a critical aim of such studies should be to identify most effective methods for an ‘experiential’ communication of expert consensus. The ‘experience’ in the current experiments was minimal, and yet we observed some support for our hypothesis that this will lead to greater belief updating in terms of consensus estimates. Future studies should seek to identify an optimum compromise between a practical communication and the amount of information provided about the individual experts providing the consensus. The ’97 hours’ project (https://skepticalscience.com/nsh/#) presents 100 characters on the screen. The user can click on any character to reveal the individual the cartoon is representing, their expert credentials, and a quote. Subsequent research could measure the influence of presentations such as this.

We hypothesised a benefit of an ‘experiential’ communication based on considerations of source credibility, and from observations in the decision making literature. It is also true, however, that the Experience condition required participants to spend more time thinking about the expert consensus. Whilst this could be seen as a confound, we see it as an integral part of experience. Experience will typically require more time and effort than simple testimony (running the experiments and analysing the data oneself – *ultimate* experience – would take longer still!). We are non-committal on the precise psychological mechanisms that could underlie the benefits of the Experience condition. As specified in the Introduction, the decisions from experience literature suggests a few contributing mechanisms (for reviews see Hertwig & Erev, 2009; Wulff et al., in press; see also Hills & Hertwig, 2010; Jarvstad et al., 2013), in addition to the potential for fundamentally different decision processes being involved (Hau et al., 2008; Hills & Hertwig, 2010). Amount of time and attention (in addition to formal considerations of source expertise) are potentially additional contributing factors. Indeed, the fact that there are many factors that could generate such an advantage, to us makes it all the more important that the potential for such a benefit is fully explored.

### *Conclusion*

Expert consensus is fundamental to the continual development of society and technological advancement. In some domains, especially those with political considerations, public support is required for necessary policy changes in light of expert consensus. Consequently, how information about expert consensus is best communicated is an important applied question. From the literatures on source credibility and decision making, we suggested that the potential to experience high levels of expert consensus would be more effective than a simple numerical description of that consensus. We observed some support for such a prediction in one of two experiments testing this prediction. We argue that the mixed nature of these results potentially stems from ceiling effects. Given the strong theoretical justification for the predictions, we encourage future researchers to seek to identify optimum methods for presenting consensus information ‘experientially.’

## References

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making, 16,* 215-233.

Bateman, I., Dent, S., Peters, E., Slovic, P., & Starmer, C. (2007). The affect heuristic and the attractiveness of simple gambles. *Journal of Behavioral Decision Making, 20,* 365-380.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67,* 1-48.

Bolsen, T., Leeper, T. J., & Shapiro, M. A. (2014). Doing what others do: norms, science, and collective action on global warming. *American Politics Research, 42,* 65-89.

Bovens, L., & Hartmann, S. (2003). *Bayesian Epistemology*. Oxford, UK: Oxford University Press.

Camilleri, A. R., & Newell, B. R. (2011a). When and why rare events are underweighted: A direct comparison of the sampling, partial feedback, full feedback and description choice paradigms. *Psychonomic Bulletin & Review, 18,* 377-384.

Camilleri, A. R., & Newell, B. R. (2011b). Description- and experience-based choice: Does equivalent information equal equivalent choice? *Acta Psychologica, 136,* 276-284.

Capstick, S. B., Demski, C. C., Sposato, R. G., Pidgeon, N. F., Spence, A., & Corner, A. (2015). Public perceptions of climate change in Britain following the winter 2013/2014 flooding. *Understanding Risk Research Group Working Paper*, 15-01.

Chaiken, S. (1980). Heuristic versus systematic information processing and the use of source versus message cues in persuasion. *Journal of Personality and Social Psychology, 39*, 752-766.

Chaiken, S. & Maheswaran, D. (1994). Heuristic processing can bias systematic processing: effects of source credibility, argument ambiguity, and task importance on attitude judgement. *Journal of Personality and Social Psychology, 66*, 460-473.

Coady, C.A.J., (1992). *Testimony: A Philosophical Study*. Oxford: Oxford University Press.

Condorcet, Nicolas Caritat de (1785). *Essai sur l'Application de l'Analyse a la Probabilite des Decisions Rendues a la Pluralite des Voix.* Paris.

Cook, J. (2016). Countering Climate Science Denial and Communicating Scientific Consensus. *Oxford Encyclopedia of Climate Change Communication*. Oxford: Oxford University Press.

Cook, J., & Jacobs, P. (2014). Scientists are from Mars, laypeople are from Venus: An evidence-based rationale for communicating the consensus on climate. *Reports of the National Center for Science Education, 34.6,* 3.1-3.10.

Cook, J., & Lewandowsky, S. (2016). Rational irrationality: Modeling climate change belief polarization using Bayesian networks. *Topics in Cognitive Science, 8,* 160-179.

Cook, J., et al. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters, 8,* 024024.

Deer, B. (2004*). Exposed: Andrew Wakefield and the MMR-autism fraud*. Retrieved from http://briandeer.com/mmr/lancet-summary.htm

Ding D, Maibach E W, Zhao X, Roser-Renouf C and Leiserowitz A (2011) Support for climate policy and societal action are linked to perceptions about scientific agreement. *Nature Climate Change, 1*, 462–465.

Fenton, N., Neil, M., & Lagnado, D. A. (2013). A general structure for legal arguments about evidence using Bayesian Networks. *Cognitive Science, 37,* 61-102.

Fox, C. R., & Hadar, L. (2006). "Decisions from experience" = sampling error + prospect theory: Reconsidering Hertwig, Barron, Weber & Erev (2004). *Judgment and Decision Making, 1,* 159–161.

Friedman, R. (1987). Route analysis of credibility and hearsay. *The Yale Law Journal, 96*, 668–742.

Gonzalez, R., & Wu, G. (1999). On the shape of the probability weighting function. *Cognitive Psychology, 38,* 129–166.

Green, C. (2016, February 26). EU referendum: Nine out of ten economists say Brexit will damage the economy. *The Independent.* Retrieved from http://www.independent.co.uk/news/uk/politics/eu-referendum-nine-out-of-ten-economists-say-brexit-will-damage-economy-a6898886.html

Hahn, U., Harris, A. J. L., & Corner, A. (2009). Argument content and argument source: An exploration. *Informal Logic*, *29,* 337-367.

Hahn, U., Harris, A. J. L., & Corner, A. (2016). Public reception of climate science: Coherence, reliability and independence. *Topics in cognitive science, 8,* 180-195.

Hahn, U., Oaksford, M., & Harris, A. J. L. (2012). Testimony and argument: A Bayesian perspective. In F. Zenker (Ed.), *Bayesian argumentation* (pp. 15–38). Dordrecht NL: Springer.

Harris, A. J. L., & Hahn, U. (2009). Bayesian rationality in evaluating multiple testimonies: Incorporating the role of coherence. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 1366-1372.

Harris, A. J. L., Hahn, U., Madsen, J. K., & Hsu, A. S. (2016). The Appeal to Expert Opinion: Quantitative support for a Bayesian Network approach. *Cognitive Science, 40,* 1496-1533.

Hau, R., Pleskac, T. J., & Hertwig, R. (2010). Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities. *Journal of Behavioral Decision Making, 23,* 48–68.

Hau, R., Pleskac, T. J., Kiefer, J., & Hertwig, R. (2008). The description-experience gap in risky choice: The role of sample size and experienced probabilities. *Journal of Behavioral Decision Making, 21,* 1–26.

Heath, Y., & Gifford, R. (2006). Free-market ideology and environmental degradation the case of belief in global climate change. *Environment and Behavior, 38*, 48-71.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science, 15,* 534–539.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in Cognitive Sciences, 13,* 517–523.

Hills, T., & Hertwig, R. (2010). Information search and decisions from experience: Do our patterns of sampling foreshadow our decisions? *Psychological Science, 21,* 1787-1792.

Hornsey, M. J., Harris, E. A., Bain, P. G., & Fielding, K. S. (2016). Meta-analyses of the determinants and outcomes of belief in climate change. *Nature Climate Change*, *6,* 622-626.

Ipsos MORI (2016, May 28). *Economists' Views on Brexit.* Retrieved from https://www.ipsos-mori.com/researchpublications/researcharchive/3739/Economists-Views-on-Brexit.aspx.

Jarvstad, A., Hahn, U., Rushton, S. K., & Warren, P. A. (2013). Perceptuo-motor, cognitive, and descriptive decision-making seem equally good. *Proceedings of the National Academy of Sciences, 110,* 16271-16276.

Jern, A., Chang, K. K., & Kemp C. (2014). Belief polarization is not always irrational. *Psychological Review. 121,* 206-224.

Kadane, J. B., & Schum, D. A. (1996). *A probabilistic analysis of the Sacco and Vanzetti evidence*. New York: John Wiley & Sons.

Kahan, D. M. (2015). Climate-science communication and the measurement problem. *Political Psychology, 36 (Supplement 1),* 1-43.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263-291.

Kirk, R. E. (1995). *Experimental design: Procedures for the behavioural sciences.* London: Brooks/Cole.

Korb, K. B., & Nicholson, A. E. (2003). *Bayesian Artificial Intelligence*. Boca Raton, FL: CRC Press.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTestPackage: Tests in linear mixed effects models. *Journal of Statistical Software, 82,* 1-26.

Ladha, K. K. (1992). The Condorcet jury theorem, free speech, and correlated votes. *American Journal of Political Science*, 617–634.

Lagnado, D. A. (2011). Thinking about evidence. In P. Dawid, W. Twining, & M. Vasaliki (Eds.), *Evidence, inference and enquiry* (pp. 183-223). Oxford, UK: Oxford University Press/British Academy.

Lagnado, D. A., Fenton, N., & Neil, M. (2013). Legal idioms: A framework for evidential reasoning. *Argument and Computation, 4*, 46–63.

Legates, D. R., Soon, W., Briggs, W. M., & Monckton of Brenchley, C. (2015). Climate consensus and 'misinformation': A rejoinder to 'Agnotology, scientific consensus, and the teaching and learning of climate change. *Science & Education, 24,* 299-318.

Leiserowitz, A., Maibach, E., Roser-Renouf, C., Feinberg, G., Rosenthal, S. (2014). *Climate change in the American Mind.* Yale University and George Mason University. New Haven, CT: Yale Project on Climate Change Communication. Retrieved from: http://environment.yale.edu/climate-communication/files/Climate-Change-American-Mind-April-2014.pdf.

Lewandowsky, S., Gignac, G., & Vaughan, S. (2013). The pivotal role of perceived scientific consensus in acceptance of science. *Nature Climate Change, 3,* 399-404.

Lewandowsky, S., Oberauer, K., & Gignac, G. E. (2013). NASA faked the moon landing—therefore, (climate) science is a hoax an anatomy of the motivated rejection of science. *Psychological science*, *24*, 622-633.

List, C. (2004). On the significance of the absolute margin. *British Journal for the Philosophy of Science, 55*(3), 521–544.

Lopes, L. L. (1983). Some thoughts on the psychological concept of risk. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 137-144.

McCright, A. M., Dunlap, R. E. (2011). Cool dudes: The denial of climate change among conservative white males in the United States. *Global Environmental Change, 21*, 1163-1172.

McCright, A. M., Dunlap, R. E., & Xiao, C. (2013) Perceived scientific agreement and support for action on climate change in the USA. *Climatic Change 119*, 511-518.

Pearl, J. (1988*). Probabilistic reasoning in intelligent systems.* Palo Alto, CA: Morgan Kaufmann.

Petty, R.E., & Cacioppo, J.T. (1984). Source factors and the elaboration likelihood model of persuasion. *Advances in Consumer Research, 11,* 668-672.

Petty, R. E., Cacioppo, J. T., & Goldman, R. (1981). Personal involvement as a determinant of argument-based persuasion. *Journal of Personality and Social Psychology, 41*, 847-855.

Pew Research Centre (2015). *Broad Public Support for Legal Status for Undocumented Immigrants*. Retrieved from: http://www.people-press.org/2015/06/04/broad-public-support-for-legal-status-for-undocumented-immigrants/.

R Core Team (2017). R: A language and environment for statistical computing. R foundation for statistical computing, Vienna, Austria. URL https://www.R-project.org/.

Rakow, T., Demes, K. A., & Newell, B. R. (2008). Biased samples not mode of presentation: Re-examining the apparent underweighting of rare events in experience-based choice. *Organizational Behavior and Human Decision Processes, 106,* 168–179.

Schum, D. A. (1981). Sorting out the effects of witness sensitivity and response-criterion placement upon the inferential value of testimonial evidence. *Organizational Behavior and Human Performance, 27,* 153-196.

Schum, D. A. (1994). The evidential foundations of probabilistic reasoning. New York: Wiley.

Shafto, P., Eaves, B., Navarro, D. J., & Perfors, A. (2012). Epistemic trust: Modeling children’s reasoning about others’ knowledge and intent. *Developmental Science, 15,* 436-447.

Speekenbrink, M., & Shanks, D.R. (2013). Decision making. In: D. Reisberg (Ed.), Oxford handbook of cognitive psychology (pp. 682-703). New York: Oxford University Press.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297-323.

Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted, when rare outcomes are experienced (rarely)? *Psychological Science, 20*, 473–479.

van der Linden, S. L., Leiserowitz A. A., Feinberg, G. D., Maibach, E. W. (2014). How to communicate the scientific consensus on climate change: plain facts, pie charts or metaphors? *Climatic Change*, 126, 255–262.

van der Linden, S. L., Leiserowitz, A. A., Feinberg, G. D., & Maibach, E. W. (2015). The scientific consensus on climate change as a gateway belief: Experimental evidence. *PloS one*, *10*(2), e0118489.

van der Linden, S. L., Leiserowitz, A. A., & Maibach, E. W. (in press). Scientific agreement can neutralize politicization of facts. *Nature Human Behaviour.*

Walton, D. (2008). *Informal logic: A pragmatic approach (2nd ed.)*. New York: Cambridge University Press.

Wells, G. L., & Olson, E. A. (2003). Eyewitness testimony. *Annual Review of Psychology, 54*, 277–295.

Whelton, S. P., Chin, A., Xin, X., & He, J. (2002). Effect of aerobic exercise on blood pressure: A meta-analysis of randomized, controlled trials. *Annals of Internal Medicine*, *136*, 493-503.

Wulff, D. U., Mergenthaler-Canseco, M., & Hertwig, R. (in press). A meta-analytic review of two modes of learning and the decision-experience gap. *Psychological Bulletin.*

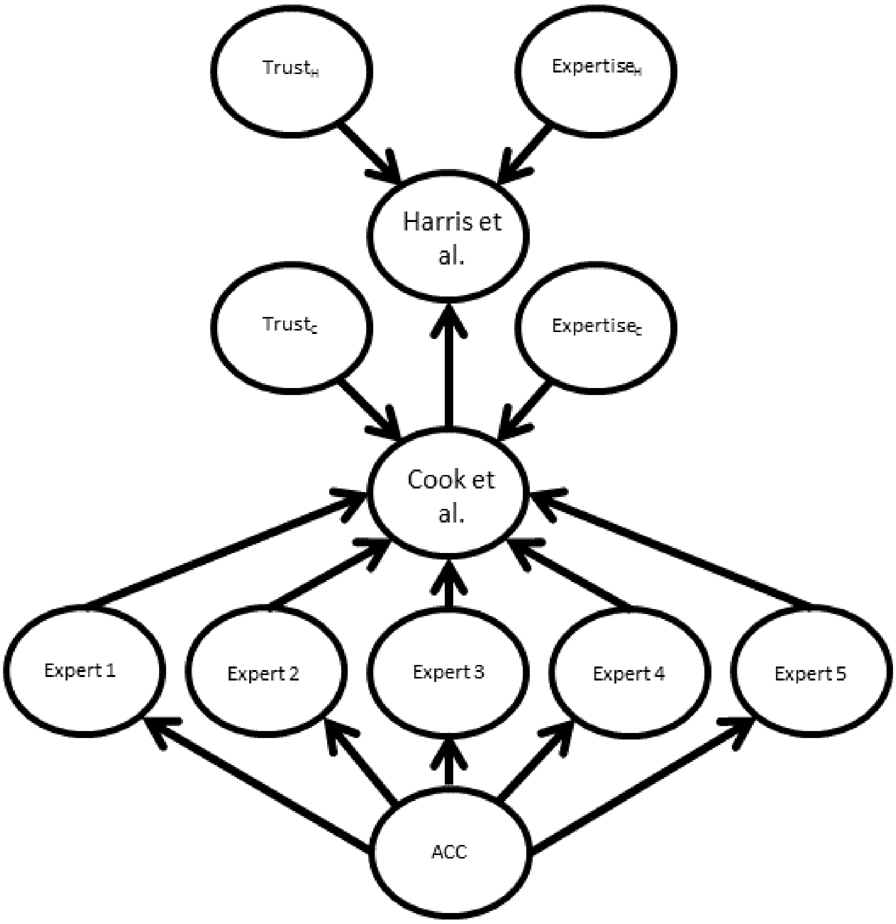



*Figure 1.* The testimonial ‘steps’ in the report at the outset of this article that “97% of climate scientists agree that humans are causing global warming.” (The 5 scientists are illustrative and kept to 5 solely due to space constraints).

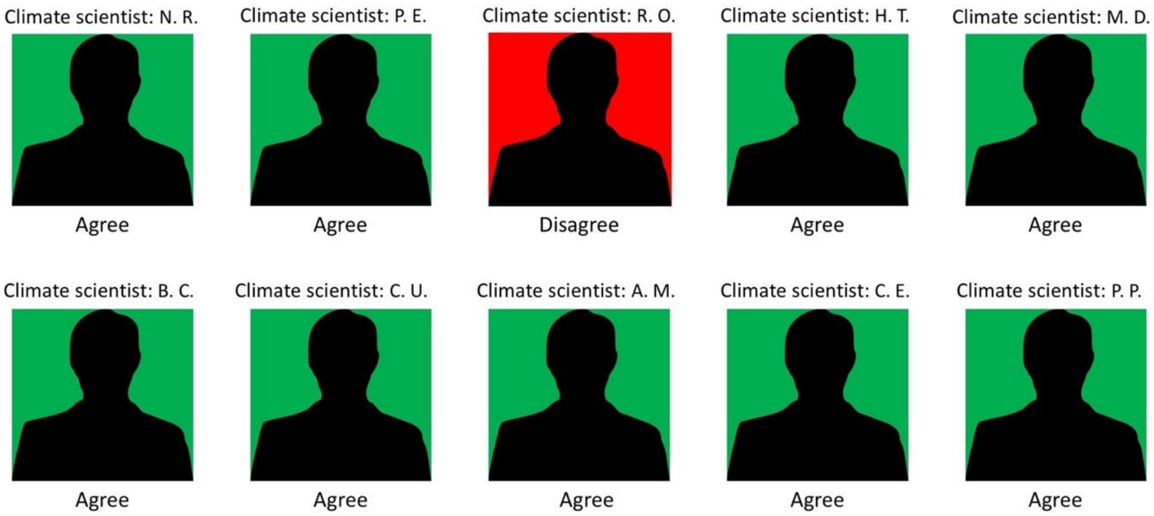
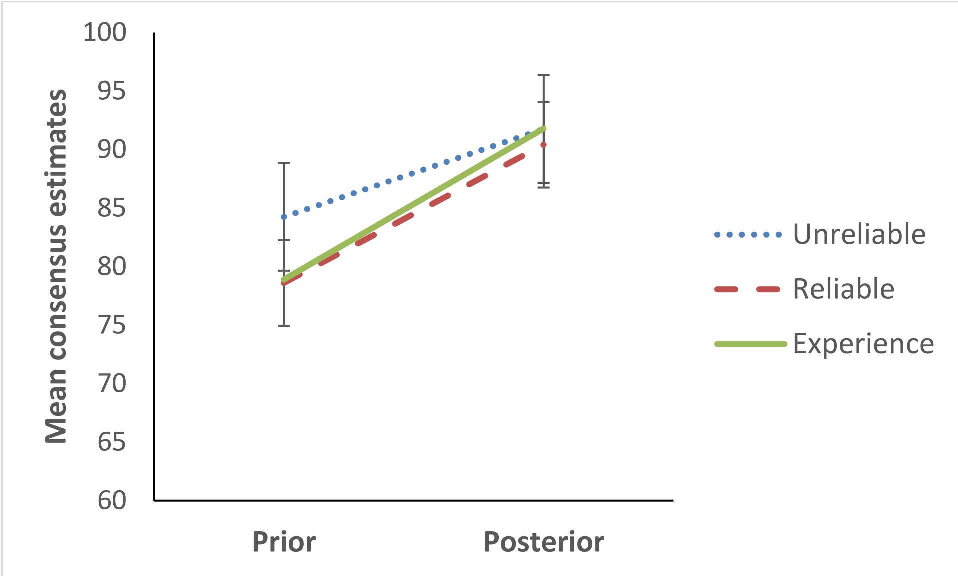
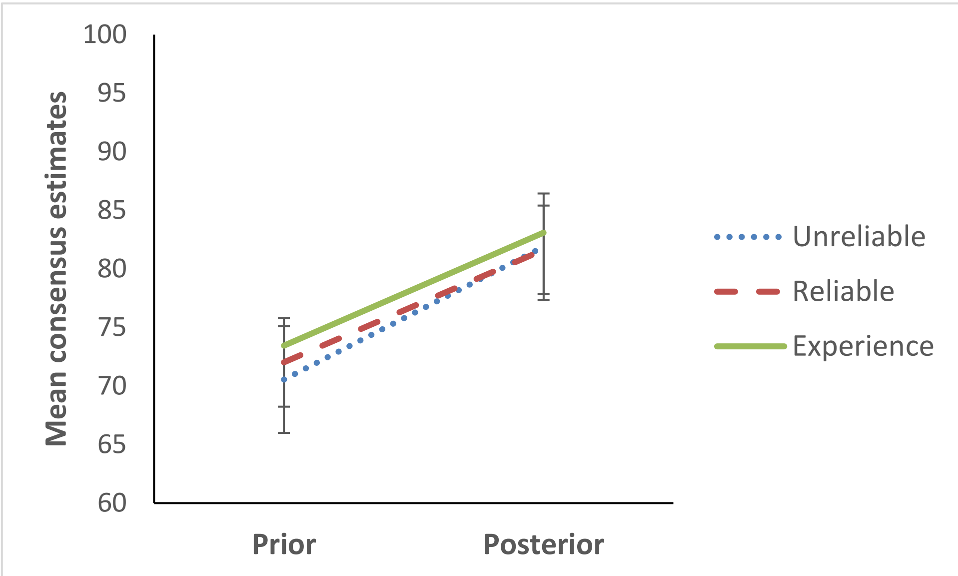



*Figure 2.* Climate scientists' silhouettes in the Experience condition (see the online version for a color version).


100
95
90
85
80
75
70
65
60
Mean consensus estimates
Prior
Posterior
Unreliable
Reliable
Experience


100
95
90
85
80
75
70
65
60
Mean consensus estimates
Prior
Posterior
Unreliable
Reliable
Experience

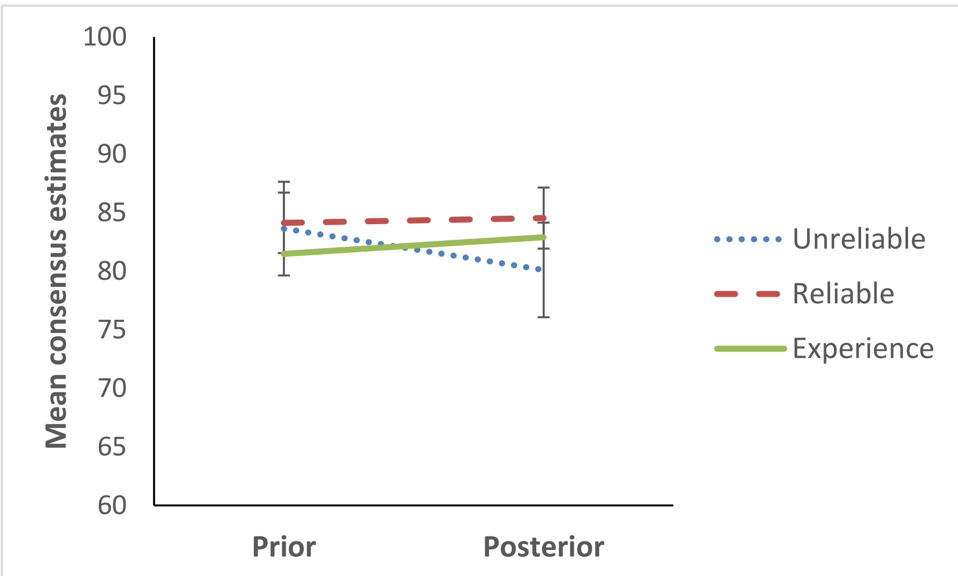
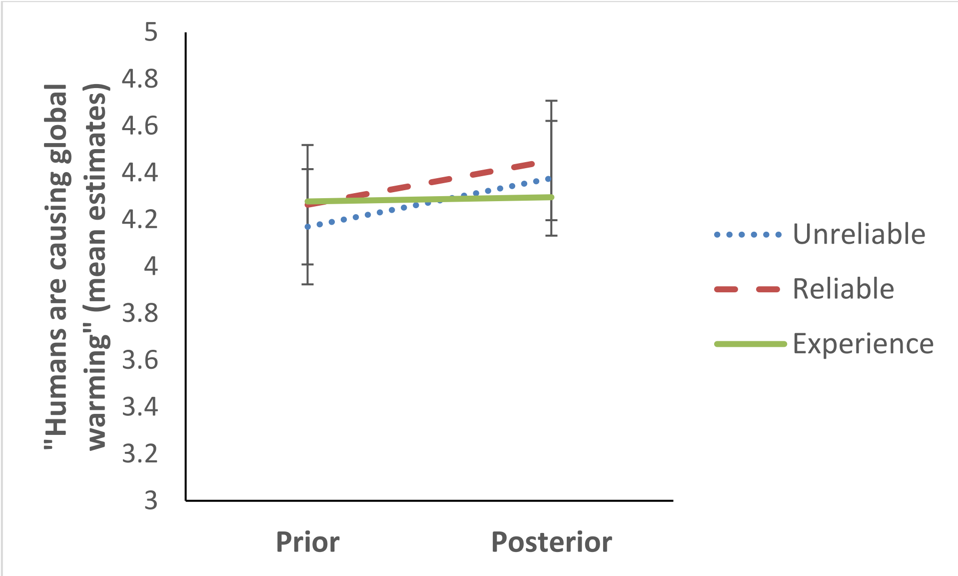
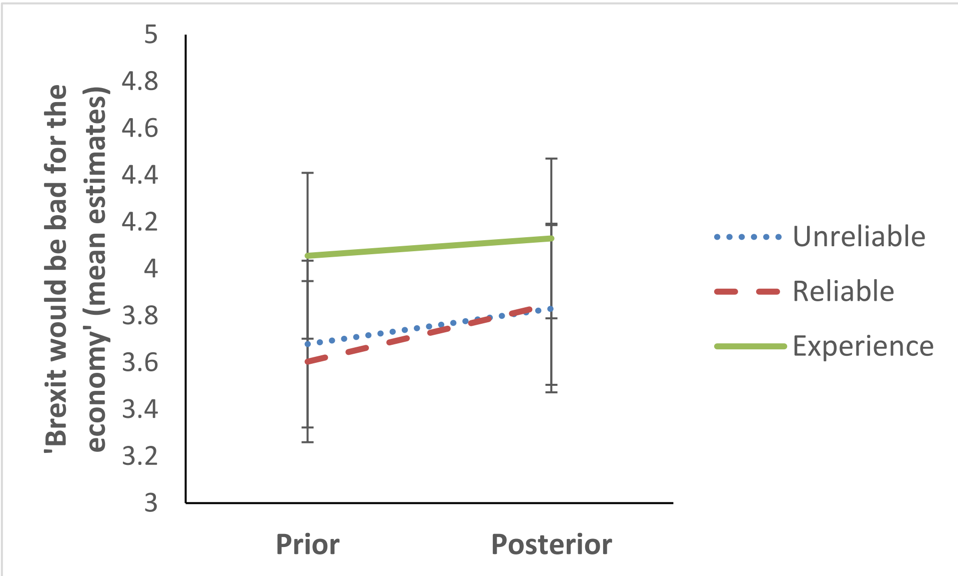

100
95
90
85
80
75
70
65
60
Mean consensus estimates
Prior
Posterior
Unreliable
Reliable
Experience

*Figure 3.* Prior and posterior estimates of the expert consensus across the 3 communication formats for the Climate (top), Political (middle) and Medical (bottom) topics in Experiment 1. Error bars represent 95% confidence intervals.


"Humans are causing global warming" (mean estimates)
5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
3.2
3
Prior
Posterior
Unreliable
Reliable
Experience


'Brexit would be bad for the economy' (mean estimates)
5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
3.2
3
Prior
Posterior
Unreliable
Reliable
Experience

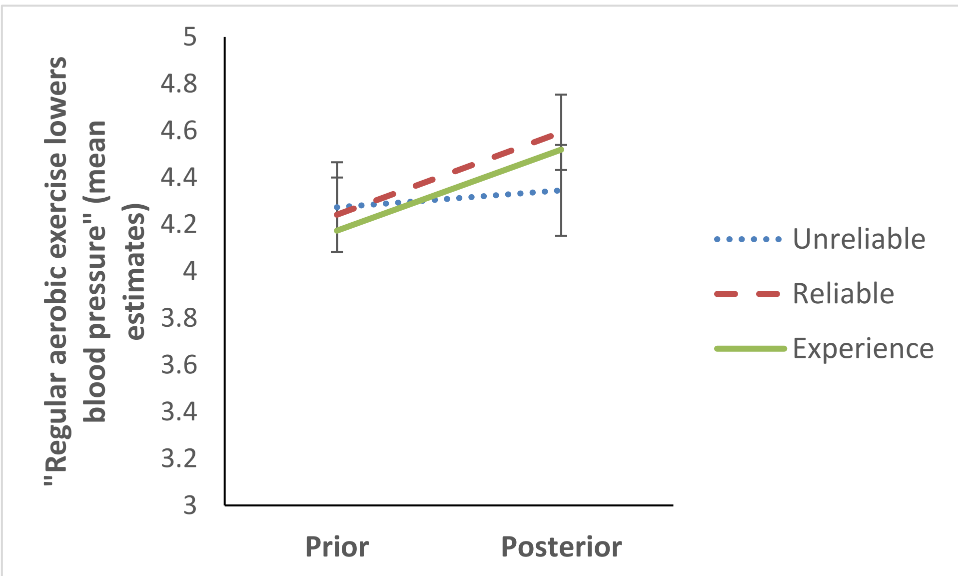
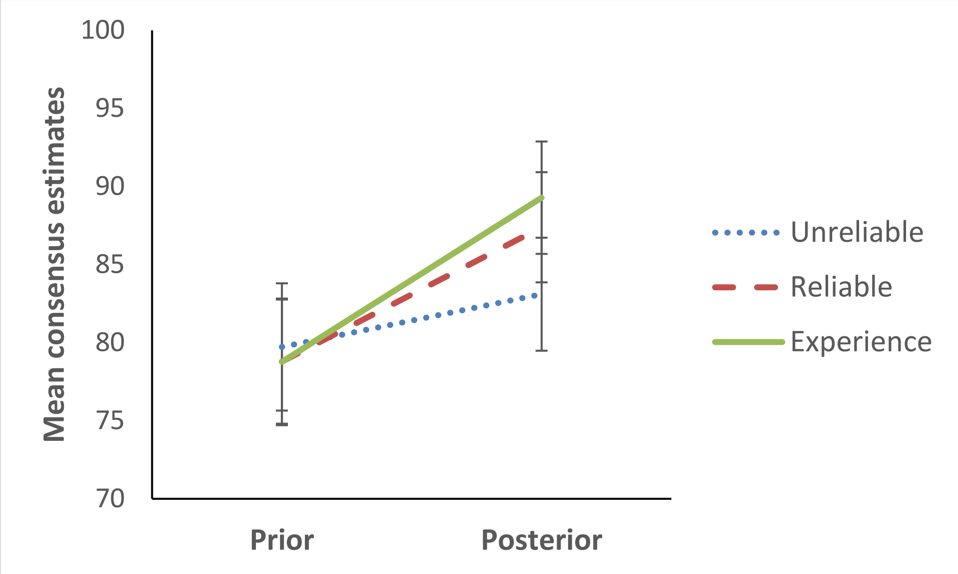
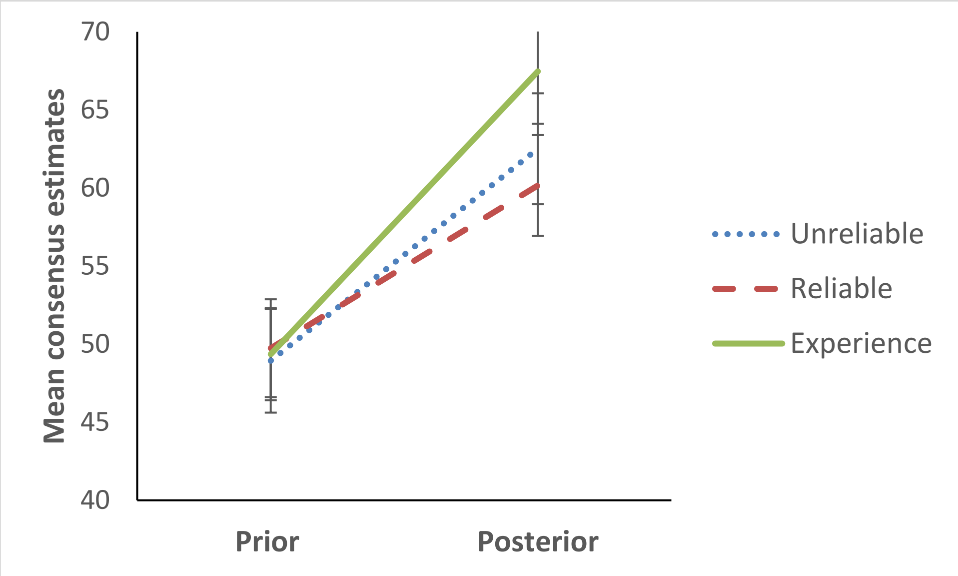

"Regular aerobic exercise lowers blood pressure" (mean estimates)
5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
3.2
3
Prior
Posterior
Unreliable
Reliable
Experience

*Figure 4.* Prior and posterior beliefs about the claims made in the consensus statements for the Climate (top), Political (middle) and Medical (bottom) topics in Experiment 1. Error bars represent 95% confidence intervals.


100
95
90
85
80
75
70
Mean consensus estimates
Prior
Posterior
Unreliable
Reliable
Experience


70
65
60
55
50
45
40
Mean consensus estimates
Prior
Posterior
Unreliable
Reliable
Experience

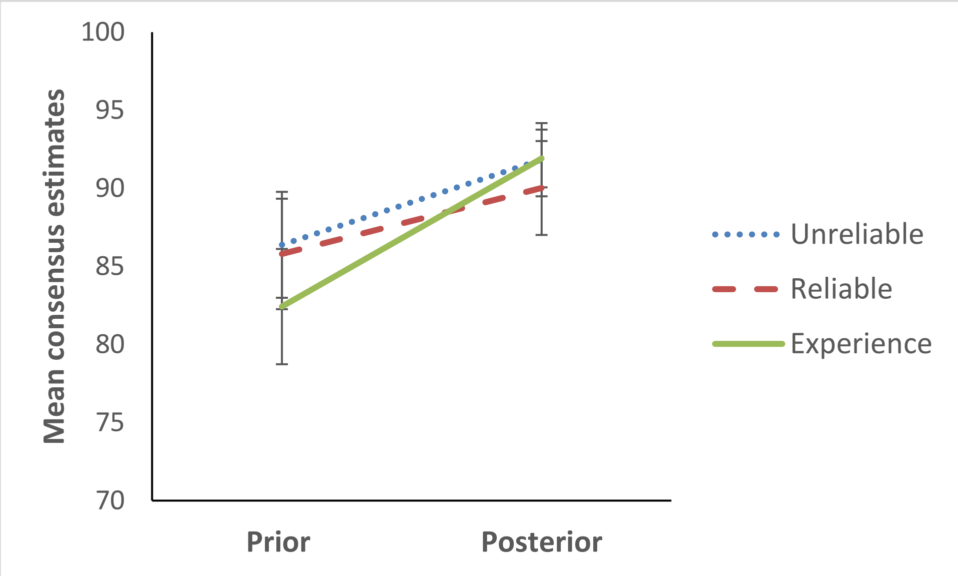
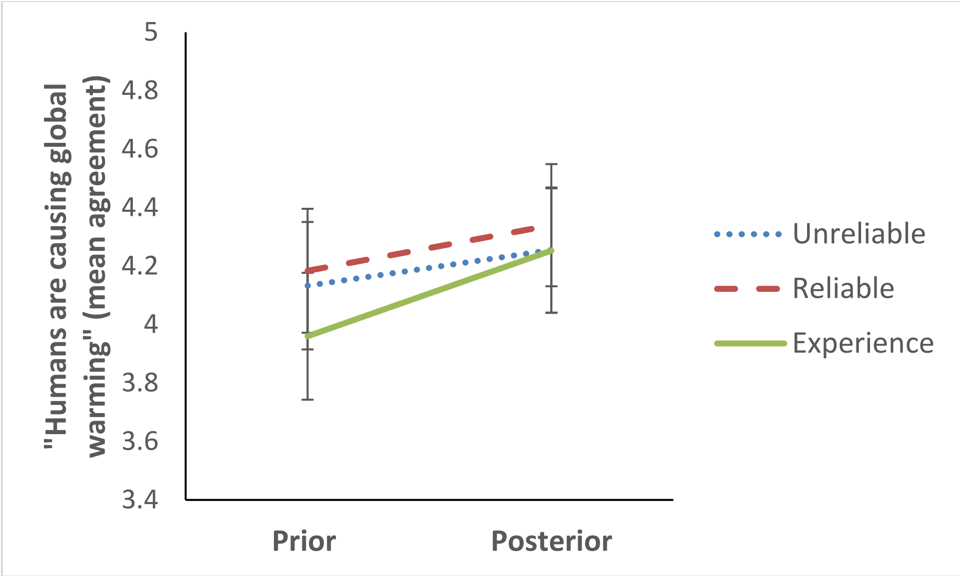
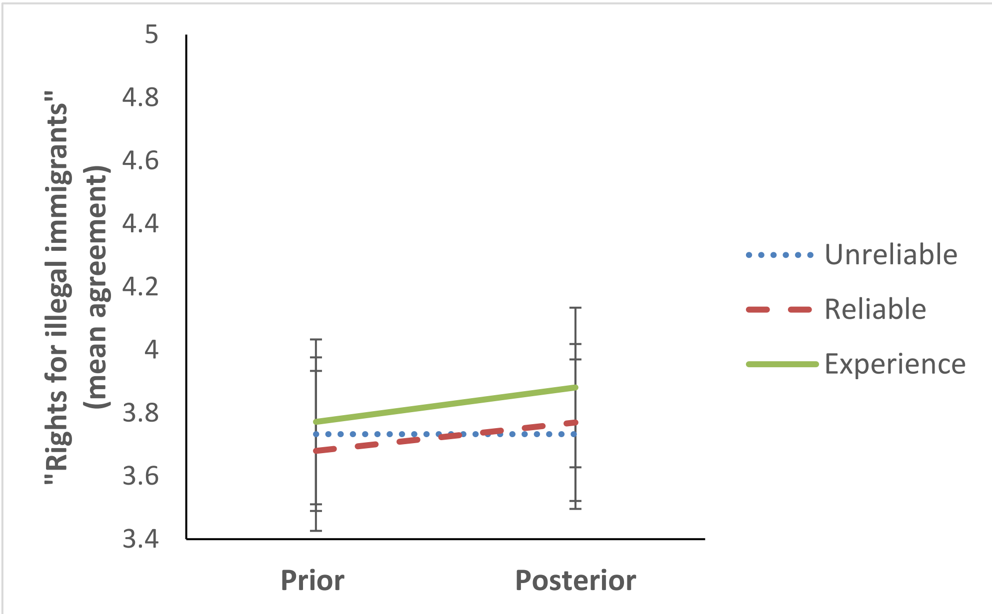

100
95
90
85
80
75
70
Mean consensus estimates
Prior
Posterior
Unreliable
Reliable
Experience

*Figure 5.* Prior and posterior estimates of the expert consensus across the 3 communication formats for the Climate (top), Political (middle) and Medical (bottom) topics in Experiment 2. Error bars represent 95% confidence intervals.


"Humans are causing global warming" (mean agreement)
5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
Prior
Posterior
Unreliable
Reliable
Experience


"Rights for illegal immigrants" (mean agreement)
5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
Prior
Posterior
Unreliable
Reliable
Experience

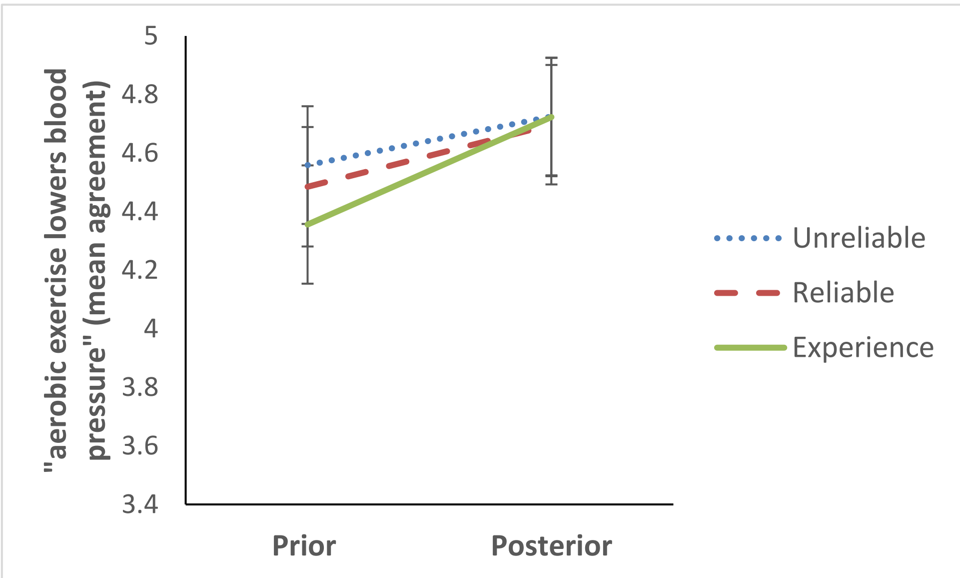

"aerobic exercise lowers blood pressure" (mean agreement)
5
4.8
4.6
4.4
4.2
4
3.8
3.6
3.4
Prior
Posterior
Unreliable
Reliable
Experience

*Figure 6.* Prior and posterior beliefs about the claims made in the consensus statements for the Climate (top), Political (middle) and Medical (bottom) topics in Experiment 2. Error bars represent 95% confidence intervals.

**Figure captions**

*Figure 1.* The testimonial 'steps' in the report at the outset of this article that "97% of climate scientists agree that humans are causing global warming." (The 5 scientists are illustrative and kept to 5 solely due to space constraints).

*Figure 2.* Climate scientists' silhouettes in the Experience condition (see the online version for a color version).

*Figure 3.* Prior and posterior estimates of the expert consensus across the 3 communication formats for the Climate (top), Political (middle) and Medical (bottom) topics in Experiment 1. Error bars represent 95% confidence intervals.

*Figure 4.* Prior and posterior beliefs about the claims made in the consensus statements for the Climate (top), Political (middle) and Medical (bottom) topics in Experiment 1. Error bars represent 95% confidence intervals.

*Figure 5.* Prior and posterior estimates of the expert consensus across the 3 communication formats for the Climate (top), Political (middle) and Medical (bottom) topics in Experiment 2. Error bars represent 95% confidence intervals.

*Figure 6.* Prior and posterior beliefs about the claims made in the consensus statements for the Climate (top), Political (middle) and Medical (bottom) topics in Experiment 2. Error bars represent 95% confidence intervals.