Consensus versus Unanimity: Which Carries More Weight?

Finnur Dellsén

Forthcoming in *The British Journal for the Philosophy of Science*
(please cite official version when available)

Abstract

Around 97% of climate scientists endorse *anthropogenic global warming* (AGW), the theory that human activities are partly responsible for recent increases in global average temperatures. Clearly, this widespread endorsement of AGW is a reason for non-experts to believe in AGW. But what is the epistemic significance of the fact that some climate scientists do *not* endorse AGW? This paper contrasts expert *unanimity*, in which virtually no expert disagrees with some theory, with expert *consensus*, in which some non-negligible proportion either rejects or is uncertain about the theory. It is argued that, from a layperson's point of view, an expert consensus is often stronger evidence for a theory's truth than unanimity. Several lessons are drawn from this conclusion, e.g. concerning what laypeople should infer from expert pronouncements, how journalists should report on scientific theories, and how working scientists should communicate with the public.

Keywords: consensus, unanimity, dissent, expertise, science communication.

1 Introduction

According to recent studies, an overwhelming plurality of climate scientists endorse *anthropogenic global warming* (AGW), the theory that human activities are causally responsible for recent increases in global average temperatures. In a particularly well-known study (Cook et al., 2013), 97.1% of abstracts in climate science journals that took a position on AGW endorsed it (see also Cook et al., 2016). The fact that so many of the relevant experts endorse AGW is a strong reason for laypeople to believe in AGW, especially in light of the fact that climate scientists are a cognitively diverse community of researchers who disagree on a host of related issues (Oreskes, 2007; Odenbaugh, 2012; Dellsén, 2018). But what is the epistemic significance, from a layperson's point of view, of the fact that around 2.9% of climate scientists do *not* endorse AGW, by rejecting it (1.9%) or by being explicitly uncertain about it (1.0%)? Put differently, what should we make of the fact that 97.1% *rather than 100%* of climate scientists endorse AGW?

More generally, we may distinguish between expert *unanimity*, in which virtually none of the relevant experts disagrees with some theory, with expert *consensus*, in which some non-negligible proportion of the experts either rejects or is uncertain about the theory.[1] Which type of opinion distribution among scientific experts, unanimity or consensus, is stronger evidence for the relevant theory from the point of view of a layperson who isn't in a position to evaluate the theory for herself? This paper argues that, all else being equal, a mere consensus on a theory is often stronger evidence for the theory from a layperson's point of view than unanimity would be. Put differently, the presence of a marginal level of dissent among the relevant experts should often increase a layperson‘s confidence in the relevant theory. For example, the 97.1% consensus on AGW is plausibly stronger evidence for AGW, from a layperson's point of view, than a 100% agreement on

[1] Thus, what I am calling ‘consensus’ is closer to Miller's (2013, 1297) definition of the term than Tucker's (2003, 511); indeed, Tucker's ‘consensus’ corresponds to my ‘unanimity’.

the theory would be in otherwise identical circumstances.[2]

As I discuss in more detail below, this conclusion has far-reaching implications for the ways in which and expert testimony, e.g. from scientists, ought to be received by the public. Rather than being somewhat skeptical of consensus theories with which some small minority of relevant experts disagrees, laypeople should typically be more confident in such theories as compared to unanimously accepted theories. The argument below also implies that science journalists, and scientists themselves, ought to communicate scientific results in a different manner than they often do. For example, the journalistic practice of 'balanced reporting', where opposing viewpoints are given equal weight in media, effectively presents consensus theories (i.e., overwhelmingly accepted theories about which there is nevertheless some disagreement) as 'open questions' to be debated further, whereas unanimously accepted theories (i.e. theories about which there is no disagreement) are presented as 'settled fact'. If the argument of this paper is sound, it should often be the other way around.

This paper is organized as follows. Section 2 sets the stage by discussing and elaborating on two well-known facts about theory assessment in science, viz. that scientists are *fallible* in their evaluation of theories, and susceptible to *conformity* due to social pressure from other scientists. The next two sections are devoted to spelling out the argument for my central thesis. In section 3, I lay out an informal version of the argument, to convey an intuitive sense of the key idea and to foreshadow the official, formal version of the argument. In section 4, I then go on to spell out this formal version of the argument in a probabilistic, or Bayesian,

[2]This argument differs from my earlier argument (Dellsén, 2018) that the extent to which expert consensus on a theory provides laypeople with a reason to believe that it is correct may be increased by expert disagreement on other theories within the same domain of expertise. The argument of this paper, by contrast, is that some expert disagreement *on the consensus theory itself* increases the extent to which laypeople have reason to believe that the theory is correct. Thus, for example, the current paper aims to show that some disagreement on AGW gives laypeople a greater reason to believe in AGW, whereas my earlier paper (Dellsén, 2018) argued that disagreements on other theories in climate science give laypeople a greater reason to believe in AGW. (These arguments are of course compatible – indeed, complementary.)

framework. Finally, in section 5, I go on to draw various morals from this argument, concerning (i) how laypeople should assess theories like AGW in light of the 97% consensus among climate scientists, (ii) what journalistic norms ought to be adopted for science reporting on scientific theories, and (iii) how scientists themselves should communicate with the public.

2 Theory Assessment: From Naïvité to Nuance

Let me start with a preliminary point. This paper concerns the general question of which type of opinion distribution among experts, unanimity or consensus, provides laypeople with the most reason to believe a given claim. As my introductory remarks suggest, however, I will take the *scientific expert* as a paradigm case of the kind of expert with which I am concerned.[3] There are a several reasons for this. First, I am interested in the kind of agents that laypeople can identify relatively easily as experts within a given domain; I'll assume that this is true of scientific experts in so far as laypeople can access information about their academic qualifications, affiliation, research output, or other relatively transparent indicators. Second, the experts I am concerned with will be able to provide testimony to non-experts concerning the truth-value of theories, or their estimation thereof; and I take it that scientific experts are generally able to do this — perhaps in contrast to, e.g., expert craftspeople. Finally, it seems to me that if the argument below is convincing when applied to scientific experts it will be even more convincing when applied to non-scientific experts, roughly because the type of considerations to which the argument appeals — viz., that such experts are individually fallibility and susceptible to social conformity effects — will be even more plausible in the case of non-scientific experts.

So how are scientific claims — or 'theories', as I shall simply call them — in

[3] I do not have any specific account of expertise; rather, what I say should be consistent with any plausible account of what it is to be an expert in a given domain (e.g. Goldman, 2001; Scholz, 2009).

fact assessed, i.e. accepted or rejected, on the basis of scientific evidence?[4] It will be instructive for what follows to start by considering a very naïve model of how such theory assessment works in science. On this naïve model, each individual scientist evaluates scientific theories in isolation from other scientists, so that one scientist's assessment of the theory has no effect on another scientist's assessment; moreover, each scientist is perfectly reliable in their assessments, so that a correct theory is necessarily accepted and a false theory is necessarily rejected. If this model was accurate, correct theories would, once they been proposed and assessed, inevitably become unanimously accepted. Given this model, unanimity among scientists would be an excellent — indeed, impeccable — indicator of a correct theory. Laypeople would thus be well advised to view unanimity among scientists as the gold standard of reliable scientific testimony.

But of course we know full well that this model is at best a crude simplification. What happens when we add some nuance to this model such that it more closely approximates actual theory assessments in science? Consider first the fact that scientists do not evaluate theories in isolation from one another. Human beings have a natural tendency towards conforming with groups and communities to which they belong, by changing their behavior or beliefs to fit those of the group (see, e.g., Cialdini and Goldstein, 2004).[5] Such conformity effects are by far strongest when the majority is unanimous, and sharply reduced as soon as there

---

[4]Here and in what follows, 'rejecting' a theory should be understood to involve either accepting the negation of the theory or being explicitly uncertain about it. Of course, scientific theories may be assessed in more nuanced ways as well, e.g. by assigning to them a subjective probability or degree of belief. I have chosen to work with a binary (accept-or-reject) picture of theory assessment in this paper since the alternative would overly complicate the discussion without adding any apparent benefits, and since scientists tend to *communicate* their opinions on scientific theories to the public in binary ways, i.e. by endorsing the theory or not (rather than e.g. announcing that the theory should be assigned a particular probability).

[5]Classic studies of social conformity effects include Jenness 1932, Sherif 1935, Asch 1952, 1956, and Deutsch and Gerard 1955.

is any dissent against the perceived majority opinion.[6] It goes without saying that scientists are not immune to such effects, even if contemporary science contains various mechanisms that are designed to counteract them (Oreskes, 2019). Indeed, there are some examples in which the social pressure towards conformity within a scientific group or community has resulted in a 'bandwagon effect' (a.k.a. 'information cascade') where a theory is widely adopted despite there being only weak or sparse empirical evidence in its favor.[7] For example, in the early 20th century the renowned French physicist René Blondlot claimed to have discovered what he referred to as 'N-rays', a form of radiation analogous to X-rays. Due presumably to Blondlot's status as an authority in French science at the time, several other French physicists subsequently claimed that they had observed N-rays as well, making Blondlot's theory of N-rays widely accepted among French physicists in the early 20th century (Nye, 1980).

Another way in which the naïve model described above is deficient concerns its assumption that scientists are perfectly reliable. Scientists, even at the best of times, are sometimes mistaken — a truism often labelled 'fallibilism' (e.g. by Peirce, 1958; Popper, 1965). This is not to deny that scientists are much more likely to correctly evaluate a given scientific theory than the average person — after all, they have access to more evidence and are specifically trained to evaluate that evidence. However, assessing a scientific theory inevitably involves some amount of 'inductive risk', in part because scientific methods and instruments are never perfectly reliable and thus liable to produce misleading evidence (Frost-Arnold, 2019), and in part because of difficulties involved in interpreting scientific data, e.g. due to unconceived alternatives to extant theories (Sklar, 1981; Stanford, 2006). Indeed, there are two separate kinds of errors to which scientists are liable in theory assessments, viz. accepting an incorrect theory ('false posi-

[6]See Asch 1951, 1955, Allen and Levine 1968, Allen 1975, Bond and Smith 1996, and Bond 2005.

[7]For a couple of other examples, along with a game-theoretic explanation of the phenomenon, see Weatherall and O'Connor 2020.

tives') and rejecting a correct theory ('false negatives'). Although philosophical discussions of science tend to focus on the first type of error, the second kind of error — exemplified, for instance, by the rejection of continental drift in the early 20th century (Oreskes, 1999) — is at least as significant for our current purposes.

Of course, the extent to which scientists are fallible varies greatly depending on the nature of the theory and the state of scientific evidence regarding the theory. For example, scientists are certainly less prone to making errors about familiar phenomena about which there is overwhelming evidence (e.g., the Earth's shape; tuberculosis) than recently-posited entities or processes about which direct evidence is relatively scarce (e.g., dark energy; schizophrenia). These cases clearly form a continuum.[8] The argument of this paper, according to which consensus provides more support than unanimity, concerns scientific theories for which there is at least some reason to think scientists would be fallible — due either to scarcity of relevant evidence or the speculative nature of the theory itself (or both). Thus this arguably excludes cases such as, for example, the fact that the Earth is roughly spherical, and that tuberculosis is caused by bacteria. Such 'theories' do not fall within the scope of the argument below because there is no plausible reason for laypeople to suppose scientists would be fallible about such matters at this point.[9]

Now, having abandoned the naïve model of scientific theory assessments in which scientists are unaffected by social influences and necessarily infallible about

[8]For the purposes of this paper, I leave it open how to locate individual theories on this continuum, since that would effectively require a general account of what makes scientists more and less reliable in their theory evaluations relative to some evidence. To provide such a general account would take us too far afield from the main focus of the current paper.

[9]That said, there was of course a time at which it would have been reasonable for laypeople to suppose that scientists would be fallible on these matters. For example, there was at least some reason to suppose scientists would be fallible regarding the cause of tuberculosis before *Mycobacterium tuberculosis* could be isolated and examined through a microscope. What has changed between then and now is of course that evidence (e.g. microscopic observations) has accumulated to such an extent that it is no longer reasonable to suppose scientists are fallible to any significant degree regarding the cause of tuberculosis.

the relevant theories, it becomes less clear how laypeople should assess scientific theories on the basis of scientific testimony. Due to the combined effects of social conformity and fallibility, an incorrect theory may gain widespread acceptance (e.g. Blondlot's theory of N-rays); and, conversely, a correct theory may fail to be widely accepted (e.g. the theory of continental drift). So what distribution of opinion among experts is such that, all else being equal, laypeople should be *most* confident that the theory is true? For all I have argued so far, the answer to this question may still be *unanimity*, but this no longer follows trivially from our model of scientific theory assessment. Indeed, my task below is to argue that this answer is incorrect, and that a weaker form of plurality of opinion among experts — viz. what I am calling *consensus* — is, all else being equal, a more reliable indicator of truth than unanimity.

3 An Informal Foreshadowing of the Argument

When a layperson considers a given scientific theory, such as AGW, she will often have little or no direct empirical evidence to which she could reliably appeal in assessing the theory. By contrast, such a layperson will often, e.g. in the case of AGW, have information about the relevant experts' opinions about the theory. In this type of situation, the layperson lacks any evidence with which to assess the theory other than an *expert opinion distribution* about the theory. Clearly, the fact that a majority of experts accepts a theory is, in the absence of some special information about the unreliability or dishonesty of those experts, some evidence for the truth of theory in this situation. However, our current concern is with the *contrast* between two different expert opinion distributions both of which constitute a majority, viz. a (strict) unanimity and a (mere) consensus. Which expert opinion distribution, unanimity or consensus (both of which go beyond a simple majority), provides a *stronger* inductive evidence for T?

Plainly, there is no simple answer to this question that generalizes to absolutely all cases. As I discuss below, there are a number of factors that influence how strongly an expert opinion distribution on a theory supports that theory. These are brought out in the formal treatment of the issue (see section 4). What I hope to do in this section, however, is to motivate — informally, for now — the thought that there is a wide range of cases in which expert consensus provides a stronger argument for a theory than unanimity would do in an otherwise identical situation. Thus the informal argument of this section is not meant as anything like a definitive proof of the paper's main conclusion. Rather, the discussion of this section serves to foreshadow the formal argument provided in the following section, and to convey an intuitive sense of the argument that is to come.

For the purposes of this section, I will be operating with qualitative rather than quantitative definitions of 'unanimity' and 'consensus'. So I do not identify 'unanimity' with 99-100% agreement, for example, or 'consensus' with 90-98% agreement. Rather, I will take 'unanimity' to refer to an opinion distribution in which there is no or virtually no disagreement with the majority position, i.e. in which all or almost all of the relevant experts have the same opinion on a theory; and I will take 'consensus' to refer to a distribution in which there is some substantial disagreement with the majority position. This qualitative pair of definitions is appropriate in the informal version of the argument given in this section since informal reasoning of the type employed here does not enable us to make the type of quantitative comparisons in evidential support that would be necessary to determine whether a specific numerical expert opinion distribution provides greater or lesser evidence than another numerical expert opinion distribution.[10]

Now, consider (informally) what seems to be inferable from expert unanimity, on the one hand, and expert consensus, on the other other. In both cases, a majority of experts favor T, which arguably counts in favor of T in both cases. But since this is common ground between the cases we are comparing (i.e. between unanimity and consensus) let us set it aside here — e.g. by supposing that our

[10]If we want to make comparisons with that level of precision, we will need to do so in a quantitative framework for non-deductive reasoning such as the Bayesian framework employed in section 4.

layperson has already adjusted her opinion in T to reflect the that a majority of the relevant experts accept T. The question, then, is how much additional support is provided by the information that the relevant experts unanimously accept T, on the one hand, versus the information that there is a mere consensus on T, on the other hand?

To get at this issue, let us start by focusing not on what can be directly inferred from this about the truth or falsity of T, but rather what can be inferred about the process by which each (majority) expert opinion distribution emerged, i.e. what causally explains each distribution. In particular, consider the hypothesis that once a simple majority of the relevant experts come to accept a theory T, all or nearly all of the the other experts (i.e. those that had not yet formed an opinion, and those who had formed a contrary opinion, on T) deferred to the majority opinion:

> **Groupthink:** Once an expert majority arose on T, virtually all of the other experts formed opinions about T irrespective of the scientific evidence, by uncritically adopting the majority opinion on T.

Note that Groupthink does not just claim that some of the other experts relied on the majority opinion to some extent; rather, it asserts that, with respect to T, nearly all of those other experts completely relied on the majority in this way. For example, in a group of 100 experts in which a majority of 51 experts accept T, Groupthink implies that virtually all of the remaining 49 experts would defer to the majority opinion in favor of T. Note also that Groupthink does not directly address the issue of whether T is correct or incorrect — which is ultimately what our layperson is interested in. That said, its correctness is clearly relevant to the correctness of T itself, as we shall soon see.

For now, consider how a layperson should evaluate Groupthink in light of each of our two types of expert opinion distribution on T, viz. unanimity and consensus. Consider unanimity first. This opinion distribution would be plausibly explained by Groupthink since a group of experts in which the majority opinion is uncritically adopted by all or almost all of the other experts will naturally be led to unanimous opinion distributions, where all or almost all experts agree. Contrast this with a mere consensus, i.e. with a situation in which there is some dissent against an otherwise overwhelming majority on T (such as the 2.9% dissent against the consensus on AGW). This opinion distribution would not be as plausibly explained by Groupthink, since (to repeat) a group of experts in which the majority opinion is uncritically adopted by all or almost all of the other experts will naturally be led to unanimous opinion distributions, which excludes the type of dissent that is involved in a mere consensus. In sum, then, if Groupthink is correct with regard to a given theory T, we should expect expert unanimity, rather than consensus, on T. It follows that expert unanimity supports Groupthink more strongly than expert consensus.

This is in turn relevant to the extent to which a layperson can reasonably take a given expert opinion distribution as an indicator that T is correct. Roughly, this is because the extent to which experts have formed their opinions independently of each other contributes to their collective reliability in assessing T, in that agreement reached by a group of more independent thinkers is more likely to be correct than an otherwise identical agreement reached by more dependent thinkers.[11] Thus, all else being equal, a layperson has less reason to believe that T is itself correct given a majority expert opinion in its favor if they have reason to think Groupthink is the correct causal explanation of the expert opinion distribution on T. Add to this our previous conclusion that laypeople have more reason to believe that Groupthink is the correct explanation in cases of unanimity on T than in cases of mere consensus on T, and we get that there is a respect in which unanimity provides less reason to believe T than consensus. Put differently, there is an epistemic advantage to expert consensus over unanimity, in that the latter

[11] See, for instance, Goldman 2001 and Dellsén 2020. This is related to the well-known upshot of the Condorcet Jury Theorem that under certain conditions a group of independent and individually competent agents is more likely to be correct in a majority vote than any of the group's proper subsets (see, e.g., List and Goodin, 2001).

but not the former supports Groupthink, which in turn undermines T.

This concludes the informal foreshadowing of the argument that consensus often provides laypeople with more support than unanimity. Although I hope it conveys an intuitive sense of how consensus could be superior to unanimity in an important respect, the argument is clearly incomplete as it stands. After all, even if there is *a respect* in which unanimity provides less reason than consensus to believe T — i.e., even if there is *an* epistemic advantage to consensus over unanimity — there might still be *other respects* in which unanimity provides a greater reason to believe T — i.e., there might still be *other* epistemic advantages to unanimity over consensus. In particular, the fact that a greater number of experts accept T in cases of unanimity is plausibly itself such an epistemic advantage to unanimity over consensus. So why not think that the epistemic advantage to consensus over unanimity is always and everywhere outweighed by an opposite epistemic advantage to unanimity over consensus?

Since this question concerns whether one kind of epistemic advantage is stronger than another, answering it requires us to appeal to a framework for inductive reasoning that allows for quantitative comparisons. That is, we must situate the argument within a framework that allows us to determine, not just whether a given piece of evidence provides a reason to believe something, but how strong that reason is as compared to some other reason. The obvious candidate for such a framework is Bayesian Confirmation Theory, so let us now turn to formulating the informal argument outlined in this section within that quantitative framework.

[12]For the purposes of this paper, there is no need to further assume that rational agents are required to to update their credences in light of new evidence in accordance with *Bayesian Conditionalization*, although such a diachronic requirement is often added.

4 A Formal Version of the Argument

For our purposes, the core claim of *Bayesian Confirmation Theory* (BCT) is that that the opinions of perfectly rational agents are, or can be represented as being, *subjective probabilities*, i.e. degrees of beliefs that satisfy the axioms of probability.[12] To this BCT adds an explication of the concept of 'confirmation' (or 'support'), according to which E confirms (supports) H just in case conditionalizing on E raises the probability of H, i.e. $P(H|E) > P(H)$. Additionally, since our concern is with the comparative support conferred on a theory from two different pieces of evidence that a layperson could come to possess, viz. expert unanimity and expert consensus respectively, we need a general, quantitative measure of confirmation.[13] For reasons given by Fitelson (2001, 40-48), I favor the *log-likelihood ratio measure* of quantitative support:[14]

$$l(H,E) := \log\left[\frac{P(E|H)}{P(E|\neg H)}\right]$$

Clearly, this relativizes confirmation, and degrees of confirmation, to a given probability function $P(\cdot)$, corresponding to the agent's probability distributions. This highlights a notorious limitation of BCT, which is that it is unclear how to dis-

---

[13] A number of ordinally non-equivalent such measures have been proposed in the literature — see Fitelson (1998) for an overview of these measures. This is not the place to argue in favor of any particular such measure, which would take us too far afield.

[14] The log-likelihood ratio measure is ordinally equivalent to the simpler *likelihood measure*:

$$l'(H,E) := \frac{P(E|H)}{P(E|\neg H)}$$

However, the log-likelihood ratio measure l has various technical and aesthetic virtues over its simplified cousin l'. (For example, $l'(H,E) = 0$ when $P(H|E) = P(H)$, i.e. when E provides no confirmation to H.) Since the two measures are ordinally equivalent, any comparative claim of the sort I will make below holds for l if and only if it holds for l'.

tinguish intuitively acceptable from unacceptable probability functions.[15] Rather than trying to solve that problem here, I will seek to identify seemingly reasonable conditions on probability functions for which the central claim of this paper — that expert consensus supports theories more than unanimity, from a layperson's point of view — is validated in BCT.

As in the informal version of the argument discussed above, the key argumentative move is to consider how unanimity and consensus support different explanations of how the experts came to have the opinion distributions they in fact have. So let G correspond to the explanatory hypothesis of GROUPTHINK, i.e. that once a majority of the experts came to accept T, virtually all of the other experts formed opinions about T by uncritically adopting the expert majority opinion on T. The denial of this claim, $\neg G$, holds that the other experts evaluated T at least somewhat independently of the majority opinion. In what follows, we take it as part of the layperson's corpus of background evidence, common to any evidential situation considered below (and therefore suppressed in the probabilities considered below), that *some* majority of experts accept T; what differs between these situations is *how large* a majority accepts T.[16] Given this background knowledge, G in effect implies that any additional endorsement of T among the relevant experts is worthless as a guide to whether T is true, since these remaining experts simply echo the majority (whose opinions have already been taken into account). Formally, where E_x is the claim that some unspecified proportion of the relevant experts, beyond a simple majority, accept T, G 'screens off' the support for T that

[15]Indeed, for this reason, some Bayesians (e.g., Howson, 2000; Strevens, 2004) hold that BCT is merely a bare-bones framework for non-deductive reasoning that will need to be supplemented with substantive assumptions about acceptable probability functions if it is to deliver any interesting results about rational non-deductive reasoning.

[16]That is, if M signifies the proposition that some majority of the relevant experts accept T, then any unconditional probability $P(X)$ discussed below should implicitly be understood to equal the conditional probability $P(X|M)$, and any conditional probability $P(X|Y)$ should implicitly be understood to equal $P(X|Y\&M)$.

might otherwise be provided by the relevant expert opinion distribution:

$$P(T|E_x \& G) = P(T|G) \qquad (*)$$

In most cases, it will be reasonable for laypeople to assign a very low probability to G, so the possibility that the experts have engaged in this type of groupthink will play a relatively minor role in a layperson's epistemic life. However, the expert opinion distribution itself can effect the extent to which G is a live possibility. Indeed, this is the key idea behind the Bayesian version of the argument: Under certain natural assumptions about layperson's probability assignments, an expert unanimity on T raises the probability of G to a greater extent than does a mere expert consensus on T; and since G screens off any opinion distribution regarding T, this in turn makes expert unanimity a less potent probability-raiser of T as compared to expert consensus. To spell out this key idea, I will use the log-likelihood measure of confirmation to compare the degree of support for T provided by one opinion distribution, E_u, as compared to another, E_c. These opinion distributions may intuitively be thought of as corresponding to 'unanimity' and 'consensus' respectively (and I will often refer to them as such in what follows).[17] To be clear, my claim is not that E_c (consensus) invariably provides more support than E_u (unanimity), i.e. that $l(T,E_c) > l(T,E_u)$ holds for any probability function $P(\cdot)$; rather, it is the much more modest claim that this holds under conditions on such probability functions that can be independently motivated as reasonable for laypeople to have in many instances.

To convey an intuitive sense of how to understand the results that follow, I will start with a simple toy example of maximally specific set of conditions in which

[17] With that said, the formal results below apply regardless of what specific levels of agreement E_u and E_c are taken to refer to in a given case. There is therefore no need to define 'unanimity' and 'consensus' for the purposes of the current section.

$l(T, E_c) > l(T, E_u)$.[18] Suppose a layperson is interested in a theory T on which there are, say, 100 experts in total. Our layperson knows already that there is a majority of some sort in favor of T, i.e. at least a simple majority of 51 experts accept T. Moreover, our layperson recognizes that it is epistemically possible for the remaining experts to accept a theory because of an extreme form of groupthink, i.e. because virtually all of the experts follow the lead of the majority rather than critically evaluating the evidence. Thus the layperson assigns a non-zero probability to G, e.g. 0.01. And since groupthink will naturally tend to lead to more extreme opinion distributions within the group, the agent rationally assigns a higher probability to more extreme opinion distributions conditional on G. For specificity, let us assume that our layperson assigns likelihoods for 80 to 100 agents accepting T, conditional on G, as shown in Figure 1. Furthermore, our layperson also acknowledges that experts are fallible, even in cases where their opinions are not explained by groupthink. Thus, conditional on $\neg G$, it is unlikely from the layperson's point of view that *all* of the experts will accept T given that T is true — although it is even less likely that they all accept T given that T is false. Again for specificity, suppose the relevant likelihoods are as shown in Figure 1.[19]

Given these assumptions about our layperson's probability assignments, one can calculate[20] the degree to which various expert opinion distributions would confirm the theory T by our layperson's lights — see Figure 2. Notably, the degree of confirmation conferred on T is not simply proportional to, or even monoton-

[18]Let me emphasize at the outset that the results that I appeal to in what follows are not derived from assumptions about this toy example, so my argument below does not rely on it in any substantial way. Rather, I include a discussion of this toy example only for illustrative purposes, to convey an intuitive sense of how the argument of this section works.

[19]In this example, our layperson expects the expert opinions, in the absence of groupthink, to form binomial distribution with modes in 90% and 10% of remaining expert acceptances (beyond a simple majority) depending on whether T is true or false. These distributions results from assuming that (i) each remaining expert evaluates the theory completely independently of the majority, (ii) each such expert is 90% likely to accept T if T is true (and G is false), and (iii) each such expert is 10% likely to accept T if T is false (and G is false).

[20]E.g., by using (†) below.

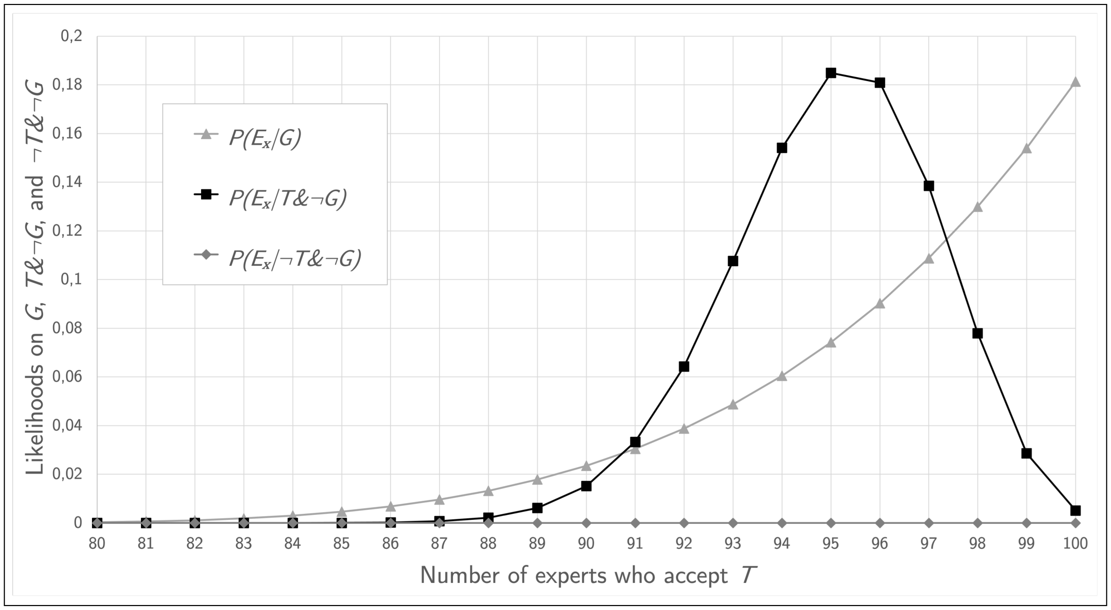

Figure 1: Likelihoods of different expert opinion distributions in a toy example.

ically increasing with, the number of experts who accept T. Rather, T is most confirmed by expert opinion distributions that can be described as a mixture of consensus with some substantial dissent (with 95 experts constituting the consensus on T, and 5 dissenters). This degree of confirmation is noticeably lower — although still clearly positive — for more uniform opinion distributions, such as complete agreement (where all 100 experts accept T). Indeed, for complete agreement, the degree of confirmation conferred on T is comparable to that provided by a comparatively modest majority of 86 out of 100 experts accepting T. So, in this toy example, opinion distributions that would intuitively be thought of as unanimous are far from being the ideal expert opinion distributions as far as confirmation of T is concerned from the laypersons's point of view. Put differently, the existence of dissent against a consensus position *increases* the extent to which T is confirmed by the expert opinion distribution in the example.

This example demonstrates that it is *possible* for expert consensus to provide more support than unanimity. But of course it tells us nothing, by itself, about

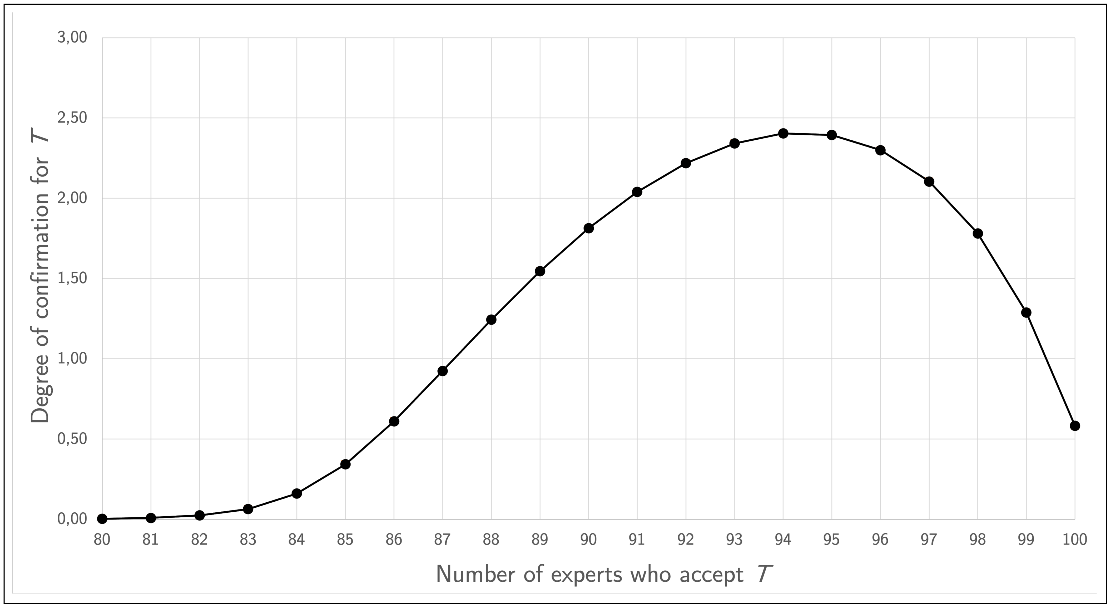

Figure 2: Degrees of confirmation for T given the likelihoods in Figure 1.

the general conditions under which this holds. To get at this more general issue, note first that given our previous assumption (*), the quantitative degree of confirmation for T provided by a given expert opinion distribution E_x can be written as a function of the probabilities and likelihoods of G and $\neg G$ as follows:

$$l(T, E_x) = \log\left[\frac{P(E_x|G)P(G) + P(E_x|T\&\neg G)(1 - P(G))}{P(E_x|G)P(G) + P(E_x|\neg T\&\neg G)(1 - P(G))}\right] \qquad (\dagger)$$

This shows that the degree to which a given expert opinion distribution confirms T depends on no less than *four* distinct probabilities, viz. $P(G)$, $P(E_x|G)$, $P(E_x|T\&\neg G)$, and $P(E_x|\neg T\&\neg G)$, the last three of which may differ between consensus and unanimity (i.e., they can take on different values when E_x is specified to be E_c or E_u). Given that $l(T, E_x)$ has so many 'moving parts', what (if anything) can be said about the general conditions under which under which expert consensus provides stronger support than unanimity, i.e. $l(T, E_c) > l(T, E_u)$?

One way to answer this question is to see what happens when we let some of these terms vary while holding others fixed. The fixed terms thus play the role of *ceteris paribus* conditions for theorems connecting the varying terms with degrees of confirmation of T given consensus and unanimity respectively.

In particular, it is not hard to see that if other terms in (†) are held constant between E_c (consensus) and E_u (unanimity), the former provides more support for T than the latter if and only if the likelihood of E_c on T and $\neg G$ is greater than the corresponding likelihood of E_u (all proofs in the Appendix):

Theorem 1. *Let $P(E_c|G) = P(E_u|G)$ and $P(E_c|\neg T\&\neg G) = P(E_u|\neg T\&\neg G)$. Then:*

$$l(T,E_c) > l(T,E_u) \iff P(E_c|T\&\neg G) > P(E_u|T\&\neg G)$$

Roughly speaking, what this tells us is that, for agents that find it likelier that non-groupthinking experts evaluating a true theory would reach a consensus on the theory than unanimity, consensus provides them with more support for T than unanimity, all else being equal. This partly explains[21] why, in our toy example, 95 experts accepting T provides so much more confirmation for T than 100 experts accepting T, since our layperson in that example assigned a higher likelihood, on T and $\neg G$, to 95 experts accepting T than to 100 experts doing so.

Interestingly, a similar theorem can be proved with regard to the likelihood of consensus versus unanimity arising among groups of experts that engage in groupthink. As noted, if the other experts in a group defer to an already-established majority opinion regarding T among their peers, this would lead to more extreme opinion distributions regarding T, so that the experts would be more inclined to unanimously accept T. It is thus of interest that if other terms in (†) are held constant between consensus and unanimity, and if the likelihood of consensus and unanimity on T is greater if T is true than if T is false (as seems eminently rea-

[21] Another part of the explanation is provided by the next theorem; these two explanations are then neatly unified in the more general explanation provided by the third theorem below.

sonable), then consensus provides more support for T than unanimity if and only if the likelihood of unanimity on G is greater than the corresponding likelihood of consensus on G:

Theorem 2. *Let* $P(E_c|T\&\neg G) = P(E_u|T\&\neg G) > P(E_c|\neg T\&\neg G) = P(E_u|\neg T\&\neg G)$. *Then:*

$$l(T, E_c) > l(T, E_u) \iff P(E_c|G) < P(E_u|G)$$

Roughly speaking, what this tells us is that, for agents that find it likelier that unanimity would occur among groupthinking experts (and who find it likelier that either kind of agreement occurs when the theory is in fact true), consensus again provides them with more support that unanimity, all else being equal. This also partly explains why, in our toy example, 95 experts accepting T provides so much more confirmation for T than 100 experts accepting T, since our layperson in that example assigned a higher likelihood, on G, to 100 experts accepting T than to 95 experts doing so.

The two previous theorems assumed as *ceteris paribus* conditions that some terms were held fixed between consensus and unanimity, while other were allowed to vary. A final theorem shows how the terms that were allowed to vary interact with one another.

For this theorem, no *ceteris paribus* assumptions are made; instead, an idealization is required to simplify (†). To motivate this idealization, notice that experts *who do not engage in groupthink* will be extremely unlikely to reach either consensus or unanimity *on false theories*. This is emphatically not to say that experts cannot reach consensus or unanimity on false theories; rather, it is to say that *if* (or *when*) this happens, it can reasonably be assumed to be because the relevant experts have engaged in some form of groupthink. In effect, this means

[22]For instance, in our toy example of experts that were assumed to be 90% reliable and independent with regard to T (see footnote 19), these likelihoods are between 10^{-40} and 10^{-50}, and thus several orders of magnitude smaller than other terms in (†) as applied to E_c and E_u.

that the likelihood of consensus on a false theory, given lack of groupthink, is negligible; and likewise for unanimity. Hence, the idealizing assumption that $P(E_c|\neg T\&\neg G) = P(E_u|\neg T\&\neg G) = 0$ can be assumed to hold as a reasonable approximation in most circumstances.[22] This gets us a lovely theorem that combines the insights gleaned from our previous two theorems:

Theorem 3. *Let* $P(E_c|\neg T\&\neg G) = P(E_u|\neg T\&\neg G) = 0$*. Then:*

$$l(T,E_c) > l(T,E_u) \iff \frac{P(E_c|T\&\neg G)}{P(E_u|T\&\neg G)} > \frac{P(E_c|G)}{P(E_u|G)}$$

Roughly speaking, this tells us is that, for agents that find it negligibly likely that non-groupthinking agents will reach either kind of agreement on a false theory, whether consensus confirms more than unanimity depends (i) *positively* on the extent to which they find it likelier that consensus would occur for true theories among non-groupthinking experts, and (ii) *negatively* on the extent to which they find it likelier that unanimity would occur among groupthinking experts. In our toy example, *both* of these factors were 'favorable' to making $l(T,E_c) > l(T,E_u)$, in that $P(E_c|T\&\neg G) > P(E_u|T\&\neg G)$ (making the first ratio on the right hand side larger than 1) while $P(E_c|G) < P(E_u|G)$ (making the second ratio smaller than 1), which is why consensus is so decisively more confirmatory than unanimity in that case.

Importantly, however, Theorem 3 also reveals that even when one of these factors is 'unfavorable' to $l(T,E_c) > l(T,E_u)$, e.g. when $P(E_c|G) > P(E_u|G)$, this can be outweighed by the other factor being sufficiently 'favorable'; in that case, the ratio of $P(E_c|T\&\neg G)$ to $P(E_u|T\&\neg G)$ simply needs to be higher than the ratio of $P(E_c|G)$ to $P(E_u|G)$. Thus, although *at least one of* these factors needs to be 'favorable' (in the above sense) in order for $l(T,E_c) > l(T,E_u)$ to hold, it is not necessary that they be both are — and, indeed, one of the factors can be 'unfavorable' as long as the other is sufficiently 'favorable' to outweigh that. To illustrate, consider a modification of our toy example in which our layperson assigns higher likelihoods to consensus opinion distributions on groupthink than to unanimity (see Figure 3). Keeping other features of the example as before, a consensus of around 95 experts still provides more confirmation for T than a strict unanimity of 100 experts (see Figure 4). Theorem 3 explains why, since in this modified toy example, we can see from the steeper downward slope at the right tail of $P(E_x|T\&\neg G)$ as compared to right tail downward slope of $P(E_x|G)$ that the ratio of $P(E_c|T\&\neg G)$ to $P(E_u|T\&\neg G)$ is greater than that of $P(E_c|G)$ to $P(E_u|G)$.

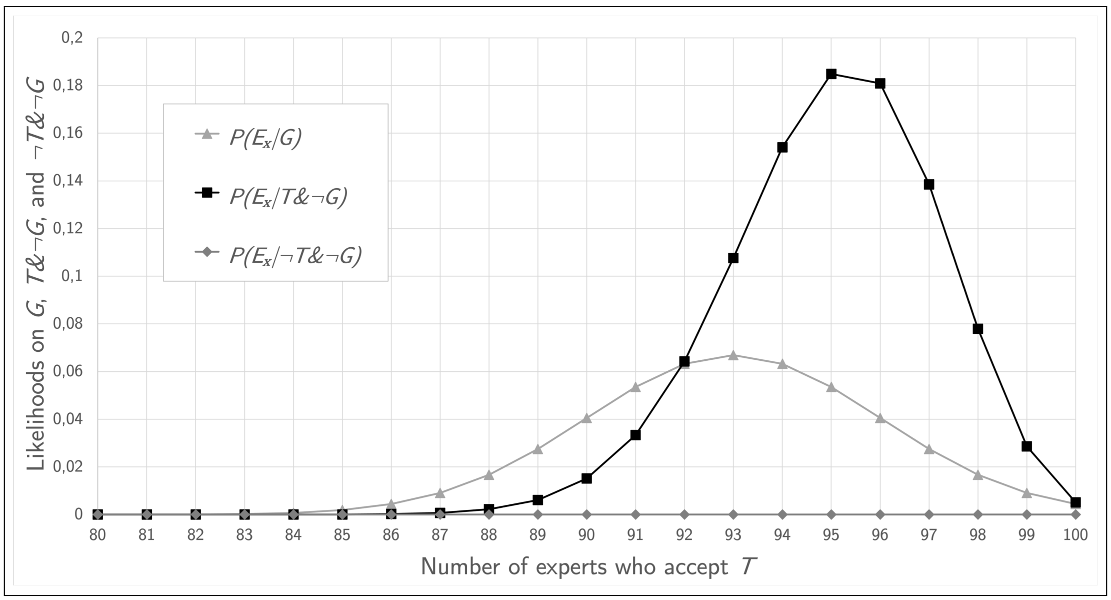

Figure 3: Likelihoods of expert opinion distributions in a modified toy example.

Let me conclude this section by briefly considering two objections. First, against the 'Bayesian' approach taken in this section, one might object that each of the above theorems requires highly unrealistic assumptions, e.g. that our layperson assigns zero probability to $P(E_c|\neg T\&\neg G)$ and $P(E_u|\neg T\&\neg G)$ as Theorem 3 requires. Thus, the objection goes, these theorems have no relevance for the question of what actual, real-life laypeople should conclude from expert consensus versus unanimity

In response, while I acknowledge that some of the assumptions made here are

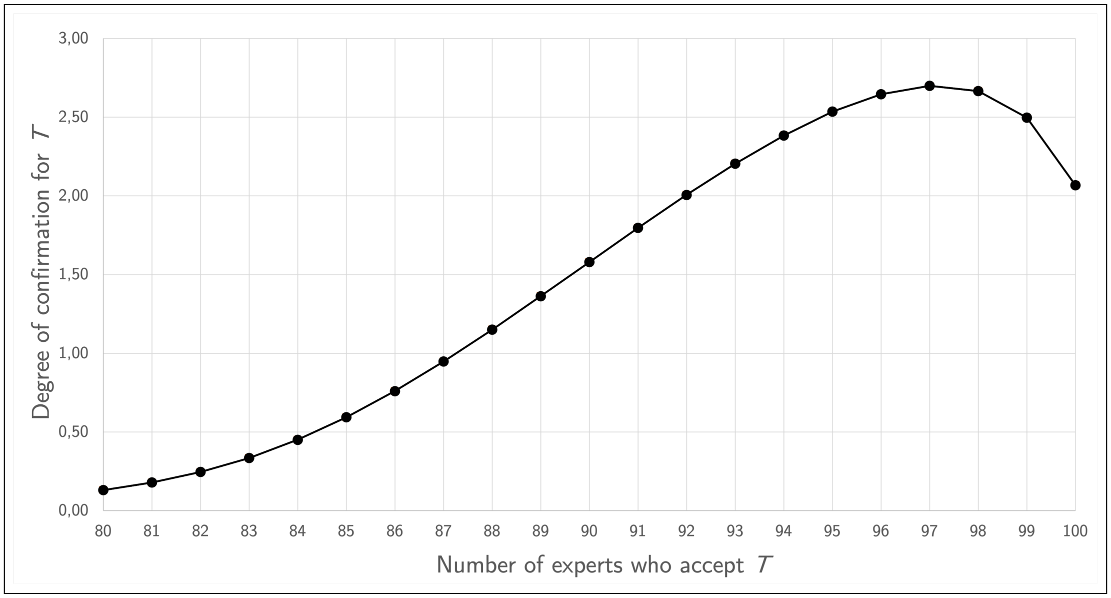

Figure 4: Degrees of confirmation for T given the likelihoods in Figure 3.

unrealistic, it does not follow that these theorems have no relevance for understanding the epistemic situation of real-life agents. These theorems reveal what sort of factors could contribute to making expert consensus provide a layperson with more confirmation for a theory than unanimity. In other words, they reveal what makes it true — when it is true — that $l(T, E_c) > l(T, E_u)$. This holds for real life agents as well as for artificial, idealized agents. *Strictly speaking*, the full answer is that $l(T, E_c) > l(T, E_u)$ is made true by *all* of the factors that occur on in (†), but one way to bring out *the way in which* it depends on each one is to hold other factors constant (as in Theorems 1 and 2) or set negligibly small factors to zero (as in Theorem 3). This is a familiar and useful strategy in science itself, e.g. in causal modeling,[23] so I make no apologies for using the same strategy to bring out how and when, from a layperson's point of view, expert consensus confirms theories more strongly than unanimity.

A second possible objection is that this argument assumes that the laypeople

[23]See, e.g., Weisberg, 2007, 2013, and Strevens, 2008, 2017.

in question do not have any information relevant to assessing T other than the expert opinion distributions regarding T. Strictly speaking, this is of course rarely, if indeed ever, the case. In the case of AGW, for example, most of us know a little about how global mean temperatures have risen since the middle of the 19th century. Thus, the objection goes, this argument applies only in artificial circumstances that are unlikely to arise in the real world.

In response, I submit that while laypeople rarely have *no* additional information with which to assess a theory, they — or rather, *we* — are quite often in evidential situations in which expert opinion is much weightier than any other more direct forms of evidence to which we might have access. Moreover, the direct evidence that one does have access to in such situations will also be available to the relevant experts, who will be better placed to assess its relevance to T, in which case expert opinion on T should *replace* our own assessment of T (Elga, 2007). Indeed, it is only in the very unusual case in which I have strong reason to believe that the experts have somehow missed an important piece of evidence that the argument given above would lose much of its force. This is certainly not true of the consensus on AGW, for instance, where the direct evidence that I have as a layperson has already been taken into account by the experts whose opinions I am consulting.

5 Conclusion and Implications

I have argued that, in a wide range of cases, laypeople have more reason to believe a scientific theory when there is some dissent against a consensus position than when all or virtually all or the relevant experts take the same position. As I have noted along the way, there are arguably exceptions to this rule — e.g., if the 'theory' in question is not something on which laypeople should expect experts to

[24] As the scare quotes indicate, however, it is something of a stretch to call the relevant claims 'theories' and the relevant agents 'laypeople'. So these cases may not constitute genuine exceptions after all.

ever be mistaken about, or if the 'laypeople' have relevant information or insights that the experts lack.[24] Generally speaking, however, a certain marginal level of dissent among experts should increase a layperson's confidence in a theory about which the experts have otherwise reached a consensus, roughly because this indicates that the agreement on the theory in question is less likely to be due to a conformity effect, i.e. 'groupthink'. As we have seen, this basic thought can be fleshed out in both abductive and Bayesian frameworks for non-deductive reasoning.

What are the practical implications of this argument? One obvious implication concerns how laypeople should form opinions about scientific theories, or other issues on they have access to expert opinion. Specifically, the argument suggests that laypeople should be most confident that theories are true when the relevant experts display a mere consensus with some dissent. For example, it is plausible that the 97% agreement among climate scientists on anthropogenic global warming (AGW) provides laypeople with a stronger reason to believe in AGW than what would be provided with a 100% agreement. Relatedly, this shows how misleading it can be to pick out and specifically listen to the opinions of those who disagree with a consensus position, e.g. the 3% of climate scientists who claim to be skeptical of AGW. The fact that *some* such experts can be found should, perhaps counterintuitively, be seen as a reason to *strengthen* our conviction in the consensus theory — provided that there are enough experts on the other side of the issue to constitute a consensus. For example, the fact that climate deniers have been able to find and promote individual scientists who disagree with the consensus on AGW — a common and effective propaganda strategy which forms part of the 'tobacco strategy' (Oreskes and Conway, 2010; Weatherall et al., 2019) — should not in the least be taken to undermine the credibility of the theory; on the contrary, it should make AGW all the more credible by laypeople's lights.

Interestingly, laypeople are not the only group of people for whom the preceding argument makes a practical difference. Another such group is science reporters. In recent years, sparked by controversial reporting on issues like climate change, journalism scholars have debated how much 'weight', e.g. column inches or on-air minutes, journalist should give to opposing viewpoints on scientific matters (e.g., Boykoff and Boykoff, 2004; Dunwoody, 2005; Dunwoody and Kohl, 2017).[25] Three types of proposals have been most extensively discussed in the literature (not necessarily under these labels):

Balanced Reporting: Give equal weight to all scientific viewpoints held by any relevant expert.

Weight-of-Evidence Reporting: Give weight to scientific viewpoints in proportion to the evidence for each.

Weight-of-Experts Reporting: Give weight to scientific viewpoints in proportion to the opinion distribution among relevant experts.

Now, we may ask, how do these journalistic norms stack up in light of the preceding argument that a consensus typically provides more support than unanimity? Which of these norms ought journalists follow in so far as they want to give their audience the most accurate impression of the epistemic status of a given theory?

Balanced Reporting is widely criticized for effectively distorting the public's perception of the epistemic status of consensus theories, since it strictly speaking implies that a single dissenter's viewpoint should be given equal weight as that of an overwhelming consensus (Gelbspan, 1998; Boykoff and Boykoff, 2004; Figdor, 2018). To this criticism we can now add the further objection that, if this paper's argument is sound, the fact that there are *some* experts on the other side of an issue does not *to any extent* suggest that the consensus is on the wrong track; on the contrary, provided there is nonetheless a sufficiently large consensus, *it suggests the opposite*. Thus, Balanced Reporting is even more misleading than previous arguments against it have suggested. For similar reasons, Weight-of-Experts Reporting will arguably also be very misleading, in so far as it would give the dissenters more weight in the media than the existence of dissenters itself warrants

[25]See also recent philosophical discussions of the issue in Figdor, 2018 and Gerken, 2020.

epistemically. For example, if a media outlet used even just 3% of its air time or column inches channeling climate skepticism, this would still present a skewed picture of the evidential situation regarding AGW in so far as unanimously accepted theories are not similarly portrayed in a skeptical light at all. Thus, the most appropriate of the three norms might well be a form of Weight-of-Evidence Reporting, especially if 'evidence' is interpreted broadly so as to include the sorts of considerations with which the argument of this paper is concerned.

Finally, the argument of this paper has implications for those whose opinions are being consulted by laypeople or the media, i.e. the experts themselves. These include scientists, but also more broadly researchers and academics who serve as epistemic authorities on various matters. In the case of scientists, there are good reasons to think that it is often normatively appropriate for working scientists to dissent in various ways, i.e. to disagree with the consensus position in science and criticize that position (Intemann and de Melo-Martín, 2018). However, in light of this paper's argument, it can also be normatively *in*appropriate for scientists to dissent with consensus position *in their roles as experts*, i.e. when communicating with the public or with policy makers, because this can give a misleading impression of the state of scientific research in their fields (on this, see also Biddle and Leuschner, 1015; Biddle et al., 2017; Leuschner, 2018). This suggests that the ideal way for scientists to interact with laypeople involves refraining from communicating their individual opinions on the epistemic status of a given theory, and instead communicate the consensus position among researchers in their field (as in, e.g., Joint Statement, 2001; American Association for the Advancement of Science, 2006); or, more generally, report on the opinion distribution of scientists working in their field (as in, e.g., Gust et al., 2008; Cook et al., 2013, 2016; Pew Research Center, 2015).[26]

[26]This paper was improved by very helpful written feedback from Dunja Šešelja and three anonymous reviewers for *The British Journal for the Philosophy of Science*, as well as verbal feedback from audiences at the Nils Klim Symposium at the University of Iceland and the philosophy research seminar at the University of York.

Appendix: Proofs of Theorems

Proof of Theorem 1

Proof. As noted in footnote 14, log-likelihood ratio measure is ordinally equivalent to the simpler measure likelihood ratio measure. Thus we have that, by (†), $l(T, E_c) > l(T, E_u)$ if and only if:

$$\frac{P(E_c|G)P(G) + P(E_c|T\&\neg G)(1-P(G))}{P(E_c|G)P(G) + P(E_c|\neg T\&\neg G)(1-P(G))} > \frac{P(E_u|G)P(G) + P(E_u|T\&\neg G)(1-P(G))}{P(E_u|G)P(G) + P(E_u|\neg T\&\neg G)(1-P(G))} \tag{1}$$

The theorem assumes that $P(E_c|G) = P(E_u|G)$ and $P(E_c|\neg T\&\neg G) = P(E_u|\neg T\&\neg G)$, so we may divide through by their denominators (which are equal); then subtract both sides by $P(E_c|G)P(G)$ (or, equivalently, by $P(E_x|G)P(G)$); and finally divide through by $(1-P(G))$; obtaining $P(E_c|T\&\neg G) > P(E_u|T\&\neg G)$, as desired. □

Proof of Theorem 2

Proof. To simplify calculations in what follows, let $\alpha = P(E_c|T\&\neg G)(1-P(G)) = P(E_u|T\&\neg G)(1-P(G))$ and $\beta = P(E_c|\neg T\&\neg G)(1-P(G)) = P(E_u|\neg T\&\neg G)(1-P(G))$. We can now write (1) as follows:

$$\frac{P(E_c|G)P(G) + \alpha}{P(E_c|G)P(G) + \beta} > \frac{P(E_u|G)P(G) + \alpha}{P(E_u|G)P(G) + \beta} \tag{2}$$

Since both denominators are positive, multiplying both sides with their product gets us:

$$P(E_c|G)P(E_u|G)P(G)^2 + P(E_u|G)P(G)\alpha + P(E_c|G)P(G)\beta + \alpha\beta > P(E_c|G)P(E_u|G)P(G)^2 + P(E_c|G)P(G)\alpha + P(E_u|G)P(G)\beta + \alpha\beta \tag{3}$$

Algebraic manipulations of this inequality gets us:

$$P(E_u|G)(\alpha - \beta) > P(E_c|G)(\alpha - \beta) \tag{4}$$

Now, note that since $1 - P(G) > 0$, it follows from the theorem's assumption that $\alpha > \beta$. Thus (4) is preserved when both sides are divided by $(\alpha - \beta)$, giving us $P(E_u|G) > P(E_c|G)$, as desired. □

Proof of Theorem 3

Proof. Our assumption that $P(E_c|\neg T\&\neg G) = P(E_u|\neg T\&\neg G) = 0$ allows us to write (1) as follows:

$$\begin{aligned} &\frac{P(E_c|G)P(G) + P(E_c|T\&\neg G)(1 - P(G))}{P(E_c|G)P(G)} \\ &> \frac{P(E_u|G)P(G) + P(E_u|T\&\neg G)(1 - P(G))}{P(E_u|G)P(G)} \end{aligned} \tag{5}$$

Since both denominators are positive, multiplying both sides by the product of their denominators now gets us:

$$\begin{aligned} &P(E_c|G)P(E_u|G)P(G)^2 + P(E_c|T\&\neg G)(1 - P(G))P(E_u|G)P(G) \\ &> P(E_c|G)P(E_u|G)P(G)^2 + P(E_u|T\&\neg G)(1 - P(G))P(E_c|G)P(G) \end{aligned} \tag{6}$$

Algebraic manipulations (in which we exploit that $0 < P(G) < 1$), then gets us:

$$\frac{P(E_c|T\&\neg G)}{P(E_u|T\&\neg G)} > \frac{P(E_c|G)}{P(E_u|G)} \tag{7}$$

as desired. □

References

Allen, V. L. (1975). Social support for nonconformity. In *Advances in experimental social psychology (Vol. 8)*. Academic Press, New York, NY.

Allen, V. L. and Levine, J. M. (1968). Social support, dissent and conformity. *Sociometry*, 31:138–149.

American Association for the Advancement of Science (2006). Statement on the teaching of evolution. Available at `https://www.aaas.org/sites/default/files/0219boardstatement.pdf`.

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgment. In Guetzkow, H., editor, *Groups, Leadership and Men*, pages 295–303. Carnegie Press, Pittsburgh.

Asch, S. E. (1952). *Social Psychology*. Oxford University Press, Oxford.

Asch, S. E. (1955). Opinions and social pressure. *Scientific American*, 193:31–35.

Asch, S. E. (1956). Studies of independence and conformity: A minority of one against aunanimous majority. *Psychological Monographs*, 70:1–70.

Biddle, J. B. and Leuschner, A. (21015). Climate skepticism and the manufactur of doubt: Can dissent in science be epistemically detrimental? *European Journal for Philosophy of Science*, 5:261–278.

Biddle, J. B., Leuschner, A., and Kidd, I. J. (2017). Epistemic corruption and manufactured doubt: The case of climate science. *Public Affairs Quarterly*, 31:165–187.

Bond, R. (2005). Group Size and Conformity. *Group Porcesses & Intergroup Relations*, 8:331–354.

Bond, R. and Smith, P. B. (1996). Culture and Conformity: A Meta-Analysis of Studies Using Asch's (1952b, 1956) Line Judgment Task. *Psychological Bulletin*, 119:111–137.

Boykoff, M. T. and Boykoff, J. M. (2004). Balance as bias: Global warming and the US prestige press. *Global Environmental Change*, 14:125–136.

Cialdini, R. and Goldstein, N. (2004). Social Influence: Compliance and Conformity. *Annual Review of Psychology*, 55:591–621.

Cook, J., Nuccitelli, D., Green, S., Richardson, M., Winkler, B., Painting, R., Way, R., Jacobs, P., and Skuce, A. (2013). Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environmental Research Letters*, 8:024024.

Cook, J., Oreskes, N., Doran, P., Anderegg, W., Verheggen, B., Maibach, E., Carlton, J., Lewandowsky, S., Skuce, A., and Green, S. (2016). Consensus on consensus: A synthesis of consensus estimates on human-caused global warming. *Environmental Research Letters*, 11:048002.

Dellsén, F. (2018). When Expert Disagreement Supports the Consensus. *Australasian Journal of Philosophy*, 96:142–156.

Dellsén, F. (2020). The Epistemic Value of Expert Autonomy. *Philosophy and Phenomenological Research*, 100:344–361.

Deutsch, M. and Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The journal of abnormal and social psychology*, 51:629–636.

Dunwoody, S. (2005). Weight-of-evidence reporting: What is it? Why use it? *Nieman Reports*, 54:89–91.

Dunwoody, S. and Kohl, P. A. (2017). Using weight-of-experts messaging to communicate accurately about contested science. *Science Communication*.

Elga, A. (2007). Reflection and Disagreement. *Nous*, 41:478–502.

Figdor, C. (2018). Trust me: News, credibility deficits, and balance. In Fox, C. and Saunders, J., editors, *Media ethics, free speech, and the requirements of democracy*, pages 69–86. Routledge, London.

Fitelson, B. (1998). The plurality of bayesian measures of confirmation and the problem of measure sensitivity. *Philosophy of Science (Proceedings Supplement)*, 64:S362–78.

Fitelson, B. (2001). *Studies in Bayesian Confirmation Theory*. PhD thesis, University of Wisconsin, Madison.

Frost-Arnold, G. (2019). How to Be a Historically Motivated Antirealist: The Problem of Misleading Evidence. *Philosophy of Science*, 86:906–917.

Gelbspan, R. (1998). *The Heat is On: The Climate Crisis, the Cover-up, the Prescription*. Perseus books, Cambridge, MA.

Gerken, M. (2020). How to balance Balanced Reporting and Reliable Reporting. *Philosophical Studies*, 177:3117–3142.

Goldman, A. I. (2001). Experts: Which Ones Should You Trust? *Philosophy and Phenomenological Research*, 63:85–110.

Gust, D., Weber, D., Weintraub, E., Kennedy, A., Soud, F., and Burns, A. (2008). Physicians who do and do not recommend children get all vaccinations. *Journal of Health Communication*, 13:575–582.

Howson, C. (2000). *Hume's Problem: Induction and the Justification of Belief*. Clarendon Press, Oxford.

Intemann, K. and de Melo-Martín, I. (2018). *The Fight Against Doubt: How to Bridge the Gap Between Scientists and the Public*. Oxford University Press, Oxford.

Jenness, A. (1932). The role of discussion in changing opinion regarding a matter of fact. *The Journal of Abnormal and Social Psychology*, 27:279–296.

Joint Statement (2001). The Science of Climate Change. *Science*, 292:1261.

Leuschner, A. (2018). Is it appropriate to 'target' inappropriate dissent? On the normative consequences of climate skepticism. *Synthese*, 195:1255–1271.

List, C. and Goodin, R. E. (2001). Epistemic Democracy: Generalizing the Condorcet Jury Theorem. *The Journal of Political Philosophy*, 9:277–306.

Miller, B. (2013). When is consensus knowledge based? Distinguishing shared knowledge from mere agreeement. *Synthese*, 190:1293–1316.

Nye, M. J. (1980). N-rays: An Episode in the History and Psychology of Science. *Historical Studies in the Physical Sciences*, 11:125–156.

Odenbaugh, J. (2012). Climate, consensus, and contrarians. In Kabasenche, W. P., O'Rourke, W., and Slater, M. H., editors, *The Environment: Philosophy, Science, and Ethics*, pages 137–150. MIT Press, Cambridge, MA.

Oreskes, N. (1999). *The Rejection of Continental Drift*. Oxford University Press, Oxford.

Oreskes, N. (2007). The scientific consensus on climate change: How do we know we're not wrong? In DiMento, J. F. and Doughman, P., editors, *Climate Change: What It Means for Us, Our Childen, and Our Grandchildren*, pages 65–99. MIT Press, Cambridge, MA.

Oreskes, N. (2019). *Why Trust Science?* Princeton University Press, Princeton.

Oreskes, N. and Conway, E. (2010). *Merchants of Doubt*. Bloomsbury Press, New York.

Peirce, C. S. (1958). *The Collected Papers of Charles Sanders Peirce*. Harvard University Press.

Pew Research Center (2015). An Elaboration of AAAS Scientists' Views. Available at `https://www.pewresearch.org/science/2015/07/23/an-elaboration-of-aaas-scientists-views/`.

Popper, K. (1965). *Conjectures and Refutations: The Growth of Scientific Knowledge*, chapter 3: Three Views Concerning Human Knowledge, pages 97–119. Routledge, London.

Scholz, O. R. (2009). Experts: What They Are and How We Recognize Them – A Discussion of Alvin Goldman's Views. *Grazer Philosophische Studien*, 79:187–205.

Sherif, M. (1935). A study of some social factors in perception. *Archives of Psychology*, 27:1–60.

Sklar, L. (1981). Do Unborn Hypotheses Have Rights? *Pacific Philosophical Quarterly*, 62:17–29.

Stanford, P. K. (2006). *Exceeding Our Grasp: Science, History, and the Problem of Unconceived Alternatives*. Oxford University Press, Oxford.

Strevens, M. (2004). Bayesian Confirmation Theory: Inductive Logic, or Mere Inductive Framework? *Synthese*, 141:365–379.

Strevens, M. (2008). *Depth: An Account of Scientific Explanation*. Harvard University Press, Cambridge, MA.

Strevens, M. (2017). How Idealizations Provide Understanding. In Grimm, S., Baumberger, C., and Ammon, S., editors, *Explaining Understanding: New Essays in Epistemology and Philosophy of Science*, pages 37–49. Routledge, New York.

Tucker, A. (2003). The epistemic significance of consensus. *Inquiry*, 46:501–521.

Weatherall, J. O. and O'Connor, C. (2020). Conformity in scientific networks. *Synthese*, DOI: https://doi.org/10.1007/s11229-019-02520-2.

Weatherall, J. O., O'Connor, C., and Bruner, J. (2019). How to beat science and influence people: Policymakers and propaganda in epistemic networks. *British Journal for the Philosophy of Science*, forthcoming.

Weisberg, M. (2007). Three Kinds of Idealization. *Journal of Philosophy*, 107:639–659.

Weisberg, M. (2013). *Simulation and Similarity: Using Models to Understand the World*. Oxford University Press, Oxford.