

When is consensus knowledge based? Distinguishing shared knowledge from mere agreement

Boaz Miller

Received: 25 January 2011 / Accepted: 4 June 2012 / Published online: 4 December 2012
© Springer Science+Business Media Dordrecht 2012

Abstract Scientific consensus is widely deferred to in public debates as a social indicator of the existence of knowledge. However, it is far from clear that such deference to consensus is always justified. The existence of agreement in a community of researchers is a contingent fact, and researchers may reach a consensus for all kinds of reasons, such as fighting a common foe or sharing a common bias. Scientific consensus, by itself, does not necessarily indicate the existence of shared knowledge among the members of the consensus community. I address the question of under what conditions it is likely that a consensus is in fact knowledge based. I argue that a consensus is likely to be knowledge based when knowledge is the best explanation of the consensus, and I identify three conditions—social calibration, apparent consilience of evidence, and social diversity, for knowledge being the best explanation of a consensus.

Keywords Social epistemology · Knowledge · Consensus · Expert testimony

1 Introduction

A consensus in a community of researchers is often perceived and regarded as a mark of knowledge shared by its members. Roman physician Galen (1963, p. 58) writes that the fact that philosophers quibble and unlike mathematicians cannot reach consensus means that they do not have knowledge. Kant (2007, pp. Bvii–xxxv) writes that metaphysics has not matured to be a proper science because metaphysicians cannot reach consensus. In science, the National Institute of Health (NIH)¹ and the

¹ <http://consensus.nih.gov/ABOUTCDP.htm>.

Intergovernmental Panel on Climate Change (IPCC)² formulate expert consensus statements to provide authoritative answers to disputed questions. *Wikipedia*'s official policy states that editorial decisions should be reached by consensus because "consensus seems to offer the best method to establish and ensure neutrality and verifiability".³ Scientific consensus is also deferred to when arbitrating between rival experts in legal trials and public policy disputes.

It is unclear, however, that such practices of consensus building and deference to it are epistemically legitimate. A long tradition in philosophy is suspicious of consensus as a knowledge indicator. By this tradition, agreement may indicate stagnation. Mill argues that the existence of dissent is necessary for correcting our views when they are wrong and justifying them when they are right (1993, pp. 83–123). Feyerabend argues that a constant stream of rival theories is essential for the advancement of science (1970, pp. 209–214), and Rescher (1993) argues against associating consensus with truth or the end of inquiry.

Given these conflicting views and considerations, the question to ask is not *whether* consensus is a mark of knowledge, but *on what conditions*. The question this paper addresses is when we may legitimately attribute a consensus in an epistemic community to knowledge that is shared by its members, when we do not have an independent way to find out whether the consensual view is true or false. I call such a consensus a "knowledge-based consensus". I identify conditions under which we may justifiably infer that it is likely that a given consensus in a real epistemic community is knowledge based. When such a situation obtains, we may justifiably believe that p from the fact that there exists a consensus that p .

My argument is twofold: (a) a consensus is likely to be knowledge based when knowledge is the best explanation thereof, (b) knowledge is the best explanation of a consensus when three conditions obtain together:

1. *The Social Calibration Condition*—all parties to the consensus are committed to using the same evidential standards, formalism and ontological schemes;
2. *The Apparent Consilience of Evidence Condition*—the consensus is based on varied lines of evidence that all seem to agree with each other;
3. *The Social Diversity Condition*—the consensus is socially diverse.

I argue that these conditions are independent. This means that an epistemic assessment of a consensus must address both the properties of the evidence supporting the consensual theory, and the social characteristics of the consensus. Additionally, these conditions are not binary; meeting them is a matter of degree. The more they are met, the more likely the consensus is knowledge based.

Let me stress what this paper is *not* about. I neither address the question of whether consensus should be the aim of inquiry, nor do I prescribe norms researchers should follow in their inquiry. In addition, this is not a full-blown conceptual analysis of the notion of knowledge-based consensus. My aim and motivation are different. Often in public debates, such as about global warming, deference to consensus is done to resolve disputes. My aim is to examine under what conditions such deference is legitimate.

² <http://www.ipcc.ch/ipccreports/ar4-syr.htm>.

³ <http://en.wikipedia.org/wiki/Wikipedia:Consensus>.

Namely, under what conditions we may infer p from there being a consensus that p , when, by supposition, we neither know whether p nor have an independent way to establish the truth or falsehood of p .

This paper consists of five sections. In Sect. 2, I discuss my methodology and assumptions. In Sect. 3, I explore the relations between knowledge, luck and inference to the best explanation, and argue that a consensus is likely to be knowledge based when knowledge is the best explanation thereof. In Sects. 4, 5 and 6 respectively I argue for my three conditions for knowledge-based consensus.

2 What knowledge-based consensus is

This paper is an exercise in normative naturalized social epistemology, which entails two central commitments. The first is that epistemic normative claims are empirical conjectures or hypothetical imperatives, linking means and ends. The second is that empirical evidence, e.g. historical and sociological studies of scientific-knowledge production and the success of different methods in achieving epistemic aims, is relevant to adjudicating between such conjectures.⁴ This paper is *not* a full-blown conceptual analysis of the notion of knowledge-based consensus. Rather, in this section and the next section, I put forward a rough conceptual account of what knowledge-based consensus means and entails. I will draw on this account in the remainder of the paper, where I discuss the conditions under which we may legitimately infer that a consensus is likely to be knowledge based.

Standard theories of knowledge take as their starting point a conception of knowledge as true justified belief. How are we to understand this conception in the context of consensus, particularly scientific consensus? Let us start with truth. As Steup (2008, §1) observes, the notion of truth simpliciter does not adequately characterize *scientific* knowledge. Scientific knowledge calls for a notion of partial or approximate truth. First, even our best scientific knowledge is never perfect or complete. It is constantly improving, evolving, and being corrected. Second, scientific theories include idealizations and fictions, which are, strictly speaking, false, but can nevertheless be regarded—under some conditions—as approximately true descriptions of their target systems (Chakravarty 2007, pp. 212–234; Ben Menahem 1988; Miller 2012, pp. 9–16).

For example, Newtonian dynamics is believed today to be, strictly speaking, a false theory, which has been superseded by Einstein's General Relativity. It is nevertheless still actively taught in universities and used for accurately describing a range of phenomena, from bridges to satellites. Newton's equations are mathematically derivable from those of General Relativity. Both theories give the same descriptions of systems at low velocities. They only come apart for systems close to light speed. Newtonian dynamics is partially true, then, in the sense that it correctly describes the behaviour of a very wide class of systems, and for those it does not, General Relativity can be used to explain its failure. Note that one can accept this claim, whether one thinks that

⁴ See Laudan (1996) at 156–157 and Solomon (2002) at 137–141 for two statements of normative naturalized social epistemology. For an influential naturalized social epistemology, see also Longino (2002).

reference to theoretical terms such as mass is preserved from Newtonian Dynamics to General Relativity or not.

What about belief? How is consensus to be characterized in terms of belief? It is tempting to characterize group beliefs as a summation of the beliefs of most of the group members. On a plausible summative account, a group G believes that p if and only if most of its members believe that p , and it is common knowledge in G that most of its members believe that p . Gilbert (1987) argues, however, that such an account is unsatisfactory. Suppose that members of a committee A happen to be the same as committee B . Suppose that after some discussion, A reaches a collective belief that q . Intuitively, it does not follow that B collectively believes that q merely because its members are the same as A . A summative account of group belief, then, does not seem to adequately characterize the phenomenon of consensus.

This paper therefore adopts Gilbert's influential non-summative account of group belief, according to which members of a population P collectively believe that p if and only if they are jointly committed to believing that p as a body (2002, p. 42). This account does not require that members of the population individually believe that p . Rather, it means, roughly, that they have agreed to let p stand as the position of the group. The joint commitment entails that they endorse p when participating in group activities, and publicly defend p when acting as group representatives. While Gilbert first required that the joint commitment be arrived at by an explicit process, as in a verbal agreement (1987, pp. 124–125), she later acknowledged that it might also be implicit (Gilbert 1994).

How are collective beliefs to be characterized as cognitive states? Here I side with Wray (2001, pp. 326–328), who argues that they are a species of *acceptance* rather than proper belief. Several differences exist between belief and acceptance. Accepting a claim is taking it for granted in one's reasoning, and it is possible to accept a claim without believing it. Acceptance often results from a consideration of one's goals, which may be epistemic or non-epistemic, while beliefs are not typically deliberately acquired to advance goals. Acceptance is voluntary, whereas belief is involuntary. Belief results in a feeling that something is true; acceptance involves no such feeling (Wray 2001, p. 325).

Given these differences, there are compelling reasons to regard collective beliefs as species of acceptance. Unlike proper beliefs, views are often adopted by plural subjects as a means of realizing the group's goals. Further, the sorts of considerations required to change a person's belief are usually epistemic—they relate to things such as the evidence that supports it, while those required to change a group's view may be pragmatic. A group may be persuaded that holding a different view may better promote its shared goals, whether epistemic or not, or its members may choose to change its goals and change its view accordingly (Wray 2001, pp. 326–327).

There are also particular reasons to characterize *scientific* consensus as a species of acceptance. First, a scientific theory sometimes face difficulties with explaining anomalies, but lacking a better theory, scientists may justifiably stick to it, though they may not take it to be literally true. Second, because theory is underdetermined by data, scientists may choose a theory based on pragmatic considerations such as which theory is easier to work with, where such considerations are detached from the question of the

truth of the theory. In such cases, acceptance, rather than belief, correctly characterizes scientists' attitude toward their theories (Cohen 1992, pp. 90–92).

Moreover, in science, the question of whether to believe certain putative facts, e.g. about unobservables, or merely accept them depends on the stance one adopts, where a stance is a set of general commitments and strategies for generating factual beliefs. The stance one adopts, e.g. a realist or an empiricist stance, depends on one's *personal* epistemic values. Individual scientists may vary in such values, and thus may vary in the propositions they believe or merely accept (Chakravartty 2007, pp. 17–26). For pursuing research and generating knowledge together, scientists are only required to accept the same basic assumptions. It is not required that *qua* individuals they all believe rather than merely accept the same claims (more on this in the next section). For example, while some 18th century scientists working within the caloric paradigm, which conceptualized heat as an unobservable fluid called “caloric”, were committed to its reality, others adopted an instrumentalist stance toward caloric. Regardless, they were all committed to using the notion of caloric in their reasoning (Psillos 1994, pp. 166–168; Fox 1971, p. 24).

A related question is the scope of agreement that is required to say that a consensus exists. Tucker (2003, pp. 509–510) argues that agreement amounts to consensus only when it is complete and there is no dissent whatsoever. I reject this view. In real-life cases, complete agreement is hardly ever present, and it is doubtful whether any wide agreement in the history of science has ever been complete. The mere existence of dissent is not enough for discounting certain wide agreements as genuine consensuses. For example, it is widely believed today that the HIV virus causes AIDS. There are few dissenting scientists who disagree, and their view is dismissed as groundless by the vast majority in the scientific community.⁵ The belief that HIV causes AIDS constitutes a basis for research, treatment, and prevention. It is taught and presented as a fact in university curricula and to the public. The dissenters are suppressed and marginalized by the mainstream scientific community, and face extreme difficulties challenging the majority view. Though a pocket of dissent exists, it is fair to say that there is a consensus that HIV causes AIDS.

Last, let me address the role of rationality and idealization in my account. Philosophers often discuss idealized models of consensus among agents converging on the truth.⁶ Such models typically assume that agents are equally rational, competent, and share the same evidence. Because ideal cases are short of the difference-making complexities of real-life cases, there are good reasons to doubt the value of ideal accounts to actual cases of agreement and disagreement in the real world (Christensen 2009, p. 765; Tucker 2003, pp. 502–504).

I therefore refrain from making such idealized assumptions. I assume that agents, scientists included, are mostly rational at best, have hot and cold psychological biases, have personal interests such as in professional advancement, social interests such as in ideology and religion, and have only partial and non-overlapping accessibility to the available evidence. This makes my theory applicable to evaluating real-life cases.

⁵ The chief dissenter is biologist Peter Duesberg and he argues for his views in his (1996). See Epstein (1996, Chaps. 3 & 4) for the history of the AIDS controversy.

⁶ For example, Peirce (1877); Lehrer and Wagner (1981), and Habermas (1984).

In this section I identified the main properties of a putative knowledge-based consensus. Roughly, it is a joint acceptance of a (hopefully) approximately true theory, such that there is very wide, but possibly incomplete, agreement on it. In the next section, I will discuss the notion of epistemic luck in relation to the role of inference to the best explanation (IBE) in my account.

3 Epistemic luck and inference to the best explanation

In this section, I argue that a consensus is likely to be knowledge based when knowledge is the best explanation thereof. I start by identifying two forms of epistemic luck the elimination of which is required for having knowledge. Then I discuss their relations to IBE.

Gettier (1963) has famously challenged the traditional analysis of knowledge. A central problem that Gettier-type problems expose is what Pritchard (2005, p. 146) calls “veritically lucky” beliefs. For example, if I form a belief by looking at a broken watch that happens to show the right time, although my belief is true and arguably justified, it is not knowledge. It is just a fluke. A veritically lucky belief is true in the actual world but false in a wide class of nearby possible worlds in which the agent forms her belief in the same way. My belief about the current time is false in most nearby possible worlds in which I form it by looking at the broken watch. A knowledge-based consensus cannot be veritically lucky.

There is another type of epistemic luck, the elimination of which is central to any account of knowledge. I call it *epistemic misfortune*. It refers to circumstances in which agents are justified or seemingly justified in their views but there are factors that systematically and/or deliberately mislead them or inhibit their gaining knowledge. Unlike veritic luck, they do not just happen to get things right, but rather they keep getting things wrong.⁷ Sceptical scenarios such as being a brain in a vat are examples of epistemic misfortune, because in them, despite their best efforts, agents are systematically and/or deliberately prevented from gaining knowledge of the world. Epistemic misfortune also exists in cases in which the available evidence points at the wrong direction, when agents share a bias or prejudice that interferes with their evaluation of the evidence, or when different biases push different agents to the same falsehood. A knowledge-based consensus cannot be a victim of epistemic misfortune.⁸

⁷ The distinction between epistemic misfortune and veritic epistemic luck is similar to the distinction that Statman (1991) draws between two forms of epistemic luck, respectively: (1) luck in the causes or circumstances that bring some subject *S* to believe *p*, or to be such a person who believes *p*; (2) luck with regard to *p* being true or false.

⁸ Analytic epistemologists, interested in conceptual analysis of knowledge, have not directed much attention to epistemic misfortune, and have not tried to rule it out from their conceptions of knowledge. This is probably because when agents are epistemically misfortunate, their beliefs are false. Hence, any conception of knowledge as true belief will have already ruled out their beliefs as knowledge. When agents are veritically lucky, on the other hand, their beliefs are accidentally true, and need to be ruled out for being a fluke. As opposed to epistemologists, philosophers of science, who are interested in the different ways beliefs or theories may seem or be justified or rational and still be false, discuss cases involving epistemic misfortune, though have not explicitly invoked this term. Recall that the aim of the theory proposed in this paper is to identify when we can infer that *p* from there being a consensus that *p*, when we do not know whether

Not all biases amount to epistemic misfortune. What distinguishes a mere bias from epistemic misfortune is that in epistemic misfortune, left to their own devices, agents' chances of correcting their misguided views are slim to null. Sometimes careful reflection may lead agents to realize that their views are biased,⁹ for example by sexist or racial prejudice, and lead them to revise them accordingly. However, if the epistemic misfortune is genuine, they will not be motivated or disposed to perform such reflection, and even when they are, they may fail.¹⁰

While in epistemic misfortune, the parties to the consensus are victims of the circumstances, sometimes they are less innocent—they deliberately form a consensus despite their awareness that the consensual view falls short of knowledge. These are cases of *non-cognitive consensus*. They are not knowledge based because people are likely to form the consensus whether the consensual view is true or not.

People form a non-cognitive consensus for many reasons. Consensus is a powerful political tool for advancing policy and promoting social aims. A group of experts may reach a consensus and mask existing disagreements within it to advance a goal. The experts may paternalistically decide it is better for the public that they speak in one voice; they might worry that their social status might be undermined if disagreements among them became public; they may wish to gain material support; etc. To promote the group's collective interests, dissenters may refrain from voicing their own views or even express views that they do not hold outside the expert community and even inside it (Beatty 2006, pp. 53–56).

As an example for such a non-cognitive consensus, Beatty (2006, pp. 55–64) discusses the U.S. National Academy of Sciences report entitled *The Biological Effects of Atomic Radiation* (1956). This was a consensus statement drafted by leading geneticists, which masked major disagreements within them about the safe ranges of atomic radiation. Beatty identifies two reasons for masking these disagreements. First, geneticists worried that if they did not present a unified front, physicists would claim expertise on questions of radiation safety. Second, they shared a conviction that the public could not cope well with scientific uncertainty.

To sum up the discussion so far, to assess whether a situation of knowledge obtains, we need to eliminate veritic luck, epistemic misfortune and non-cognitive reasons as possible good explanations of the consensus. This brings us to IBE. Under the IBE model, we legitimately infer the truth of our best explanation of some given facts. For example, if I hear squeaky sounds in my kitchen, the cheese in my kitchen cupboard

Footnote 8 continued

p and cannot establish the truth or falsehood of *p* in an independent way. We therefore must go in a roundabout way and exclude veritic epistemic luck and epistemic misfortune as likely to be present.

⁹ This is what distinguishes epistemic misfortune from Pritchard's notion of reflective epistemic luck (2005, p. 175).

¹⁰ For example, in *The Decent of Man*, Darwin (1871) repeatedly draws on the accepted belief of his time that savages are cognitively inferior to Europeans. He often portrays them as a brutal animal-like intermediate step in the evolution of man as evidence for human evolution. Today such claims are regarded as false and inadequate evidence for evolution. Yet, it is unclear whether a white upper-class man brought up in Victorian England could have seen these people in a different light. By contrast, Darwin's (1871, p. 326) depiction of women as cognitively inferior to men is arguably not a case of epistemic misfortune, as Darwin knew women of comparable cognitive abilities to men, and his argument from evolutionary theory is designed to counter claims by his contemporaries that women are cognitively equal to men.

is chewed, and I see a mousehole under the cupboard, I legitimately infer that there is a mouse in my kitchen, as this is the best explanation of these facts. According to IBE, explanatory merits such as simplicity, scope and elegance determine which explanations are the best. Explanatory considerations determine plausibility, so the best explanation is the likeliest explanation (Lipton 2004, pp. 60–62).

Why are we entitled to treat explanatory merits as likeliness indicators? In a manner consistent with this paper's being an exercise in normative naturalized social epistemology, I side with the view that our justification to do so is context relative, specifically, relative to our background knowledge. On this view, IBE does not name a formal inference, but an abstract pattern whose force and success depend on the specific background assumptions involved, and whose specific form changes with context (Day and Kincaid 1994, p. 282). Explanatory merits are empirical observations. For example, if a court infers a defendant's guilt from the fact that her fingerprints were found in the crime scene, a person matching her description was seen fleeing the crime scene, etc., this inference reflects the court's knowledge of how crimes are usually committed. Generally speaking, we observe regularities, see that some events are more frequent than others, and develop our explanatory merits based on these observations (Ben Menahem 1990, pp. 322–334).

In most contexts, the connection between explanatory merits and credibility is hardly disputed. What is disputed is which of the available explanations is best, and whether it is good enough. For example, in a criminal trial, the defendant's defence attorney will try to raise doubts as to his client's guilt being the best explanation of the crime, but he will not try to doubt IBE as such (Ben Menahem 1990, p. 325; Lipton 2004, p. 192). Similarly, I take my first claim that a consensus is likely to be knowledge based when knowledge is the best explanation thereof to be unproblematic and trivial. What is problematic and disputable, and hence the focus of this paper, is *when* knowledge is the best explanation of a consensus.

The three conditions I identify aim to rule out veritic luck, epistemic misfortune and non-cognitive reasons as the best explanation of a consensus. When these conditions obtain, we may legitimately attribute the consensus to shared knowledge. Because IBE is a fallible inference method, my account does not *guarantee* that in all cases where my conditions obtain the consensus is knowledge based, only that it is *likely* to be so. That is, when we eliminate such alternative explanations, knowledge remains the best explanation of the consensus. In such cases, we may legitimately infer that the consensus is likely to be knowledge based, i.e., likely to be neither veritically lucky nor epistemically misfortunate nor non-cognitive.

In this section, I discussed anti-luck conditions for knowledge and their relation to the role of IBE in my argument. I now turn to argue for social calibration, which is the first of the three conditions for knowledge being the best explanation of a consensus.

4 The social calibration condition

Sometimes an alleged agreement exists, but a closer look reveals that it is loose, superficial, set around vague terms that are susceptible to multiple interpretations, and hence not genuine. Such an alleged agreement is not a putative knowledge-based consensus. An example is the consensus document of the United Nations International

Conference on Population and Development (ICPD) drafted in Cairo in 1994. Halfon (2006, p. 788) characterizes it as “a functional consensus” that allows “disparate communities of people to act ‘as if’ cognitive agreement were in place”. He argues that it was achieved mainly by standardizing demographic data across many countries while maintaining interpretive flexibility in the meaning of key terms such as “women’s empowerment”, “family planning”, and “reproductive rights” (2006, 791–801). If this is indeed the case, the ICPD consensus cannot be knowledge based (cf. van der Sluijs et al. 1998).

To avoid such a scenario, agents must agree on some minimal content, for example, the meaning of key terms they use. What is this minimal content? The social calibration condition¹¹ addresses this question. It requires that agents agree regarding what it is they agree upon and share the same *fundamental* background assumptions. It ensures that the agreement is sufficiently genuine and not superficial.

Following Kuhn (1970, pp. 182–191), we can identify three types of objects of such consensual meta-agreement. The first is *shared formalism* (or in Kuhn’s terms, “symbolic generalizations”), for example, $f=ma$ in Newtonian physics and $2H_2O \rightarrow 2H_2 + O_2$ in analytic chemistry. The second type is *ontological schemes* (“metaphysical models”) which are descriptions or models of the building blocks and furniture of the world, such as that matter is composed of particles. As Kuhn notes—a point which will be of significance—members of a community may vary in their strength of commitment to such models along a spectrum from mere heuristics to realistic models of the world. The third is evidential standards (“exemplars”) which are model solutions that show how to apply the formalism to solve specific problems and define what solutions are acceptable.¹²

Shared evidential standards are not necessarily linked to a Kuhnian paradigm. They may also be cross-paradigm methods of reasoning. As Hacking (1992) argues, statistical reasoning comes equipped with an ontological scheme that describes the world as composed of populations, which have properties such as distribution and standard deviation. The procedures for calculating an average or standard deviation are defined independently of what the population represents—be it people or particles. Similarly, the method of scientific reasoning with an analogical model comes equipped with an ontological scheme that describes the unobservable microscopic world as analogous to macroscopic mechanical models. Here as well, scientists may vary in their beliefs about the reality of the objects such ontological schemes describe, while all using the schemes in their reasoning.

Fuller (2002, pp. 208–210) worries that a joint commitment restricted only to shared formalism, ontological schemes and evidential standards in the sense just discussed does not satisfy the social calibration condition for knowledge-based consensus, and that true unity of conceptual beliefs is required as well. I reject this view. I argue that such a joint commitment satisfies the social calibration condition, even if members of the consensus differ in their conceptual interpretations of these schemes.

¹¹ I borrow the term “social calibration” from Goldberg (2007, p. 61), who argues that a great overlap in the meaning of lexical terms between the idiolects of members of a speech community is necessary for satisfying the conditions on successful transmission of knowledge through testimony.

¹² See Kusch (2002, pp. 152–157) for an account of evidential standards as shared communal exemplars.

To understand why Fuller's worry is unsubstantiated, let us examine it in more depth. Fuller distinguishes two types of consensus: essential and accidental. In an essential consensus, a group forms a collective decision for the same thing using shared standards of evidence and sense of relevance. In an accidental consensus, each individual forms the same belief *on her own and for her own reasons*. A paradigmatic accidental consensus is the agreement found among a random group of surveyed people. Fuller argues that it cannot be knowledge based, because pollsters do not check that everyone surveyed understands the question in the same way, and by paraphrasing a question, the extent of consensus may be manipulated. Recall that knowledge excludes veritic luck. Accidental consensus typically fails to do that. Because it is contingent and easily manipulatable, even if it is right, it may have easily been wrong. Hence, it is not knowledge-based.

Fuller regards a scientific consensus that is restricted to shared formalism and evidential standards as an accidental consensus. Fuller claims that "to establish that the scientists have agreed upon a certain mathematical formalism [...] is hardly enough to show that they have decided to pursue something in common" (2002, p. 219). He worries that such a consensus may mask deep conceptual differences due to incommensurable assumptions that may surface later (2007, p. 10). For a consensus to be essential, Fuller requires that the parties to it also have uniformity of interpretation and conceptual understanding of the theory.

While I agree with Fuller that only an essential consensus can be knowledge based, I disagree that a uniformity of interpretation and conceptual understanding is required for an essential consensus. Pace Fuller, agreement on shared formalism, ontological schemes, and evidential standards provides a sufficient basis for pursuing "something in common". A mature mathematical formalism constitutes a representation of specific target systems that "*mediates* between theoretical 'first principles' and descriptions of empirical situations in their pure complexity" (Gelfert 2011, p. 282). As such, it constrains researchers in that it constitutes a rigid set of rules of manipulation of symbols, derivation, and inference. A commitment to a formalism is *ipso facto* a commitment to accepting derivations and inferences done with it. Such formalisms, then, constitute the syntax and minimal semantics required to calibrate collective research (Gelfert 2011; cf. Morrison 1999). This argument may also be extended to non-mathematical formalisms and ontological schemes, such as those of analytic chemistry.¹³

For example, all interpretations of quantum mechanics, e.g. the Copenhagen, many-worlds, and hidden-variables interpretations are committed to using the same mathematical formalism that represents quantum states as vectors in a Hilbert space. *It is exactly this commitment that ensures that they are interpretations of the same theory, and latch on to the same empirical phenomena in the same way.* At the same time, they significantly diverge in their metaphysical interpretations of these phenomena.

It follows from that, of course, that when a controversy exists about the metaphysical interpretation of a shared formalism, such as the interpretation of quantum

¹³ It is not required that all parties to the consensus actively use the same formalism, only that they all accept each others' formalisms. For example, organic chemists use a specific formalism that is an extension of the more general formalism of analytic chemistry.

mechanics, we should qualify the scope of the consensus and limit it to the empirical content of the theory as expressed by the shared formalism. Namely, what all interpretations of quantum mechanics accept, and only it, is arguably the object of a knowledge-based consensus. Similarly, when differences in conceptual understanding make some members reject certain evidential standards despite their commitment to the shared formalism, such as in the case of mathematical intuitionists' rejection of non constructive proofs, we cannot attribute a knowledge-based consensus to the group.

Fuller, however, may resist the conclusion that agreement on shared formalism, ontological schemes, and evidential standards satisfies the social calibration condition by providing another argument, which I call *the argument from logical conjunction*: In a consensus that is restricted to shared formalism, if we look at each person and construct a logical conjunction of her reasons for holding the consensual view, we will most likely get a logically consistent set of beliefs. Hence, each person's view may be rational. However, because members of the consensus differ in their conceptual interpretations, if we construct a logical conjunction of *all the members'* reasons for holding the consensual view, we will most likely get an incoherent set. Hence, the consensus is irrational and cannot be knowledge-based (Fuller 2002, pp. 220–221; cf. Pettit 2006).

Recall, however, that in Sect. 2, we characterized consensus as group belief. A group belief that p does not require that members of the group individually believe that p , only that they agree to let p stand as the group's position. Recall also that consensus is a species of acceptance, rather than proper belief. Acceptance only requires that one take certain claims for granted for the sake of reasoning, without necessarily taking them to be true.

If consensus is a species of acceptance rather than belief, then Fuller's argument from logical conjunction is misplaced. It states that the set of all the group members' reasons for *believing* the consensual view may be incoherent. It does not follow that the conjunction of their reasons for *accepting* it is incoherent. Reasons for acceptance depend on the group's collective goals and its members' individual goals. These reasons only need to be logically coherent inasmuch as people share the same goals. For example, in a group of scientists, different persons may have different views about the aims of science, but as a group they may agree on an epistemic aim such as empirical adequacy as a common denominator. Their collective set of reasons for accepting the consensus view need only be logically coherent with respect to that aim. True uniformity of conceptual beliefs is not required for social calibration.

Social calibration understood, under my suggestion, as joint acceptance of fundamental evidential standards, ontological schemes, and shared formalism strikes the right balance between preventing the consensus from being accidental, which disqualifies it from being knowledge-based, and allowing the parties to it to maintain a diversity of perspectives, views and interpretations, which, as I will argue in the next sections, is *required* for knowledge-based consensus. An essential consensus can exist despite conceptual disagreements among the group, and may therefore be knowledge based.

In this section, I presented the social calibration condition, which specifies the minimal content parties need to agree on for their consensus to qualify as knowledge based. In the next sections I discuss two additional conditions.

5 The apparent consilience of evidence condition

There is a common intuition, which states, roughly, that it is likely that a consensus is knowledge based when the parties to it are independent of one another in the ways they form their views. I will explore this intuition and argue that it amounts to two separate conditions. The first, to which this section will be devoted, is apparent consilience of evidence. The second, which is social diversity, will be discussed in Sect. 6.

5.1 First approximation: Goldman's causal independence condition

Goldman asks: "If a putative expert's opinion is joined by the consensual agreement of other putative experts, how much warrant does this give a hearer for trusting the original position?" (2001, p. 98). Goldman argues that while at first blush, simply trusting the majority may seem like a good idea, it is actually not. For example, if a group blindly follows a guru, its consensual view carries no more credence than the guru's view.

Goldman (2001, pp. 99–102) argues that a person X agreeing with Y carries weight only inasmuch as X is more likely to agree with Y that p if p is true than if p is false. This happens when X 's and Y 's routes to believing that p are at least partly causally independent, for example, when X 's and Y 's beliefs are based on independent experiments or eye witnessing. When X comes to believe that p after hearing Y saying that p and critically reflecting on it, X 's belief is partly causally independent of Y 's.¹⁴

Goldman's view seems intuitive, but faces difficulties. It counters the common wisdom that critical deliberation between individuals makes their final conclusion more warranted than the conclusion at which each individual would have arrived alone, even if they all formed the same belief in a causally dependent way. This intuition underpins the common practice of drafting expert consensus statements.

Additionally, large groups can effectively inter-divide the cognitive labour and reach more warranted results than isolated individuals. Evidence suggests that because of that, large groups do much better in that respect than small groups (Thagard 1999, p. 176). Although members of such groups rely on each other's results, thus they make their respective views causally interdependent, we have seemingly good reasons to trust their consensual view. Last, causally isolated agents may reach the same conclusion for different and incoherent reasons, thus violating the social calibration condition.

¹⁴ Goldman's reasoning echoes with Condorcet's jury theorems, which state, roughly, that sufficiently large groups of individuals in which there is a sufficiently large subgroup of individuals who have a higher than 0.5 probability to form a correct belief on a given matter will reach the correct decision on that matter by the method of majority voting. The application of these mathematical theorems to concrete real-world cases, however, is far from trivial. It is not clear in which concrete cases we would expect the conditions of statistical independence and higher than 0.5 probability to obtain, or even how to judge whether they obtain (Vermeule 2009, pp. 28–33).

Goldman's causal independence condition, then, is unsatisfactory. In the next subsection I present Tucker's "knowledge hypothesis" and argue that it manages to overcome the difficulties with Goldman's formulation, but as I argue in Sect. 5.3, it does not overcome the challenges posed by Solomon's Social Empiricism. This will allow me to develop my alternative.

5.2 Second approximation: Tucker's knowledge hypothesis

Tucker (2003) identifies three conditions under which shared knowledge is the best explanation of a consensus: First, the consensus is uncoerced; second, it is uniquely heterogeneous; and third, it is sufficiently large. Let us look at each of them in turn.

Tucker argues that when a consensus is coerced, namely a group is threatened, intimidated or bullied into holding a view, its shared view is not knowledge. An example is the consensus in the USSR on Lysenko's theory of biological adaptation; scientists who did not support it were at risk of being sent to the Gulag.

Tucker's second and main condition is that the consensus is uniquely heterogeneous, namely no subgroup shares an extraneous property that may otherwise explain the agreement within it. Tucker draws an analogy between uniquely heterogeneous consensus and a controlled scientific experiment. In a controlled experiment, if members of a test group do not share any property other than the one being tested, an observed effect may be attributed to the tested property. Similarly, when there is a consensus in a group of people who do not all share a property such as a mutual power relationship, joint interest, shared ideology or bias, the consensus may be attributed to knowledge (2003, p. 506). This is why the third condition, that the group is sufficiently large, is required. While an accidental consensus in a small uniquely heterogeneous group is likely, it is unlikely in a large heterogeneous group, or so Tucker argues.

How is Tucker's unique heterogeneity condition an improvement of Goldman's causal independence condition? The unique heterogeneity condition may be seen as an application of Goldman's causal independence intuition to the group level, avoiding the difficulties with Goldman's individualistic formulation of it. Rather than looking at each individual and see to what extent she formed her view in a causally independent way, we look at the level of the larger epistemic community and examine the extent each *subgroup* formed its position in a causally independent way. Each subgroup may correspond to a discipline or sub-discipline, a school of thought, etc.

Tucker's formulation of the condition allows us to take into consideration the size of the group along with its members' causal dependence. When we examine subgroups rather than individuals, we may give more credence to the joint view of a subgroup than the view of an individual working in isolation, and to large groups more than small groups. Thus, the unique heterogeneity condition manages to capture the intuition behind Goldman's causal independence condition, while avoiding the difficulties with it.

But is it indeed unlikely that members of a sufficiently large uniquely heterogeneous group would form a consensus even if they did not have knowledge? Solomon argues

that it is not. In the next section, I present Solomon's rival account, according to which an accidental aggregation of views toward a non-knowledge-based consensus is plausible. I will present Tucker's objections and argue that they are unsuccessful. Drawing on the lessons from this discussion, I will propose my condition.

5.3 A challenge for Tucker: Solomon's accidental aggregation theory of consensus

In *Social Empiricism* (2002) Solomon develops a normative social epistemology of division of cognitive labour and knowledge in science. She conceptualizes cognitive diversity in terms of factors ("decision vectors") that influence individuals' and communities' theory choice. She identifies many types of decision vectors that affect scientists' theory choice, such as motivational, social, cognitive, religious, and ideological decision vectors. She distinguishes between empirical and non-empirical decision vectors. Empirical decision vectors are all the factors that make scientists prefer theories with empirical success (2002, pp. 51–63), where empirical success is any scientific success that is contingent on how the world is, such as predictive, retroductive, experimental, explanatory, or technological success (2002, pp. 17–22). Non-empirical decision vectors are other reasons or causes for theory choice, such as ideology, pride, or preference for simpler theories (2002, pp. 51–63).

Solomon argues that dissent is epistemically normatively appropriate, when theories on which there is dissent each have associated empirical success, and empirical decision vectors are equitably distributed among theories, i.e. in proportion to the theories' empirical success. For instance, if some theory has some technological success, a proportional number of scientists in the relevant community should be drawn to it *because of this success*. It follows from Solomon's conditions that a consensus is epistemically justified exactly when one theory has all of the empirical success. Such cases are typically rare. Thus, permanent dissent is typically desirable. Dissent is not a temporary glitch to be eventually overcome, and consensus is not the end of inquiry (2002, pp. 117–120).

Solomon argues that the history of science shows that consensus often emerge out of an *accidental aggregation of non-empirical decision vectors, and are thereby unjustified*. Thus, there are cases in which a theory enjoys a consensus because of a combination of factors although there are rival empirically successful theories (2002, pp. 121–135). Therefore for Solomon, even if a consensus is uncoerced, uniquely heterogeneous, and sufficiently large—as Tucker requires—it is not necessarily likely to be knowledge based.

Tucker rejects Solomon's accidental aggregation hypothesis as a good *ceteris paribus* explanation of a consensus. He interprets Solomon as arguing that consensus is likely even given conflicting biases. If C stands for a consensus, K stands for knowledge and $B_1 \dots B_n$ stand for biases, then *ceteris paribus*, he takes Solomon to argue that it is likely that

$$[\star] P(C|B_1) \times P(C|B_2) \times \dots \times P(C|B_n) > P(C|K).$$

Tucker takes Solomon to argue that *ceteris paribus*, if we divide a group of people into n subgroups, where each subgroup shares a bias,¹⁵ the probability that a subgroup that shares a bias B_1 and a subgroup that shares a bias B_2 , and so on, would all reach a consensus is greater than the probability that the group would reach a consensus if the consensus view amounted to knowledge.

Against Solomon, Tucker argues that $[\star]$ is false. For $[\star]$ to be true, we need to assume that:

- (a) the likelihood of the consensus given each bias is initially very high;
- (b) the probability of the consensus given shared knowledge is very low.

Tucker (2003, p. 503) regards these two assumptions as implausible, thus he rejects Solomon's accidental aggregation hypothesis.

There are several difficulties with Tucker's objection. As for (a), Tucker is wrong to assume that the likelihood of consensus given each bias is low. It is plausible that different biases would all pull at the same direction. Consider the following example. Starting from the mid 1950's and lasting about thirty years, there was a misguided scientific consensus that excess acidity in the stomach caused peptic ulcers. This consensus lasted in spite of various attempts to challenge it by proponents of the rival theory according to which ulcers were caused by bacteria. In 1984, scientist Barry Marshall had to perform the dramatic and extraordinary act of infecting himself with the ulcer-causing bacteria, and curing himself with antibiotics in an attempt to overturn the consensus. (Marshall worked with Robin Warren, and they eventually won the 2005 Nobel Prize in Medicine for their work.) If we examine the factors responsible for the endurance of the misguided consensus, we find several independent ones. One factor is that Palmer, who published in 1954 an influential paper that allegedly refuted the bacterial theory, was a prominent scientist whose work was considered definitive, while proponents of the bacterial theory were of significantly lower status and peripherally located. Another factor was that accepted physiological theories gave plausibility to the view that the stomach was too acidic an environment to sustain living bacteria. Last, it was hardly in the interest of pharmaceutical companies to support the bacterial theory, as it meant that ulcers could be cured with a one-time treatment of antibiotics, rather than being a chronic illness that required the continuous take of antiacids, which were patent-protected until 1994 (Fukuda et al. 2002).¹⁶

Are such *ad hoc* coalitions so unlikely that they can be ruled out as good *ceteris paribus* explanations of a consensus? The answer is negative. When people and groups have a variety of interests and hold views on many issues, they are almost bound

¹⁵ This formulation may already not accurately reflect Solomon's claim, since biases are usually not mutually exclusive. I will not delve into this point.

¹⁶ My explanation of the endurance of the consensus over the excess acidity theory differs from other accounts. Thagard (1999, pp. 64–69; 2000, pp. 230–237) argues that the consensus over the excess acidity theory endured because until the mid 1980s, the rival bacterial theory did not exhibit explanatory coherence. Zollman (2010) argues that the scientific community failed to converge on the truth due to the mutual effect of two factors: the prevalence of extreme views within the community, and rapid information sharing. Both Thagard and Zollman, however, do not take into account the effect of social factors, such as commercial interests and differences in social status between actors, without supporting their exclusion of these factors with argument. For a useful timeline of the events surrounding the discovery see: http://en.wikipedia.org/wiki/Timeline_of_peptic_ulcer_disease_and_Helicobacter_pylori.

to agree on occasion on *some things*. Different groups often have mutual interests which will cause them to reach a consensus on occasion on a particular matter despite disagreeing on other things.

As for (b), the prior probability of consensus given knowledge is what is in dispute. Even if people individually have knowledge, they might not reach consensus for reasons such as mutual misunderstanding, personal rivalries, and difficulties to communicate (Thagard 2000, pp. 236–237). The prior probability of consensus given knowledge cannot be assumed to be high without begging the question.

In response to these two claims, Tucker may still argue that in [\star], when n is large enough, the left side (“biases”) goes to zero. Namely, even if the probability of consensus given each bias is initially high and the prior probability of consensus given knowledge is low, when there are enough biases, their convergence toward a consensus is still very unlikely.

This claim is problematic because the expression on the left side of [\star] is misleading. It describes the probability of *complete agreement with no dissent whatsoever*, which is very low anyway. For example, imagine a group of 100 people, where each person is 0.95 likely to believe that p . By Bernoulli’s theorem, the probability that exactly all 100 people will agree that p is close to zero,¹⁷ but the probability, for example, that more than 90 will believe that p is close to one.¹⁸ In fact, Solomon’s (2002, Chap. 7) historical counterexamples of accidentally formed and epistemically unjustified consensuses are of wide agreements with only minor dissent. Tucker argues that they do not refute his claim because they are not examples of a genuine consensus: If there is a dissenting view, the agreement does not amount to a consensus. Recall from Sect. 2, however, that it not required that absolutely no dissent be present to say that a consensus exists. It is enough that a very wide agreement exists, and dissenting views are marginalized and excluded from the mainstream. Tucker’s objection therefore fails.

5.4 The apparent consilience of evidence condition

So far I argued that although Tucker’s account improves Goldman’s causal independence condition, it does not overcome Solomon’s accidental aggregation hypothesis. Does this mean that an IBE account of knowledge-base consensus is hopeless? Not necessarily. It is still intuitively plausible that *ceteris paribus*, when many people who do not share much in common reach consensus, shared knowledge is its best explanation. This just shows that Tucker’s criteria fail to adequately capture this intuition. Let us, then, examine what underpins it.

Notice that Tucker and Goldman are not interested in social unique heterogeneity and causal independence as such. If under Goldman’s model, different individuals happened to form their belief in exactly the same way despite their causal independence, their combined consensual view would carry no more credence than the each

¹⁷ It is $0.95^{100} = 0.0059$.

¹⁸ The probability that m or more people will believe that p is $\sum_{i=0}^m \binom{100}{i} 0.05^i \cdot 0.95^{100-i}$. For example, if $m = 90$, the probability is 0.9885.

individual's view alone. This would be just like the guru case. Goldman and Tucker take such a scenario to be very unlikely. Their respective conditions are *proxy conditions*. Ultimately, what matters for them is that different agents form their views in different ways, and they assume that causally independent individuals or subgroups that do not have an extraneous property in common tend to form their views in significantly different ways.

Underlying their accounts is the thought that the more varied the ways people form the same view, the more credence it deserves. If many people come to hold that p in significantly different ways, this is probably because that p is true or approximately true. Knowledge would be the best explanation of why they all hold the same view.

This is an application of the general notion of *robustness* to the social context. Robustness is the idea that “hypotheses are better supported with plenty of evidence generated by multiple techniques that rely on different background assumptions” (Stegenga 2009, p. 650). Producing evidence using multiple techniques and under different background assumptions aims at eliminating influences that are *accidental to the particular way* a hypothesis is tested. For example, if the same pattern is observed when the same sample is placed in different types of microscopes, it is likely that the observed pattern is accurate, rather than a by-product of the particular way a certain microscope operates.

In the social context, the robustness principle holds that when a consensus is built on an array of evidence drawn from a variety of techniques and methods, it is less likely to be an accidental by-product of one technique—and all the more likely to be knowledge based. I suggest that in the social context, convergence of multiple techniques should mean *apparent consilience*¹⁹ of different lines of evidence.

Apparent consilience of evidence is what Tucker's unique heterogeneity condition ultimately strives for. How is that so? Typically, different groups—different disciplines and sub-disciplines—use different methods and different types of evidence. For example, in mass torts, evidence is used from animal studies and from human epidemiological studies. Both of them constitute different types of evidence that correspond to localized groups—epidemiologists and toxicologists—that have their own journals, societies, etc.

Not only do disciplinary boundaries make researchers use different types of evidence, but also geographic, national and other social barriers. For example, historians talk about “national styles” in science. A famous example is the divide in the nineteenth century between French physicists and chemists, who favoured abstract mathematical reasoning, and English scientists, who favoured concrete mechanical models and visual diagrams (Duhem 1954, pp. 70–71; Nye 1993). Similarly, Collins (1998) attributes significant differences in interpreting gravity-waves data between contemporary American and Italian physicists to their having different “evidential cultures”. Because different social groups tend to use different evidence, Tucker's unique heterogeneity condition is a proxy for apparent consilience of evidence.

¹⁹ I borrow the term “consilience” from Whewell, who talks about the principle of “consilience of inductions”, according to which hypotheses are more supported when they independently stem from different inductive inferences (1858, pp. 87–90).

While unique heterogeneity is a proxy for apparent consilience of evidence, it is not good enough. Different social groups and different disciplines do not always use different methods and different types of evidence. It follows from the argument in 4.3 that we cannot assume it will be so unlikely to find a uniquely heterogeneous group whose members happen to use the same methods and rely on the same evidence. *We should therefore explicitly specify that we are interested in apparent consilience of evidence, rather than unique heterogeneity, as a condition for knowledge-based consensus.*

Notice that the condition of apparent consilience is different from actual consilience. It only requires that all existing evidence *seem* to support the consensual view, namely that *there be no seemingly contradictory available evidence that is ignored, overlooked or otherwise suppressed by the members of the consensus group by social or other means.*

One may wonder why only *apparent* consilience of evidence is required, rather than actual consilience. This is because actual consilience is too demanding a condition. The difficulty with determining whether the evidence is actually consilient, as opposed to seemingly consilient is threefold. First, *de facto* multiple methods for combining and weighing different lines of evidence are available, both generally in science and in specific disciplines. Douglas (2012) reviews commonly used methods, such as expert elicitation, meta-analysis, causal machine learning, and explanatory approaches. As she notes, different methods and even different implementations of the same method may give different outcomes for the same body of evidence.

For example, proponents of Evidence-Based Medicine (EBM) advocate the “hierarchy of evidence” method for combining and ranking different types of evidence. Roughly, according to the EBM hierarchy, a randomized controlled trial (RCT) is better evidence than an observational study, and a systematic meta-analysis of RCTs is better evidence than a single RCT. Yet, in practice, different implementations of the EBM hierarchy weigh and rank evidence differently, and may lead to different conclusions for the same body of evidence (Upshur 2003); and different methods of meta-analysis may also diverge in their outcomes (Stegenga 2011). Given such an abundance of methods, which method, if any, should be chosen to evaluate whether the evidence is *actually* consilient?

Second, even if there were only one widely agreed upon method for combining and weighing evidence, appealing to it for answering the question of when consensus is knowledge based merely because it enjoys a wide consensus would be question-begging. There must also be strong independent epistemic rationales supporting it. As Douglas’ review reveals, however, all existing methods have their pros and cons. None is epistemically superior to all others in all respects. For example, despite the EBM hierarchy’s popularity and its proponents’ claims that it is the best and only scientific way to assess medical evidence (Straus et al. 2010), philosophers of science strongly doubt the epistemic rationales underpinning it (Worrall 2002; Borgerson 2009; La Caze 2009).

Third, our current best methods for determining evidential support leave room for subjectivity. For example, Bayesianism allows agents to diverge in their prior probabilities, and IBE allows agents to diverge in their ranking of the features of a good explanation. Consequently, people may rationally disagree on whether there is

consilience of a given body of evidence. (Douven 2009, pp. 347–350). Similarly, all the methods of evidence weighing that Douglas reviews, qualitative and quantitative alike, include a subjective element, because they crucially require the use of judgement in their correct application. That is, our best confirmation theories and evidence weighing methods cannot be used dogmatically. They do not provide algorithms whose outcome is invariant to the agents employing them. Rather, they provide guidelines for agents for reaching an *informed judgement*, which partly depends on their competence and background assumptions, which, in turn, partly depend on their social, cultural, and disciplinary background. Thus, apparent consilience of evidence is the best we can hope for.

Since evidential judgements partly depend on their agents' social situatedness, it seems plausible that a theory of knowledge-based consensus needs to evaluate the social features of the consensus as well as the epistemic features of the evidence on which it is based. In the next section, I make this argument. I argue that social diversity is also required for a knowledge-based consensus, and that this condition is irreducible to apparent consilience of evidence.

6 The social diversity condition

In the last section, I argued that apparent consilience of evidence captures the robustness intuition applied to the question of when a consensus is knowledge based. In this section, however, I argue that it does not *fully* capture the robustness intuition, and social diversity must be added as an independent condition.

Recall, robustness is the notion that hypotheses are better supported with plenty of evidence generated by multiple techniques that rely on different background assumptions. In social epistemology, the multiple backgrounds condition is often cashed out in terms of social diversity, as diverse people tend to have different background assumptions. Longino writes that a “diversity of perspectives is necessary for vigorous and epistemically effective critical discourse [...] When consensus exists, it must be the result [...] of critical dialogue in which all relevant perspectives are represented” (Longino 2002, p. 131).

Diversity has many epistemic benefits. Diversity may generate new research questions, identify limitations with existing models, propose new models, propose alternative hypotheses and interpretations of data, open up new lines of evidence, reveal “loaded” language in descriptions of phenomena, and more adequately identify and weigh potential risks (Intemann 2009).

Mill writes that a person's beliefs are a product of his social background (1993, p. 86). Feminist epistemologists add that because certain perspectives are often inseparable from certain social identities, even in open and critical settings, there are limits to people's ability to transcend their background and free themselves of their biases and prejudice. Longino (2002, p. 132) argues that the absence of women and ethnic minorities from a scientific consensus, even if not intentional, constitutes a serious cognitive flaw, which reduces the community's critical resources. Thus, *de facto* social diversity rather than mere openness to different views is required for knowledge-based consensus.

One may argue that the condition of social diversity is reducible to apparent consilience of evidence. Ultimately, we want to realize the robustness condition for consensus. Therefore, while diversity may be instrumental in bringing about a variety of background assumptions required for robustness, it is not necessary. It is sufficient that the different converging lines of evidence seem to be based on sufficiently varied background assumptions, where social diversity is just one means of achieving that.

I reject this objection. Social diversity is a standalone condition irreducible to apparent consilience of evidence. The production and assessment of evidence is a social process. Therefore, the appearance of consilience of evidence may itself be due to some underlying social reality, rather than the existence of knowledge in the consensus community.

Consider the following thought experiment. Suppose you discover that a scientific consensus exists that passive smoking does not raise the chances of lung cancer. Suppose this consensus exhibits apparent consilience of evidence. Studies of different types support this conclusion: Epidemiological studies show no significant correlation between passive smoking and lung cancer, structural-analysis studies suggest that cigarette smoke undergoes some chemical reaction in the open air that reduces its carcinogenic effects, etc. Suppose further these studies do not seem to be based on some common problematic background assumptions. This consensus, so the objection goes, is knowledge based.

Suppose you later discover that all these studies were supported or partly supported directly or indirectly by tobacco companies. Is knowledge still the best explanation of the consensus? Not any more. Regardless of what you thought of the consensus and the evidence before, they now become suspect. A better explanation of the consensus may be that the tobacco industry is responsible for bringing it about. For example, it may have given leading researchers incentives to produce evidence that supports its interests, and this evidence convinced other members of the scientific community who formed a consensus.

The upshot is that it is not enough for a consensus about the harmlessness of passive smoking to exhibit apparent consilience of evidence. Rather, it must be also socially diverse, namely shared by researchers from both the private and public sectors, with different financial ties, smokers and non-smokers, etc. Hence, social diversity is irreducible to apparent consilience of evidence.

To be clear, diversity controls for two types of influences—expected and unexpected. In the passive smoking case, we *expect* that corporate interests and personal smoking habits may bias research, so we require that a knowledge-based consensus include researchers with no financial ties with tobacco companies and non-smokers. Similarly, we *expect* views on gender to be influenced by a person's own gender, therefore a knowledge-based consensus on gender-related issues should include both men and women. But diversity is also required for controlling for *unexpected* influences. There may be influences of which we are unaware, and the more diversified the consensus, the less likely they will affect it. This is similar to the rationale behind the design of randomized clinical trials, which are required to control for both known confounders (sex, age, etc.) and unknown confounders.

One may object that the condition of general social diversity is too strong. Women and ethnic groups concentrated in developing countries, for example, have been

historically underrepresented and excluded from science, but it still managed to produce good theories. This shows, so this objection goes, that while diversity may be epistemically beneficial, it not necessary.

Against this I argue that lack of sufficient diversity is an epistemic problem, and scientific consensuses that were not sufficiently diverse may have indeed been less knowledge-based than they were thought to be. The prevalence of sexist and other biases in science has been extensively studied by historians and philosophers of science. I will mention a few prominent examples. [Martin \(1991\)](#) argues that because researchers were blinded by a stereotypical Sleeping Beauty/Prince Charming model of the egg and the sperm, they overlooked the active role of the egg in fertilization. Feminist anthropologists argue that social stereotypes about man as inventive and woman as passive contributed to the development of “man the hunter” theory of cultural development, where alternative theories that fit the same empirical evidence and attribute positive contributions to women were not seriously considered ([Slocum 1975](#); [Tanner and Zihlman 1976](#)). Gould (1996, p. 240) shows how intelligence testing conducted by the U.S. government in the 1920s on newly arrived immigrants presupposed knowledge specific to American culture. [Keller \(1983\)](#) argues that biologist Barbara McClintock’s unique methods and theories in heredity, which were initially unrecognized but ultimately won her the Nobel Prize, stem partly from her being a woman.

As [Okruhlik \(1998\)](#) argues, the importance of such examples is not which theories ultimately proved right, but that flaws in established orthodox theories and possible alternative theories were not seriously considered due to lack of sufficient social diversity in the scientific community. A central problem in theory choice is a failure to conceive of alternative explanations. Hence if for a putatively successful theory T , it very easily could have been the case that had we thought of an alternative T^* , we would not have accepted T , then if T is true, we are lucky to have accepted it. Social diversity increases the number of alternatives we consider, and hence we can be more confident that if T is true, it is not merely veritically lucky.

One may argue that such examples only show that diversity is needed in the biological and social sciences, but not in the hard sciences, namely physics and mathematics. This objection, however, is problematic. First, though the examples above are from the special sciences, they do not all directly address questions of gender. McClintock, for example, studied heredity in maize rather than humans. Second, the influence of social stereotypes extends beyond the special sciences. For example, [Wagner \(2009\)](#) argues that the use of gendered language in the formulation and proof of the “stable marriage” theorem has blinded mathematicians from some of its mathematical implications. While I do not want to overstate the case and claim that mathematics is male-biased, it seems that the burden of proof lies with those who hold that diversity only matters in special sciences. After all, the theories in the examples above were thought to be objective and value free *until* the cultural perceptions embedded in them were pointed out. Women are still underrepresented in the hard sciences, thus the potential epistemic benefits of more diversity in these sciences are yet to be discovered.

Moreover, diversity comes in degrees. It is not “all or nothing”. The exclusion of women does not mean lack of diversity, but less diversity. Recall, the conditions for

knowledge-based consensus are a matter of degree. The more they are met, the more likely a consensus is knowledge based.

7 Conclusion

I addressed the question of when a consensus is knowledge based. I have identified three conditions under which this is likely to be the case: social calibration, apparent concision of evidence, and social diversity. I assume that knowledge-based consensus is an achievable target, but whether knowledge-based consensus are ubiquitous or rare in actual scientific practice is an open question which calls for empirical investigation.

As I have stressed, the statement that a consensus over p is *not* knowledge based does not mean that p is false or unwarranted. Rather, it means that the fact that an agreement that p exists in an epistemic community does not carry additional credence for p . It does not give us a reason to infer that p from the fact that a consensus exists that p . The reasons of the parties to the consensus to hold that p may still be compelling and sufficient for justifiably holding that p or they may not. When there is evidence that a consensus is knowledge based, however, the fact that the consensus exists gives us an independent reason to justifiably hold that p . Only then, is deference to a view because it is a consensual view justified.

Acknowledgments I thank Hagit Benbaji, Joseph Berkovitz, Jim Brown, Anjan Chakravartty, Steve Fuller, Yves Gingras, Sandy Goldberg, Arnon Keren, Kareem Khalifa, Laszlo Kosolovsky, Yakir Levin, Isaac (Yanni) Nevo, Isaac Record, Jacob Stegenga, Eran Tal, Brad Wray, and two anonymous reviewers for useful comments and suggestions. I am grateful to the Azrieli Foundation for an award of an Azrieli Fellowship. Special thanks to Meital Pinto for her support and inspiration.

References

- Beatty, J. (2006). Masking disagreement among experts. *Episteme: A Journal of Social Epistemology*, 3(1), 52–67.
- Ben Menahem, Y. (1988). Models of science: Fictions or idealizations? *Science in Context*, 2(1), 163–175.
- Ben Menahem, Y. (1990). Inference to the best explanation. *Erkenntnis*, 33, 319–334.
- Borgerson, K. (2009). Valuing evidence bias and the evidence hierarchy of evidence-based medicine. *Perspectives in Biology and Medicine*, 52(2), 218–233.
- Chakravartty, A. (2007). *A metaphysics for scientific realism: Knowing the unobservable*. Cambridge: Cambridge University Press.
- Christensen, D. (2009). Disagreement as evidence: The epistemology of controversy. *Philosophy Compass*, 4(5), 756–767.
- Cohen, L. J. (1992). *An essay on belief and acceptance*. Oxford: Clarendon Press.
- Collins, H. M. (1998). The meaning of data: Open and closed evidential cultures in the search for gravitational waves. *The American Journal of Sociology*, 104(2), 293–338.
- Darwin, C. (1871). *The descent of man, and selection in relation to sex*. London: John Murray.
- Day, T., & Kincaid, H. (1994). Putting inference to the best explanation in its place. *Synthese*, 98(2), 271–295.
- Douglas, H. (2012). Weighing complex evidence in a democratic society. *The Kennedy Institute of Ethics Journal*, 22(2), 139–162.
- Douven, I. (2009). Uniqueness revisited. *American Philosophical Quarterly*, 46(4), 347–360.
- Duesberg, P. (1996). *Inventing the AIDS virus*. Washington, DC: Regnery Publishing.
- Duhem, P. (1954). *The aim and structure of physical theory*. Princeton: Princeton University Press.

- Epstein, S. (1996). *Impure science: AIDS, activism, and the politics of knowledge*. Berkeley, CA: University of California Press.
- Feyerabend, P. (1970). Consolations for the specialist. In I. Lakatos & A. Musgrave (Eds.), *Criticism and the growth of knowledge* (pp. 197–230). Cambridge: Cambridge University Press.
- Fox, R. (1971). *The caloric theory of gases: From Lavoisier to Regnault*. Oxford: Clarendon Press.
- Fukuda, Y., Shimoyama, T., Shimoyama, T., & Marshall, B. J. (2002). Kasai, Kobayashi and Koch's postulates in the history of *Helicobacter Pylori*. In B. Marshall (Ed.), *Helicobacter pioneers* (pp. 15–24). Oxford: Blackwell.
- Fuller, S. (2002). *Social epistemology* (2nd ed.). Bloomington, IN: Indiana University Press.
- Fuller, S. (2007). *The knowledge book: Key concepts in philosophy, science and culture*. Montreal: McGill-Queen's University Press.
- Galen (1963). On the passions and errors of the soul, trans. Paul W. Harkins. Columbus, OH: Ohio State University Press.
- Gelfert, A. (2011). Mathematical formalisms in scientific practice: From denotation to model-based representation. *Studies in History and Philosophy of Science*, 42(2), 272–286.
- Gettier, E. (1963). Is justified true belief knowledge? *Analysis*, 23, 121–123.
- Gilbert, M. (1987). Modeling collective belief. *Synthese*, 73(1), 185–204.
- Gilbert, M. (1994). Remarks on collective belief. In F. F. Schmitt (Ed.), *Socializing epistemology: The social dimensions of knowledge* (pp. 235–256). Lanham, MD: Rowman Littlefield.
- Gilbert, M. (2002). Belief and acceptance as features of groups. *Protosociology*, 16, 35–69.
- Goldberg, S. C. (2007). *Anti-individualism: Mind and language, knowledge and justification*. Cambridge: Cambridge University Press.
- Goldman, A. I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research*, 63(1), 85–110.
- Gould, S. J. (1996). *The mismeasure of man* (rev ed.). New York: Norton.
- Habermas, J. (1984). *The theory of communicative action, Vol. 1 and 2*. Boston: Beacon Press.
- Hacking, I. (1992). Statistical language, statistical truth and statistical reason: The self-authentication of a style of scientific reasoning. In E. McMullin (Ed.), *The social dimensions of science* (pp. 130–157). Notre Dame, IN: University of Notre Dame Press.
- Halfon, S. (2006). The disunity of consensus: International policy coordination as socio-technical practice. *Social Studies of Science*, 36(5), 783–807.
- Intemann, K. (2009). Why diversity matters: Understanding and applying the diversity component of the National Science Foundation's broader impacts criterion. *Social Epistemology*, 23(3–4), 249–266.
- Kant, I. (2007). *Critique of pure reason, trans. N. K. Smith* (2nd ed.). New York: Palgrave Macmillan.
- Keller, E. F. (1983). *A feeling for the organism: The life and work of Barbara McClintock*. San Francisco: W.H. Freeman.
- Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: The University of Chicago Press.
- Kusch, M. (2002). *Knowledge by agreement: The programme of communitarian epistemology*. New York: Oxford University Press.
- La Caze, A. (2009). Evidence-based medicine must be. *Journal of Medicine and Philosophy*, 34(5), 509–527.
- Laudan, L. (1996). *Beyond positivism and relativism: Theory, method, and evidence*. Boulder, CO: Westview Press.
- Lehrer, K., & Wagner, C. (1981). *Rational consensus in science and society: A philosophical and mathematical study*. Dordrecht: D. Reidel.
- Lipton, P. (2004). *Inference to the best explanation* (2nd ed.). London: Routledge.
- Longino, H. (2002). *The fate of knowledge*. Princeton: Princeton University Press.
- Martin, E. (1991). The egg and the sperm: How science has constructed a romance based on stereotypical male-female roles. *Signs*, 16(3), 485–501.
- Mill, J. S. (1993). On liberty. In J. M. Dent (Ed.), *Utilitarianism, on liberty, considerations on representative government* (pp. 69–187). London: Everyman.
- Miller, B. (2012). The rationality principle idealized. *Social Epistemology*, 26(1), 3–30.
- Morrison, M. (1999). Models as autonomous agents. In M. S. Morgan & M. Morrison (Eds.), *Models as mediators: Perspectives on natural and social science* (pp. 38–65). Cambridge: Cambridge University Press.

- Nye, M. J. (1993). National styles? French and English chemistry in the nineteenth and early twentieth centuries. *Osiris*, 8, 30–49.
- Okruhlik, K. (1998). Gender and the biological sciences. In M. Curd & J. A. Cover (Eds.), *Philosophy of science: The central issues* (pp. 192–208). New York: Norton.
- Peirce, C. S. (1877). The fixation of belief. *Popular Science Monthly*, 12, 1–15.
- Pettit, P. (2006). No testimonial route to consensus. *Episteme: A Journal of Social Epistemology*, 3(3), 156–165.
- Pritchard, D. (2005). *Epistemic luck*. New York: Oxford University Press.
- Psillos, S. (1994). A Philosophical study of the transition from the caloric theory of heat to thermodynamics: Resisting the pessimistic meta-induction. *Studies in History and Philosophy of Science*, 25(2), 159–190.
- Rescher, N. (1993). *Pluralism: Against the demand for consensus*. New York: Oxford University Press.
- Slocum, S. (1975). Woman the gatherer: Male bias in anthropology. In R. R. Reiter (Ed.), *Toward an anthropology of women* (pp. 36–50). New York: Monthly Review Press.
- Solomon, M. (2002). *Social empiricism*. Cambridge, MA: MIT Press.
- Statman, D. (1991). Moral and epistemic luck. *Ratio*, 4(2), 146–156.
- Stegenga, J. (2009). Robustness, discordance, and relevance. *Philosophy of Science*, 76(5), 650–661.
- Stegenga, J. (2011). Is meta-analysis the platinum standard of evidence? *Studies in History and Philosophy of Biological and Biomedical Science*, 42(4), 497–507.
- Steup, M. (2008). The Analysis of Knowledge. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2008 Edition). <http://plato.stanford.edu/archives/fall2008/entries/knowledge-analysis>.
- Straus, E., Glasziou, P., Richardson, W. S., & Haynes, R. B. (2010). *Evidence-based medicine: How to practice and teach it* (4th ed.). Edinburgh: Churchill Livingstone.
- Tanner, N., & Zihlman, A. (1976). Women in evolution. Part I: Innovation and selection in human origins. *Signs*, 1(3), 585–608.
- Thagard, P. (1999). *How scientists explain disease*. Princeton: Princeton University Press.
- Thagard, P. (2000). *Coherence in thought and action*. Cambridge, MA: MIT Press.
- Tucker, A. (2003). The epistemic significance of consensus. *Inquiry*, 46, 501–521.
- Upshur, R. E. G. (2003). Are all evidence-based practices alike? Problems in the ranking of evidence. *Canadian Medical Association Journal*, 169(7), 672–673.
- van der Sluijs, J., van Eijndhoven, J., Shackley, S., & Wynne, B. (1998). Anchoring devices in science for policy: The case of consensus around climate sensitivity. *Social Studies of Science*, 28(2), 291–323.
- Vermeule, A. (2009). *Law and the limits of reason*. New York: Oxford University Press.
- Wagner, R. (2009). Mathematical marriages: Intercourse between mathematics and semiotic choice. *Social Studies of Science*, 39(2), 289–308.
- Whewell, W. (1858). *Novum organon renovatum: Being the second part of the philosophy of the inductive sciences*. London: John W. Parker and Son.
- Worrall, J. (2002). What evidence in evidence-based medicine? *Philosophy of Science*, 69(3), S316–S330.
- Wray, K. B. (2001). Collective belief and acceptance. *Synthese*, 129(3), 319–333.
- Zollman, K. J. S. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, 72, 17–35.