⋆

Deliberation, Judgement and the Nature of Evidence

⋆

Jon Williamson

In *Economics and Philosophy*
Version of December 9, 2014

Abstract

One kind of deliberation involves an individual reassessing the strengths of her beliefs in the light of new evidence. Bayesian epistemology measures the strength to which one ought to believe a proposition by its probability relative to all available evidence, and thus provides a normative account of individual deliberation. This can be extended to an account of individual *judgement* by treating the act of judgement as a decision problem, amenable to the tools of decision theory. A normative account of *public* deliberation and judgement can be provided by merging the evidence of the individuals in question and calculating appropriate Bayesian probabilities and judgement thresholds relative to this merged evidence.

But this formal epistemology for deliberation and judgement lacks substance without an account of how evidence can be merged. And in order to provide such an account, we need in turn an account of what the evidence *is* that grounds Bayesian probabilities. This paper attempts to tackle these two concerns. After finding fault with several views on the nature of evidence (the views that evidence is knowledge; that evidence is whatever is fully believed; that evidence is observationally set credence; that evidence is information), it is argued that evidence is whatever is rationally taken for granted. This view has consequences for an account of merging, and it is shown that standard axioms for merging need to be altered somewhat.

§1
Deliberation and Judgement

Bayesianism provides a natural account of certain questions to do with judgement and doxastic deliberation:

Individual Deliberation. How strongly should I believe θ?

Individual Judgement. Should I judge θ as true?

Public Deliberation. How strongly should *we* believe θ?

Public Judgement. Should *we* judge θ as true?

Let us consider these questions in turn.

¶ *Individual Deliberation.* Turning first to the question of individual deliberation, the Bayesian would answer: believe θ to degree $P_E(\theta)$, where P_E is your rational belief function, relative to your total evidence E, which assigns a degree of belief to each proposition θ that is expressible in your language. Bayesians agree that to count as rational, your belief function must be a probability function and must satisfy constraints imposed by evidence E. (Bayesians disagree about exactly what constraints evidence imposes, about how belief functions should be updated in the light of new evidence, and about whether further norms impose constraints on the belief function, but these disagreements will not be relevant to the concerns of this paper.) Since $P_E(\theta)$ varies only with E and θ, for individual deliberation to result in a change in strength of belief in θ, the evidence base E must change, either through introspection or through observation or other forms of external interaction.

¶ *Individual Judgement.* To address the question of individual judgement, the Bayesian can treat the decision as to whether or not to judge θ to be true as she would treat any other decision, by invoking the apparatus of Bayesian decision theory. The Bayesian's answer to this question would then be: judge θ as true if the expected utility of judging θ is greater than that of not judging θ. Bayesian decision theory presumes that an individual is equipped with a utility function as well as a belief function, and that decision matrices can be populated with these utilities in order to evaluate the prospects of alternative choices. For example, a decision matrix for judging θ might look like:

	θ	$\neg\theta$
Judge θ	5	-4
Don't judge θ	-1	3

Here the utility of judging θ, should θ be true, is 5, and so on. The expected utility of judging θ is $5P_E(\theta)-4(1-P_E(\theta))$, which is greater than the expected utility of not judging θ, $-1P_E(\theta)+3(1-P_E(\theta))$, just when $P_E(\theta)>7/13$. In this example the Bayesian would say, then, that you should judge θ to be true when your rational belief function relative to your total evidence ascribes degree of belief greater than 7/13 to θ.

¶ *Public Deliberation.* The Bayesian can answer the question of public deliberation—how strongly should *we* believe θ?—by suggesting that we merge our evidence bases $E_1,\ldots,E_n$ to yield an evidence base E representing the evidence of the group, treat the group as an agent in its own right, and believe θ to degree $P_E(\theta)$ where P_E is the group's rational belief function relative to E.

Note that the Bayesian would say that if we as individuals act together as a group, i.e., with a common interest, then we as a group ought to adopt a single rational belief function, sometimes called our ***intersubjective probability*** (Gillies, 1991).[1] If the evidence base of the group or the interest of the group differs from those of its individual members then the group can adopt a rational belief function that differs from each of the individual rational belief functions, though of course the act of merging our individual evidence bases may spur us to revise those individual

[1]There is no need to assume a common prior probability function or common knowledge here (Aumann, 1976), merely a common interest.

evidence bases, leading to an iterative process of merger and reflection under which convergence between individual and group beliefs may occur.

Why not merge our rational belief functions directly, instead of merging the evidence bases which constrain them? This is for three reasons.

First, our individual belief functions may contain subjective elements that are irrelevant to group deliberation; merging our evidence, rather than our belief functions, helps to filter out these subjective contingencies. Now, different varieties of Bayesianism disagree as to the pervasiveness of subjectivity. The *subjective* Bayesian account is very permissive: for instance it deems me to be perfectly rational if I strongly believe that a particular football team will win its next match *in the total absence of any evidence either way*—all that is important is that my beliefs be consistent with the evidence that I have. On the other hand, the *objective* Bayesian account would insist that I be much more equivocal in such a situation, giving probability $\frac{1}{2}$ or thereabouts to the proposition that the team will win its next match, in the total absence of relevant evidence (assuming measures in place to break a draw in the match). Nevertheless, even on the objectivist account, P_E is not always uniquely determined by E (Williamson, 2010). Thus all varieties of Bayesianism agree that some room is left for arbitrariness with respect to the choice of belief function. If individual prior belief functions were directly merged, their arbitrary aspects would be inherited to some extent by the group belief function. This would leave a curious position in which only one belief function—the merger of the individual belief functions—is deemed rational as a group prior belief function, but where that belief function may be more opinionated than the evidence demands. For instance, if, under a subjectivist account, we all strongly believe that Forest Green Rovers will win its next match, through nothing short of blind optimism—i.e., in the absence of evidence—, then our merged belief function will presumably inherit this strong belief. Although individually we would have been equally rational had we each believed that Rovers will lose, if the group's rational belief function is the merger of the individual functions then the group is rationally compelled to believe that Rovers will win. By merging our evidence bases rather than our belief functions, on the other hand, the group will be given precisely the same leeway for subjective choice as the individuals which compose it: since we have no evidence regarding Rovers, neither will the group, and the individuals and the group will be judged by the same standards—permissive standards on the subjectivist account, but much less so on the objectivist account.

The second reason for merging at the level of evidence rather than belief is that important information in the evidence can be lost if merging takes place at the level of belief. One individual's degree of belief in θ may be based on little or no evidence, while another's may be based on strong evidence. If merging were to take place at the level of belief, i.e., if the merged belief function were determined solely by the individual belief functions, then both individuals would typically be given given the same weight in determining the extent to which the group believes θ, since the evidence is not taken into account. But then the influence of the stronger evidence is watered down by giving undue weight to groundless beliefs.

The third reason for not directly merging belief functions is that there appears to be no such merging operator that is entirely satisfactory. Suppose one wanted to merge individuals' belief functions and utility functions and use the merged functions for group decision making. Then one would presumably want to impose at least the following conditions on the belief and utility merging operators. (i) Pareto optimality: if the group chooses an option, there should be no alternative option that offers higher expected utility for all individuals. (ii) Unanimity: if each individual has the same belief function then this function is the group's belief function. (iii) Non-dictatorship: there should be no individual such that the group's belief function is always set to that individual's belief function. As Hylland and Zeckhauser (1979) show, there are no belief and utility merging operators that satisfy the conditions (i)-(iii).

¶ *Public Judgement.* Finally, let us consider the question of whether we should judge θ as true. As in the individual case, the Bayesian can apply the machinery of decision theory here. First, aggregate the utilities that populate the individuals' decision matrices for θ (see, e.g., Hild et al., 2008). As in the case of individual judgement, θ should be judged true just if the expected utility of so doing exceeds the expected utility of not so doing, except here the expectation is determined by the group's rational belief function P_E, where E is the merger of the individuals' evidence bases.

The question again arises: why not aggregate the individuals' judgements directly, rather than proceed indirectly by merging the individuals' evidence bases and then deriving group judgements via decision theory? There are three principal reasons—reasons that are analogous to the three reasons outlined above for not merging belief functions. First, individual judgements may be attributable to subjective elements, and merging at the level of evidence prevents this arbitrariness from propagating to the group judgements. Second, merging at the level of judgement may pass over valuable information encapsulated in the evidence, leading to sub-optimal group judgements. Third, there are notorious problems with the direct aggregation of judgements; indeed the literature on the direct aggregation of judgements is dominated by impossibility theorems (see, e.g., List, 2012).

¶ In order to lend substance to an account of group deliberation and judgement such as that sketched above, the Bayesian must of course say something about how evidence is to be merged. Which ways of merging evidence bases are rational?

In Williamson (2009) I suggested that the theory of *belief merging* (see, e.g., Konieczny and Pino Pérez, 2011) might be applied to the task of merging evidence. (Indeed, I argued there that this theory is better suited to merging evidence than merging beliefs.) In this paper we will look more closely at that proposal. Clearly, in order to evaluate a normative theory of evidence, one needs to have some sort of rudimentary understanding of what evidence is. We shall consider in turn several accounts of evidence—evidence is knowledge (§3); evidence is what is fully believed (§4); evidence is observationally set credence (§5); evidence is information (§6); evidence is what is rationally granted (§7)—and argue for the latter account. §8 will develop that account in more detail. Finally, §9 will evaluate the theory of belief merging as a theory of merging evidence, arguing that since evidence is what is rationally granted, three of the six axioms of the theory of merging need to be reformulated.

First, some general comments on the Bayesian notion of evidence.

§2
The Nature of Evidence

The Bayesian, whose principal concern is the relation between evidence and belief, customarily says rather little about evidence itself. Two truisms usually suffice.

First, evidence is what constrains rational belief. Here we should distinguish a *body of evidence* from one proposition being *evidence for* another: an agent's body of evidence is what constrains her rational belief function, while one such proposition is evidence for another if those propositions stand in the appropriate relation (positive probabilistic dependence, perhaps). We will solely be concerned with the agent's body of evidence here; not with the relational notion of evidence. An *evidence base* is a basis from which one can derive a body of evidence: for instance, $\{\theta, \theta \rightarrow \varphi\}$ is an evidence base for the body of evidence $\{\theta, \theta \rightarrow \varphi, \varphi\}$. In general, a body of evidence may have several evidence bases; in §9, in order to discuss the theory of belief merging, we shall need to suppose that each agent under consideration possesses some specific evidence base that gives rise to their body of evidence. While in this paper we will mainly be concerned with bodies of evidence and evidence bases that are sets of propositions, we do not need to assume that evidence is always propositional.

Bayesians have various views about how evidence constrains rational belief. For instance, some argue in favour of a *Calibration norm* (often called *Miller's principle* or the *Principal Principle*), which says that degrees of belief should be calibrated to any physical probabilities of which there is evidence. Other Bayesians disagree. But one constraint that all Bayesians adhere to is that any evidence should be fully believed:

Presumption. If $\theta \in E$ is expressible in the agent's language and $E \cup \{P(\varphi) = 1 : \varphi \in E\}$ is consistent, then $P_E(\theta) = 1$,

where E is the agent's whole body of evidence and P_E is her rational belief function.[2]

The second truism is that, when deliberating as to what belief function to adopt, one should take all one's evidence into account. This is the *principle of total evidence.* It does not assert that all one's evidence will be relevant to beliefs in question, merely that an agent's rational belief function is constrained by her whole body of evidence. Hence the customary notation P_E for an agent's rational belief function, where E is the agent's whole body of evidence.

One can use these two truisms as a springboard from which to say a bit more about evidence.

One can distinguish, for example, *rational evidence*—that which constrains rational belief—from *actual evidence*—that which constrains actual belief. The Bayesian will want to say that an agent's actual beliefs are rational if they are formed in the right way from her actual evidence (i.e., they adhere to the norms of Bayesianism, in that the strengths of her beliefs are representable by probabilities, they are appropriately constrained by her actual evidence, and whatever else is required) and the agent is rational to take as evidence what she actually takes as evidence (i.e., she adheres to the norms of evidence, whatever they turn out to be). Since Bayesian

[2] The consistency condition is required to rule out pathological cases such as that in which E contains both θ and $P_E(\theta) < 1$. Consistency is taken to mean deductive consistency with the axioms of probability.

epistemology is concerned with Bayesianism as a normative theory, we can cut to the chase by focusing attention on rational evidence and rational belief, leaving the question of the ways in which actual evidence and actual beliefs depart from the rational for psychology to answer. Thus in this paper an unqualified use of the word 'evidence' will be used to signify *rational evidence*, i.e., that which constrains rational belief and satisfies norms of evidence.

Furthermore, Bayesianism is supposed to be a guide to life: Bayesian epistemology is intended to offer advice as to how to determine rational degrees of belief that can feed into decision making, at which point Bayesian decision theory offers advice as to which decision to take. Given this practical goal, it should be possible to follow the norms of Bayesianism: Bayesian epistemology is subject to the *ought implies can* principle, which says that one can only require of an agent what is possible for her to achieve. For example, in discussions of logical omniscience, Bayesians agree that in cases such as mathematics, where one cannot expect an agent to determine all the logical consequences of all expressible propositions, the requirement of logical omniscience should be relaxed in some appropriate way (Corfield, 2001). The *ought implies can* principle also requires that an agent's evidence should be accessible to that agent. For if an agent cannot—perhaps after some appropriate process of introspection, calculation or deliberation—determine what her (rational) evidence is, how can her (rational) degrees of belief be expected to satisfy constraints imposed by that evidence?[3]

Thus any account of evidence should render evidence accessible to the agent in question, if it is to flesh out the Bayesian theory of deliberation and judgement outlined in §1. This implies, in particular, that the account of evidence must avoid epistemic circularity: it cannot say that $E = X$ if the X in question can only be determined by appealing to the agent's evidence E. In particular, it cannot say that $E = X$ if the X in question can only be determined by appealing to the beliefs that are grounded by the agent's evidence E, since one would need to determine E in order to determine those beliefs which are to be used to determine E.

In sum, we are concerned here with evidence in the sense of an agent's whole body of rational evidence (or an evidence base for that body). This evidence should be accessible to the agent, and an account of evidence should avoid epistemic circularity.

Let us turn, then, to some accounts of evidence, and evaluate them in the light of the above discussion.

§3
E=K, what is Known

The question of the nature of evidence is not only of interest to Bayesian epistemologists, but to epistemologists more generally, whose concept of evidence is much the same:

> It seems to be a platitude about evidence that a subject's evidence is what that subject has to go on in trying to arrive at a view. I would

[3] As we shall see, this demand for accessibility of evidence sets Bayesian epistemology apart from some epistemological theories, e.g., that of Timothy Williamson (2000). On the other hand, the accessibility requirement is line with other theories, such as that of Skorupski (2010, p. 21), who says, 'if autonomy is possible it must be possible to audit one's reasons by reflective self-examination and thereby give a warranted answer to the question, "Do I have sufficient reason for this belief, sentiment, or action?"'

> go further and add that a subject's evidence is limited to those things that the subject can properly treat as reasons for belief without needing antecedent justified beliefs to justify treating these considerations as proper starting points for deliberation. (Littlejohn, 2011, p. 242)

Littlejohn goes on to characterise an agent's stockpile of evidence as the body of propositions that fix evidential probabilities (Littlejohn, 2011, p. 248). Here 'evidential probability' does not refer to Kyburg's theory of evidential probability (see, e.g., Kyburg Jr, 1991)—rather, it is the name given by Timothy Williamson (Williamson, 2000, Chapter 10) to something akin to Keynesian logical probability (Keynes, 1921). Its proponents argue that this notion of probability is to be taken as primitive, and is not identical to rational degree of belief, but nevertheless typically grounds rational degree of belief in the sense that, except in pathological cases, the evidential probability of θ relative to E is the degree to which a perfectly rational agent with total evidence E would believe θ. Hence it is fair to say that under this view of evidence, evidence is also intended to ground rational degree of belief, if indirectly via logical probability.

The view taken by Williamson (2000, Chapter 9) and endorsed by Bird (2007, §2), Littlejohn (2011) and others is that evidence is knowledge, $E = K$. This is meant to work both ways: everything that you know is in your body of evidence and you know everything that is in your body of evidence.

This account of evidence seems at first sight to be well suited to Bayesian epistemology, since Bayesians often use 'background knowledge' interchangeably with 'evidence', deeming Bayesian probability to be relative to background knowledge in the same way that Bayesian probability is relative to evidence. On the other hand, this account of evidence is somewhat at odds with the intuitive understanding of evidence since it renders some evidence far from *evident*: if $E = K$ then one cannot always determine what one's evidence is because one cannot always tell whether or not one knows a given proposition (Williamson, 2000, §9.3).

While these considerations concern common usage and hence are rather inconclusive, the second consideration points to a serious objection: evidence, on this account, is inaccessible. Knowledge is true belief that has some authority in virtue of being justified, or reliably obtained, or suchlike. One may think one knows θ but be mistaken about its truth or authority, and one may be in no position to correct oneself. (While one may mistakenly think one believes θ, a Bayesian might say that in principle one can decide this question with a bit of effort, by observing one's own betting tendencies.) Hence, given the *ought implies can* principle, it is hard to see how one could be deemed irrational for failing to accord one's degrees of belief with one's evidence, if $E = K$. Williamson (2000, §9.3) bites the bullet here, partly because his version of evidence is used to constrain his version of evidential probability, which is equally inaccessible. But this move is no consolation to the Bayesian who remains in need of accessible evidence so that the norms of Bayesian epistemology can have normative force in practical applications.

The proponent of $E = K$ might respond that while we *may* be mistaken about our evidence we are typically not, in which case evidence is accessible enough to constrain rational degree of belief. But this is not so. What one takes to be one's evidence may be thought of as a conjunction θ of a great many propositions, at least a small proportion of which turn out false. So in practice θ is bound to be false and will not qualify as knowledge. If evidence is to constrain rational degree of belief and $E = K$ then θ should not constrain rational degree of belief. Thus if

$E = K$, we are always wrong to rely on what we take to be our evidence in order to determine our degrees of belief. Worse still, if θ is a sufficiently large conjunction of propositions, at least one of which is false, arguably one *knows* that θ is false, i.e., one knows $\neg\theta$.[4] If $E = K$, then one's evidence contains $\neg\theta$. Thus not only does θ fail to qualify as evidence, but one's beliefs should be constrained by it's negation: one's evidence is exactly the opposite of what one takes it to be!

A second serious concern for the Bayesian is that of epistemic circularity. If $E = K$, then in order to decide whether some proposition θ is evidence, one needs to determine whether θ is known and hence whether θ is a (qualitative) rational belief. And to determine rational beliefs one needs to isolate the evidence and investigate the constraints that it imposes. Hence, in order to determine whether θ is evidence one needs to have determined whether θ is evidence.

The proponent of $E = K$ might respond that there is no circle here because knowledge requires, not *rational* belief, but justified belief, or reliable belief, or belief that is qualified in some other way. But it still seems plausible that if a belief is to make the grade for knowledge, it will have to be suitably constrained by evidence. Indeed an apparent minimal condition on a knowledge-grade belief is that it should be a strong rational belief in the Bayesian sense: for if θ is known and $E = K$, then $\theta \in E$ and by the Presumption principle (§2), θ must be fully believed—one is not rationally entitled to disbelieve θ. Hence the circle seems not so easily avoided.

Bayesian epistemology is widely applied to the philosophy of science, and a third concern that the Bayesian might have about the $E = K$ claim is that it does not appear to accord with the concept of evidence used in science, which is not factive. In science, gold-standard evidence takes the form of observations or experimental outcomes corroborated by several independent trials. It is widely recognised that items of evidence are often false. The ideal is not that evidence all be true, but that it be obtained by methods deemed sound, in the hope that such methods are reliable and yield a high probability of truth. (Truth remains a goal of enquiry, if not a necessary condition on evidence.) In many cases, more or less precise error bounds are introduced in order to quantify the proportion of the evidence, or data, that is erroneous. There is no suggestion that these data points fail to qualify as evidence, as they would if $E = K$, and it is widely acknowledged that belief should be apportioned on the basis of total evidence, false evidence as well as true.

One area which highlights the non-factive nature of evidence in science is that of meta-analysis. A meta-analysis pools evidence from multiple studies in order to come up with an overall recommendation. For example, a meta-analysis for the question of whether exposure to a toxin causes a particular disease will examine multiple studies in the literature that bear on that question, merging the data of the studies where possible and aggregating their results, in order to produce an answer to the question that is based on all available evidence. A meta-analysis of multiple well-conducted randomised controlled trials is placed at the top of the so-called 'evidence-hierarchies' that pervade evidence-based medicine and other evidence-based movements across the sciences. So, without question, the data merged by a meta-analysis qualifies as evidence. But the reason why meta-analysis is so important is that evidence of any particular study can be erroneous and/or

[4] This is particularly clear when, as is frequently the case in scientific contexts, what one takes to be one's evidence is obtained by methods with a known level of reliability—say 99%. Then the chance of θ being true can be known to be some fixed number, small enough to render a true belief in $\neg\theta$ knowledge. We shall return to this sort of case below.

misleading. It is precisely because evidence may be false that such evidence is best placed in the context of a meta-analysis; the hope being that, through merging, the size of the merged body of evidence will be large enough that there will be proportionally little erroneous evidence.

There are two principal ways in which the proponent of $E = K$ might respond to the charge that evidential propositions in science are not always true. First, the defender of $E = K$ can distance herself from the evidence, arguing that the evidence in question is not the evidence as reported by a particular study, but instead E consists of propositions of the form *study S reports data D*; Williamson (2000, pp. 198–199) adopts a similar strategy in the case of erroneous perceptual evidence. Of course this is at odds with common parlance. Moreover, that sort of evidence would fail to license the conclusions that scientists typically draw: data D taken together with a claim of the form $D \rightarrow C$ will license the conclusion C, but the meta-evidence 'study S reports D' fails to license that conclusion in the presence of $D \rightarrow C$ without some further assumption concerning the credibility or reliability of the study. Since the reliability of any individual study is just what is in question in contexts such as meta-analysis, that reliability cannot be assumed. Instead, the approach taken in such cases is to take the data of the study as evidence and draw the conclusion defeasibly, in the awareness that evidence provided by further studies may lead the conclusion to be reversed. Thus this first response fails to account for the use of evidence in science.

Secondly, the defender of $E = K$ might offer the response that evidence must be true to make sense of basic intuitions about evidence:

> Why is it bad for an assertion to be inconsistent with the evidence? A natural answer is: because then it is false. That answer assumes that evidence consists only of true propositions. For if an untrue proposition, p, is evidence, the proposition that p is untrue is true but inconsistent with the evidence. (Williamson, 2007, p. 209)

But such a claim is too quick. An alternative answer to the question of why it is bad for an assertion to be inconsistent with evidence is: because the assertion is then unlikely to be true. This alternative answer does not assume that evidence consists of true propositions but merely that it consists of propositions that are probable in some appropriate sense—an assumption that accords much better with the use of evidence in science.

Thus far we have seen that the Bayesian might object to the $E = K$ claim on the grounds of inaccessibility, epistemic circularity, and incompatibility with scientific practice. A fourth concern for the Bayesian is that knowledge may be of too *weak* an epistemological standing to qualify as evidence. An astronomer may, for instance, know that an astronomical event has taken place merely by observing it with the naked eye, but this may be inadequate for it to count as evidence in an astronomical study, which can require more in the way of authority (justification, reliability, or whatever) than does knowledge. Similarly, the norms of evidence in a legal setting can require dimensions of authority other than those required by knowledge: knowledge that is irrelevant to the case in question, or hearsay, or opinion, may not count as evidence.[5]

[5] Legal evidence might even be non-propositional (and hence not knowledge): a murder weapon constitutes 'physical evidence' for example.

The proponent of $E = K$ might reply here that what counts as knowledge is context-relative, and in the context of an astronomical study, knowledge demands more than observation with the naked eye. But this is disputable. Arguably, the astronomer does know that the event has taken place by observing it with the naked eye—indeed, if anyone's naked eyes are reliable enough to yield knowledge of astronomical events, those of a trained astronomer surely are. Moreover, astronomical evidence can be expensive to collect, and knowledge (obtained, e.g., via the naked eye) that the event has taken place may be a prerequisite to justify the cost of collecting evidence to support the claim that the event has taken place. So it appears that in such cases it is evidence but not knowledge that is context relative.

§4
E=B, what is Believed to degree 1

Proponents of $E = K$ cite the claim that $E = B$ as their official opposition, and a standard Bayesian position (Williamson, 2000, p. 222; Littlejohn, 2011, p. 241). However, it is hard to find advocates of this latter view, so this may be a straw man.[6]

For sure, Bayesians would endorse the claim that if a proposition is a part of an agent's body of evidence, then that agent ought to believe it to degree 1—c.f., the Presumption principle of §2. But even here there are qualifications: the agent is hardly compelled to believe the evidential proposition to degree 1 if she cannot express or entertain that proposition. A robot agent may be able to express (i.e., have a belief function defined on) propositions about its location in a room, but not be able to express propositions encapsulating the evidence it receives in its sensors. I may see something clearly that I cannot describe—my evidence is not then captured by a proposition that I believe to degree 1, such as the proposition that I see something that I cannot describe; rather, my evidence consists of propositions about what it is that I see, propositions that are currently beyond my capacity to capture in language.

On the other hand, the Bayesian should be reluctant to endorse the converse claim: not everything that is believed to degree 1 need qualify as evidence. This is largely a consequence of the Bayesian thesis that degrees of belief are probabilities; probability theory forces probability 1 on many propositions that are highly contingent and not always evident. Examples are provided by the strong law of large numbers, the various zero-one laws, the claim that the Bayesian is forced to believe to degree 1 that her degrees of belief are perfectly calibrated with empirical frequencies (Dawid, 1982), however implausible that proposition might be, and the claim that the Bayesian is forced to believe to degree 1 that her degree of belief in θ is x, if indeed it is the case that $P_E(\theta) = x$ (Milne, 1991). Even where probability 1 is not forced, the Bayesian may still want to insist on awarding probability 1 to propositions that cannot be construed as evidence: for example the proposition that a dart will not hit a given point on a dartboard viewed as a continuous disc.[7]

---

[6] While Michael Bratman seems to claim that $E = B$ when he says that 'An agent's beliefs provide the *default cognitive background* for further deliberation and planning' (Bratman, 1992, p. 10), it is clear that he does not understand qualitative belief as that which is believed to degree 1.

[7] One might object that the mark made by a dart is not a point in the geometric sense because it has a positive area; hence this area does not have measure zero. But instead of the mark made by the dart we can consider the horizontal projection of the centre of mass of the dart onto the dartboard: this is a point in the geometric sense and does have Lebesgue measure zero.

Another reason for the Bayesian to reject the $E = B$ claim is that epistemic circularity rears its head again. Evidence is that which determines belief, so to find out whether one is rational to believe θ to degree 1, one needs to ascertain whether that strength of belief is compatible with one's evidence, given the norms of Bayesianism. In particular, one needs to know whether θ is in one's body of evidence, for if it is, then one should believe it to degree 1. But if $E = B$ then in order to determine whether θ is in one's body of evidence, one needs to determine whether it has (rational) degree of belief 1.

The proponent of $E = B$ might try to avoid this circularity by appealing to actual degree of belief 1 rather than rational degree of belief 1 (c.f., §2). Then, in order to determine whether one ought to believe θ to degree 1, one needs to determine whether θ is a part of one's evidence, i.e., whether one actually believes θ to degree 1. But such a move trivialises Bayesianism as a normative theory: actual degree of belief 1 becomes sufficient for rational degree of belief 1, since $P_E(\theta) = 1$ if $\theta \in E$. Thus if I am certain that Forest Green Rovers will win then I am rationally compelled to be certain that they will win, regardless of what information I have about them. Such a consequence is clearly unsatisfactory.

Finally, degree of belief 1 is too strong a condition on evidence, as one can arguably take as evidence propositions about which one which would be less than perfectly certain. It may appear to me that a jacket is orange, though I view the jacket by candlelight. While norms of belief may deem this insufficient grounds for awarding the proposition that the jacket is orange degree of belief 1 given my other evidence (including the evidence that the jacket appears to me to be orange), norms of evidence may deem the observation sufficiently reliable as to count as evidence henceforth. One might argue that every observation has this status: there will normally be some grounds for doubt which would preclude giving what is observed probability 1 relative to other evidence, but that does not preclude one taking what is observed as evidence, at least provisionally. (This view requires that evidence be defeasible. But on any view according to which evidence is accessible, evidence turns out to be fallible: for any—or almost any—proposition that one uses as grounds for one's beliefs, one can mistakenly take that proposition to be true. In which case evidence must be defeasible if one's beliefs are to continue to track the truth.)

§5
E=C, Credences set by observation

In order to deal with cases such as observation by candlelight, Richard Jeffrey put forward the thesis that credences that are the direct effect of observation—for example, my 0.8 degree of belief that the jacket is orange—are evidence in the sense that they motivate changes in other degrees of belief (Jeffrey, 1968, §2; Jeffrey, 2004, §3.2).[8]

A key advantage of the $E = C$ thesis is thus that it handles evidence that would otherwise be less than certain. A second advantage is that this position sidesteps the problem of inexpressible evidence, by treating credences as a proxy for that evidence: 'there is *something* about what is seen that leads the observer to have

[8] For Jeffrey, such a credence need not be evidence in the sense of a *reason* for belief: θ is a reason for believing φ if $P(\theta)$ and $P(\varphi|\theta)$ are both high (Jeffrey, 1968, p. 38). An observationally-set credence is part of a body E of evidence but need not be evidence *for* another proposition (c.f., §2).

the indicated degrees of belief ... but there is no reason to think this something expressible by a statement in the observer's language.' (Jeffrey, 1968, p. 35.)

Epistemic circularity is also less of a problem on the $E = C$ account, for the following reason. In order to determine what your evidence is, you can proceed by isolating those of your actual degrees of belief that have been directly causally determined by observation. Since you have no control over those causally-determined degrees of belief, and ought implies can, they can hardly be deemed irrational. Hence they are rational degrees of belief as well as actual degrees of belief, and can constrain the rest of your rational belief function P_E. So, although it may not be easy to isolate which of your actual degrees of belief are directly determined by observation, there is no obvious circularity here. In contrast to the $E = B$ case, this appeal to actual degrees of belief does not trivialise Bayesianism as a normative theory: observationally-set degrees of belief are indeed rational as well as actual degrees of belief, and in any case there remains scope for the normative and the descriptive to come apart with regard to other degrees of belief.

This move to observationally-set credences does come at a cost however: observation trumps other ways of setting degrees of belief and this is not always appropriate. Suppose that an agent knows that she is colour blind, seeing both orange and green as the same hue of green (protanopia). An observation might initially cause a credence of .99 that a jacket is green. If $E = C$ then the agent is rational to believe that the jacket is green to degree .99. The Bayesian should agree with this: it is rational because it is compulsory, having been causally determined by observation. But if $E = C$, the agent is rational to continue with degree of belief .99 that the jacket is green, even after ample time for reflection, as long as that credence remains directly set by observation—i.e., as long as she has not engaged in a conscious process of reflection and deliberation which culminates in setting credence .99 by choice rather than by direct observation. The Bayesian should disagree with this verdict: the agent should know better. She knows that she is colour blind, and this knowledge should lead her to engage in a process of reflection and deliberation after direct observation, and thereby to adopt a more moderate degree of belief in this case. However, the proponent of $E = C$ cannot take this route: the agent's knowledge that she is colour blind does not even count as evidence if that knowledge was obtained by testimony, say, rather than direct observation. Hence the advocate of $E = C$ must hold that such knowledge should not constrain the agent's degrees of belief.

In sum, neither do all observationally-set credences count as evidence (e.g., the observationally-set credence that the jacket is green, after ample time for reflection), nor are all items of evidence observationally-set credences (e.g., the evidence that one is colour blind). $E = C$ fails in both directions.

§6
E=I, Information

All the views of evidence outlined above appeal to belief in one way or another: $E = K$ sees evidence as true qualitative belief with some further authority; $E = B$ sees evidence as full belief; $E = C$ sees evidence as partial belief. Another option is to depart from belief altogether, construing a body of evidence as a body of information:

> we may be concerned not only with what scientists actually believe (or

> believed), but also with other information accessible to the community … it is clear that information which has been collected as a result of human effort but never believed *can* bear a confirmation (or corroboration) relation to hypotheses in which we are interested (Rowbottom, 2014, §2).

As Rowbottom argues, it can be reasonable for an agent to take as evidence the contents of a notebook, as well as the contents of her memory, even if the information in her notebook is not present in her mind and thus not actually believed. Although the agent does not actually believe the contents of her notebook, she may be disposed to believe them, and there is a clear sense in which they may constrain her rational degrees of belief.

One attractive feature of this view is that it offers an account of evidence that applies equally in the case of a group's body of evidence as in the case of an agent's body of evidence. In contrast, if $E = C$ it is by no means obvious how the notion of evidence can apply to a group whose members have incompatible credences.

There are various worries that one might have about this view though. One concern is that while it may not be clear what evidence is, it is, if anything, less clear what information is, with little agreement in the literature as to the nature of information. As Rowbottom acknowledges, it is far from obvious how best to answer the simple question of whether or not information need be true. Hence an account of evidence in terms of information may be less than helpful for the purposes of guiding formal epistemology.

Another concern is that much hinges on how accessible the information must be in order to count as evidence. One might think that if ought implies can, an agent should not be deemed irrational when her degrees of belief neglect information that she has forgotten or lost or that was destroyed, because she cannot access that information. But if $E = I$ then in certain cases she can be deemed irrational. A cook may have forgotten whether Mornay Sauce contains Gruyère cheese and may be deliberating as to how strongly she should believe that proposition. If $E = I$ and that information is in her notebook, then she should fully believe or disbelieve it, because it is part of her evidence, as long as the notebook is accessible to her. But how accessible? She may not have time to look at her notebook; she may have left it at home; she may have lent it to a friend far away. Indeed it seems that many of the propositions about which we deliberate can be answered by delving into information that is accessible in some remote sense—does that mean that we should have firm beliefs about all these propositions? By divorcing evidence from any sort of propositional attitude of the agent, the advocate of $E = I$ has no immediate answer to these questions about accessibility.

§7
E=G, what is rationally Granted

Another option, mooted in Williamson (2010, §1.4.1), is to understand an agent's body of evidence as consisting of everything she takes for granted in her current operating context. (Since we are interested here in *rational* belief we shall use 'granted' in an unqualified way for what is *rationally* granted, saving 'actually granted' for other cases—c.f., §2.)

On the one hand, whatever constitutes evidence must be taken for granted: $E \subseteq G$. This is just the familiar point that one cannot reason *to* anything without presuming something else. Similarly, whatever is taken to be the basis for one's beliefs needs to be granted, at least provisionally. If this were not true—if some item of evidence were open to question in the current context—then Bayesianism would lose its normative force: why should anyone satisfy the constraints imposed by their evidence if the evidence itself is currently up for grabs? In particular, the Presumption principle of §2, which is at the very core of Bayesianism, would fail. As C.I. Lewis notes,

> If what is to confirm the objective belief and thus show it probable, were itself an objective belief and hence no more than probable, then the objective belief to be confirmed would only probably be rendered probable. Thus unless we distinguish the … belief in which experience may render probable, from those presentations and passages of experience which provide this warrant, any citation of evidence for a statement about objective reality, and any mentionable corroboration of it, will become involved in an indefinite regress of the merely probable—or else it will go round in a circle—and the probability will fail to be genuine. (Lewis, 1946, p. 186.)[9]

On the other hand, whatever is granted must be taken as evidence: $G \subseteq E$. If you grant θ then your degrees of belief must satisfy the constraints imposed by θ, for otherwise your degrees of belief are compatible with $\neg\theta$ and you cannot be said to grant θ after all—you might be said to have mixed views about θ, or simply to have suspicions that $\neg\theta$.

In particular, it is clear that what you take for granted should satisfy the Presumption principle: if $\theta \in G$ is expressible and G is consistent with $P_G(\theta) = 1$, then $P_G(\theta) = 1$. For otherwise θ is taken for granted but believed to degree less than one, which is problematic in the sense of Moore's paradox: one cannot be said to grant θ but doubt it at the same time; those propositional attitudes are inconsistent.

In sum, then, $E \subseteq G$ and $G \subseteq E$, so $E = G$. This view has certain immediate advantages over the preceding views. That evidence need not be true is a problem for $E = K$ but not for $E = G$. Indeed, we often do defeasibly grant false propositions and we can be rational to grant false propositions: e.g., we are rational to grant propositions that are presented to us by some sufficiently reliable process, though a small proportion of those propositions will be false. If $E = G$ then there is no need to assume that evidence is articulable in an agent's language, as there is if $E = B$. $E = G$ does not give undue weight to unreliable observations, as does $E = C$. Moreover, accessibility of evidence is less of a problem for $E = G$ than for

[9] Lewis continues, 'if anything is to be probable, then something must be certain' (Lewis, 1946, p. 186), and Jeffrey (1968, p. 36) takes issue with this demand for certainty when he takes as evidence observationally-set credences, which are apparently less than certain (§5). But an observationally-set credence is a credence of the form $\widehat{P}(\theta) = x$ that has been directly caused by observation—the observation of a jacket by candlelight, for example—, where $\widehat{P}$ signifies *actual* degree of belief, as opposed to rational degree of belief P_E. It is thus $\widehat{P}(\theta) = x$ that is in the agent's body of evidence, not the claim that the jacket is orange, nor the act of observation itself. And the agent cannot but grant that $\widehat{P}(\theta) = x$ at the instant after observation, because she cannot but believe θ to degree x, that belief having been caused by the observation. So observationally-set credences do accord with the claim that $E \subseteq G$.

One reason Jeffrey objects to Lewis is that, for evidence to be certain, i.e., to have degree of belief 1, that evidence must be expressible, which is not always the case. Here I do not insist that all evidence be certain, merely that it be granted, and we clearly take things for granted that we cannot articulate. Moreover, the principle of Presumption only requires that evidence be certain where that evidence is expressible. So we can endorse Jeffrey's claim that evidence need not be expressible in the agent's language.

$E = K$ or $E = I$. To make this last point clear, however, we will need to look more closely at granting and norms for granting.

§8
Granting

We take many kinds of things for granted. When arguing, we can grant the premisses of an argument. Granting that there is evil, and that an omniscient, omnipotent, benign god would not permit evil, one can argue that there is no omniscient, omnipotent, benign god. By granting the premisses we can focus on the question of the validity of the argument; if we don't grant the premisses then the soundness of the argument comes into play. We grant observations. I might grant that the jacket I'm looking at is orange, for instance. We grant metaphysical presuppositions. Thus I take for granted that the person writing the last word of this sentence is the same person as the person who wrote the first word of the sentence. We grant scientific theory. One might grant the laws of quantum mechanics when working on quantum cryptography, even if one doesn't know precisely what these laws are. One may also grant the laws of Newtonian mechanics—in the knowledge that they are, strictly speaking, false—when one thinks that it will be fruitful to do so. We grant our assumptions. Giving a talk at an academic conference, for instance, one might assume in the absence of contrary evidence that the audience will understand English. We take people for granted: we grant that those we know well will behave in certain ways. (We also grant favours, wishes and requests, but this is a different sense of the word 'grant' to that used here. There is also a negative connotation to taking someone for granted that goes beyond what is intended here.)

By taking something for granted we presume its truth, taking the question of its truth beyond the context of enquiry.[10] It is worth stressing the context relativity of what is taken for granted. A medical practitioner may rightly take the efficacy of a particular medicine for granted, while an efficacy auditor may be right to call that efficacy into question, and a philosophical sceptic may be right to take very little for granted. What should be taken for granted depends on the agent's current operating context, including the questions that the agent is seeking to answer and the standards that are imposed on her.[11]

Bayesianism offers a permissive account of rational belief: we are rational to believe what we do, given our evidence E, unless one of the norms of Bayesianism is contravened. As we saw in §1, this leaves room for subjectivity, in that a rational belief function P_E need not be uniquely determined by E. Similarly, one can develop a permissive account of rational granting, an account which can admit more than one body of evidence as rational for an agent to adopt, by providing a set of norms by which the agent should abide. In what follows we shall take G to be a set of propositions that are rationally granted, and shall propose norms that G satisfies. Equivalently, we can view G as a set of propositions that are *actually* granted, and the following norms as norms that ought to be satisfied if G is to count as *rational*. G thus plays a role analogous to that of the rational belief function P_E in Bayesian theory. (Here G is a set of propositions. Colloquially, one can take for

---

[10] However, that a proposition is taken for granted does not imply that it is a 'hinge proposition' in Wittgenstein's sense. Propositions that are granted in one context may be open to debate in other contexts, while hinge propositions are 'exempt from doubt' (Wittgenstein, 1969, §341) and 'incontestable' (Wittgenstein, 1969, §655).

[11] See Bratman (1992, §3) for a consideration of other ways in which granting is relative to context.

granted non-propositions—one's spouse, for example. But in so doing one grants certain propositions about one's spouse's behaviour, and it is these propositions that are included in the set G.)

G1: G is consistent.

This norm appeals to a broader notion of consistency than logical consistency. Not only should G not contain any logical contradictions, but it should be satisfiable in principle by the relevant propositional attitudes. Thus $G = \{\theta, P_G(\theta) < 1\}$ is inconsistent because one cannot grant a proposition yet, having granted it, grant that one believes its negation to some extent.[12] This norm might seem trivial: if one cannot have propositional attitudes that satisfy certain constraints, then one need hardly be prohibited from adopting such attitudes. But 'satisfiable in principle' is rather different to 'satisfiable in practice'. In practice, one might grant θ yet less than fully believe φ, not realising that θ and φ are the same proposition. In principle, however, such attitudes are as incompatible as granting θ and granting $\neg\theta$. Norm G1 is intended to deem irrational those attitudes that are incompatible in principle.

G2: G is closed under its consequences.

Similarly, norm G2 appeals to a notion of consequence that is broader than that of logical consequence. For example, one consequence of θ being a member of G is that $P_G(\theta) = 1$, so this latter proposition is also in G if θ is. One can thus formulate G2 as $G = Cn(G)$, where Cn is the relevant supraclassical consequence operator. G2 entails closure under logical consequence, and it is intended in a spirit similar to the assumption of logical omniscience in Bayesian theory: it is an ideal that is sensible in a wide variety of applications, but that needs to be taken with a pinch of salt elsewhere—in mathematics, for instance.

G3: If θ is actually granted and the agent in question cannot help but take θ for granted then $\theta \in G$.

Typically, if one cannot help but grant θ then that is because one's granting θ has been directly caused by some cognitive process such observing or remembering. Such processes are normally reliable—indeed we have evolved to depend upon them because they are reliable—hence it is reasonable to continue to depend upon them. But this norm stands even where the processes are unreliable, such as cases of colour blindness or systematic memory failure. This is because ought implies can (§2): one can hardly be deemed irrational for granting what one cannot help but grant.

After adequate time for deliberation—e.g., deliberation about one's colour blindness—one *can* help but take the propositions in question for granted and the norm no longer applies. The question then arises as to whether any diachronic norms can offer guidance in such a situation. Plausibly,

[12] One might question the legitimacy of this sort of self reference—G containing a proposition that refers to G—and argue that the proposition $P_G(\theta) < 1$ is not eligible for membership in G. But a similar point applies to $G = \{\theta, \widehat{P}(\theta) < 1\}$, where $\widehat{P}$ is actual degree of belief, which makes no reference to G itself. One simply ought not grant a proposition of the form θ *but I don't fully believe it.*

G4: If $\theta \in G$ in some particular operating context, and proposition φ is subsequently granted but the operating context remains the same, then $\theta \in G'$, the new set of granted propositions, unless the retraction of θ resolves some tension.

Here we shall allow φ to be a tautology, to cover the case in which nothing new of substance has been granted but where there has been an opportunity to deliberate about what one grants.

One obvious kind of tension is that in which the new evidence casts doubt on θ. The natural way of explicating this casting of doubt is in terms of degree of belief: we can say that the new evidence casts doubt on θ if $P_{G\star\varphi-\theta}(\theta)$ is significantly lower than $P_{G-\theta}(\theta)$, for some suitable contraction operator $-$. ($G-\theta$, the *contraction* of G by θ, is the result of removing appropriate propositions from G so that θ is no longer granted. On the other hand, $G\star\varphi$, the *revision* of G by φ, is the result of adding and removing appropriate propositions to G to include φ yet remain consistent. We shall defer discussion of the question of which contraction and revision operators are suitable to the end of this section.)

Another kind of tension arises when it is found that $G\star\varphi$ is unnecessarily complex or incoherent, or lacking in some other epistemological virtue. The picture here is analogous to the Quinean notion of web of belief (Quine and Ullian, 1970): changes can be made at various places to the web of what one grants in order to improve its epistemological qualities. Note that G4 allows some propositions to be granted just on the grounds of their contribution to the quality of the body of what one grants, not on the grounds of their probability. For instance, Bayesian accounts sometimes give universal hypotheses probability zero, yet it may be rational to grant these hypotheses just on account of their explanatory power, or the simplicity that they yield (see, e.g., Williamson, 2013, §3).

Another sort of tension arises in the colour-blindness example: I may initially grant that the jacket in front of me is orange—call this θ—, and then, later, come to realise that I may well be mistaken. In this case it is not that *new* evidence has arisen that casts doubt on θ; I previously granted that I was colour blind and it is that which is in tension with θ. It is just that I wasn't previously able to take the colour blindness into account since my granting θ was directly caused by observation. The tension was already there but I wasn't in a position to do anything about it. Now I can take it into account because there has been time for reflection. So θ is to be retracted from G not because $P_{G\star\varphi-\theta}(\theta)$ is significantly lower than $P_{G-\theta}$ (this is a case in which φ is tautologous and $G\star\varphi-\theta = G-\theta$) but because $P_{G-\theta}(\theta)$ is sufficiently low and θ offers no mitigating epistemological virtues in this case. This, then, provides a third way in which a tension can be resolved by retracting θ.

G5: If θ is contentious in the context of enquiry then $\theta \notin G$.

Norm G5 makes explicit the fact that what is appropriate to grant depends to some extent on the context of enquiry. Thus if one is trying to determine whether or not θ is true, one should not simply grant θ: one should not put beyond the context of enquiry what is up for grabs in that enquiry.

One might also propose the converse norm—if θ is *un*contentious in the context of enquiry then $\theta \in G$—, but such a proposal would be rather questionable: in the context of many early navigation problems, it was uncontentious that the earth is flat, but it is not clear that someone who didn't take this claim for granted would be irrational, especially given the paucity of evidence supporting the flat-earth claim.

If $E = G$, norm G5 underwrites the commonly-held view that evidence is uncontentious or uncontroversial: 'Think of your evidence as the propositions that are uncontroversial for you to use in arguing for or against other propositions. They are uncontroversial given your own perspective … They are the bases of argument and inquiry, not the object of them' (Foley, 1993, p. 192). However, what is taken for granted should not simply be relative to the agent's own perspective, but to the general context within which the agent operates. Thus an agent may be wrong to grant something that her peers find contentious, especially if she is engaged in trying to convince her peers or is subject to standards imposed by her peers. And we shall maintain in §9 that if a group is engaged in public deliberation then the group should not take propositions for granted that are contentious to the individuals in the group.

¶ *Decision-Theoretic Granting.* G1–5 are compelling norms of granting, but this list is no doubt incomplete. Moreover, conflicts between norms can arise. For example, by G3 one may end up granting something that makes G inconsistent, contravening G1. While this tension will later be eradicated by G4, there remains a time when G1 is infringed. As the discussion of G2 implies, some of these norms are to be treated literally as standards of *normal* epistemological behaviour, and it should be expected that deviations from the norm can arise in certain circumstances.

Nevertheless, one can try to reconcile some of the conflicting aspects of these norms by moving to a more unified view of rational granting. One such unification—one that accords well with the account of judgement given in §1—appeals to Bayesian decision theory. Suppose that, within some fixed context, an agent is deciding whether or not to grant some proposition θ that may or may not already be taken for granted. If θ is not already in G then the options for the new set G' of granted propositions are $G \star \theta$ (revising G to include θ) or $G - \theta = G$ (leaving G as it is). On the other hand if θ is already in G then the options for G' are $G \star \theta = G$ (leaving G as it is) or $G - \theta$ (retracting θ from G). Arguably,

Decision-Theoretic Granting. $G' = G \star \theta$ if and only if either

- the agent cannot help but grant θ, or
- θ is uncontentious in the current context and the expected utility of her granting θ exceeds that of her not granting θ, relative to the context and relative to $G - \theta - \neg\theta$.

Note that for a contraction operator, $G - \theta - \neg\theta \stackrel{\text{df}}{=} (G - \theta) - \neg\theta = (G - \neg\theta) - \theta$. If $\theta \not\in G$ and θ is consistent with G then $G - \theta - \neg\theta = G$. If $\theta \not\in G$ and θ is inconsistent with G then $G - \theta - \neg\theta = G - \neg\theta$. If $\theta \in G$ then $G - \theta - \neg\theta = G - \theta$.

This account of granting presumes an appropriate utility function and a belief function $P_{\cdot}(\cdot)$ that maps a set of granted propositions and a proposition under consideration to a number in the unit interval. Suppose for example that the agent initially takes G for granted and is deciding whether or not to add to G some proposition θ that is not already granted but is consistent with what is already granted. In this case utilities and probabilities are relative to G. She might formulate the following sort of decision table:

	θ	$\neg\theta$
Grant θ	5	-4
Don't grant θ	-1	3

The expected utility of granting θ is then $5P_G(\theta)-4(1-P_G(\theta))$, which is greater than the expected utility of not granting θ, $-1P_G(\theta)+3(1-P_G(\theta))$, just when $P_G(\theta)>7/13$. On the other hand, if $\theta \not\in G$ were inconsistent with G then assuming those elements of G that are inconsistent with θ would prejudge the question of whether or not to grant θ; therefore one would need to consider probabilities and utilities relative to $G-\neg\theta$, rather than G. Similarly, if θ were already granted then assuming θ would prejudge the question of whether or not to grant θ, so utilities and probabilities are relativised to $G-\theta$ rather than G; in this case $G\star\theta=G$.

One advantage of this decision-theoretic account is that the utility function can encapsulate a multitude of virtues: it can award high utility to those acts of granting that maintain consistency (G1), closure under consequences (G2), coherence, simplicity, unification and so on (G4).

A second advantage of the decision-theoretic account is that one can fruitfully analyse connections between utility and truth. Many propositions θ are such that the utility of granting θ when θ is true is greater than the utility of not granting θ when θ is true; we shall call such propositions *trutile*. Many θ are such that the utility of not granting θ when θ is false is greater than that of granting θ when θ is false; these will be called *faltile*. θ in the above decision table is both trutile and faltile, for example. Note, though, that not all propositions need satisfy these conditions. It might be more useful to grant a universal hypothesis, for instance, than not to grant it, even if it is strictly-speaking false, as long as it admits relatively few counterexamples. This might be because of its simplicity, strength, and/or unifying power in the context of other granted propositions. In which case such a hypothesis is not faltile.

One consequence is that an uncontentious proposition θ should be granted if it is trutile and $P_{G-\theta-\neg\theta}(\theta)=1$. To see this consider a general decision table:

	θ	$\neg\theta$
Grant θ	w	x
Don't grant θ	y	z

If θ is open to deliberation but uncontentious, then by Decision-Theoretic Granting, θ should be granted if the expected utility of granting θ exceeds that of not granting θ:

$$wP_{G-\theta-\neg\theta}(\theta)+xP_{G-\theta-\neg\theta}(\neg\theta)>yP_{G-\theta-\neg\theta}(\theta)+zP_{G-\theta-\neg\theta}(\neg\theta),$$

i.e., if

$$(w-y)P_{G-\theta-\neg\theta}(\theta)>(z-x)P_{G-\theta-\neg\theta}(\neg\theta) \qquad (1)$$

If θ is trutile then the left-hand side of Equation 1 is positive; since $P_{G-\theta-\neg\theta}(\theta)=1$, the right-hand side is zero; hence, θ should be granted. This fact allows one to put beyond the context of enquiry many propositions that are certain on the basis of what one already grants, allowing one to focus one's deliberation on those propositions that are less secure in the sense of being contentious or less certain. Arguably, this consequence captures what is plausible behind the converse of G5, discussed above.

Another consequence of such an analysis is that an uncontentious proposition θ should be granted if it is trutile but not faltile, as long as $P_{G-\theta-\neg\theta}(\theta)>0$. (The left-hand side of Equation 1 is positive and the right-hand side is not.) In the case of the universal hypothesis, it may well be that $P_{G-\theta-\neg\theta}(\theta)=0$, so the left-hand side of Equation 1 is zero. But if θ is *strictly* non-faltile, i.e., $x>z$, then the right-hand side of Equation 1 is negative and one should in any case grant θ.

Note that, by presupposing a belief function P_G as well as a utility function, the second disjunct of the Decision-Theoretic Granting criterion presupposes a previous set G of granted propositions. When it comes to determining an initial set of granted propositions, only the first disjunct can apply: $\theta \in G$ iff the agent cannot help but grant θ. So, before starting to deliberate about what one should grant, the suggestion is that one should begin with what one actually and unavoidably grants; the process of deliberation transforms this set of propositions into one that can have greater overall utility.

¶ *Accessibility of evidence.* Having made some tentative first steps towards a normative account of granting, we can return to main theme of this paper, namely the nature of evidence. The question remains as to whether, under the account provided in this paper, an agent's evidence is accessible to her. If not, she cannot be expected to use her evidence to constrain her degrees of belief via the Bayesian norms.

But if, as argued, $E = G$, then it should be clear that an agent's body of evidence is in principle accessible to her, in a way that it couldn't be if $E = K$, say.

Note first that there are two chief ways in which one might demand that evidence be accessible. As far as the Bayesian is concerned, in order to deliberate about how strongly to believe various propositions of interest, the agent does not need to be able to articulate her body of evidence in propositional form—she merely needs to be able to identify the constraints that evidence imposes on her degrees of belief via the Bayesian norms. On the other hand, in order to deliberate about whether or not to grant a proposition θ, she needs a belief function and utility function relative to a prior body of evidence E. This second task requires perhaps a stronger sense of accessibility, since not only does it require accessibility in the Bayesian sense (in order to determine the belief function), but also the agent needs to be able to assess epistemic qualities of evidence, such as its consistency or coherence, in order to determine the utility of adding θ to her body of evidence. In neither case, though, is there a requirement that the agent be able to articulate each and every item of evidence.

In §3 we saw why, if $E = K$, evidence is not sufficiently accessible: the agent cannot be expected to correctly identify what she knows, because for any proposition that she thinks she knows she might be mistaken about its truth or authority, for reasons beyond her ken. If the agent is wrong about what she knows, she may be wrong about the constraints that her evidence imposes on her belief function. And she may be wrong about the epistemic qualities of her evidence and the utility of adding a further proposition to her evidence. Hence $E = K$ can satisfy neither the Bayesian demands on evidence nor the demands of a decision-theoretic normative account of evidence analogous to that given above.

If, on the other hand, $E = G$, as argued in §7, then the agent's evidence is sufficiently accessible. The agent can reasonably be expected to determine what she ought to grant, since, under the above decision-theoretic account, this depends only on what she unavoidably grants, what is contentious in the current context, her belief function relative to what she currently grants, and her utility function; these are all in principle accessible to the Bayesian agent who has time to deliberate and introspect.

¶ *Revision and Contraction.* For the sake of completeness, we shall now turn to the question of which contraction and revision operators are suited to the needs of the Bayesian. Consider some given propositional attitude that involves somehow *endorsing* the propositions in question—an attitude such as judging, granting, believing, accepting, predicting etc. Fix E for the moment and suppose A_P is the set of propositions that an agent with Bayesian belief function $P = P_E(\cdot)$ currently endorses, for that particular endorsing attitude. (We need not assume that this set of propositions depends only on the agent's belief function; it may depend on utilities as well, for instance, or on the question being asked.) A *question* Ψ asks which member of a partition $\Psi = \{\psi_1, \ldots, \psi_k\}$ of propositions is true. An agent might try to answer this question by endorsing some proposition from this partition or by endorsing some logically complex proposition built up from propositions in Ψ.

Lin and Kelly (2012) put forward the following desiderata that one might expect an endorsed set of propositions A_P and a revision operator $\star$ to satisfy. (While they focus on acceptance as the endorsing attitude in question, their discussion can be generalised to other endorsing attitudes.)

Non-Skeptical. Each complete answer ψ_i to Ψ is endorsed over some open neighbourhood of probability functions. This precludes A_P such that ψ_i is endorsed if and only if $P(\psi_i) = 1$.

Non-Opinionated. There is some open subset of probability functions over which a disjunction of members of Ψ is endorsed.

Consistent. For all P, A_P is consistent.

Corner Monotone. If ψ_i is endorsed at P then it is endorsed for every convex combination of P and P^i, where P^i is defined as the probability function that gives probability 1 to ψ_i.

Track Conditioning. $A_P \star \theta = A_{P(\cdot|\theta)}$ if $P(\theta) > 0$. Thus if one revises the set of endorsed propositions on learning θ, the resulting set of propositions should tally with that that would have been endorsed had the agent conditionalised on θ.

Note that not all Bayesians accept Bayesian conditionalisation as a *universal* rule of updating: some argue that conditionalisation fails in certain cases (Earman, 1992, p. 51; Howson, 1997; Williamson, 2011). Nevertheless, all Bayesians would want belief updating to agree with conditionalisation in other cases—Bayesians would not want a revision operator that forced disagreement with conditionalisation even in non-pathological cases.

It turns out that when the question has at least three answers, the standard approach to belief revision—*AGM revision* (Alchourrón et al., 1985)—doesn't satisfy these desiderata (Lin and Kelly, 2012, Corollary 1). This is because AGM revision satisfies the following condition which prevents it from both tracking conditioning and being non-opinionated:

Accretion. If θ is consistent with A_P then $A_P \star \theta = Cn(A_P \cup \{\theta\})$, the closure of $A_P \cup \{\theta\}$ under its consequences.

On the other hand, certain implementations of *Shoham revision* (Shoham, 1987) do satisfy these desiderata. For some given threshold t_{ij}, define:

$$\psi_i \prec \psi_j \Leftrightarrow \frac{P(\psi_i)}{P(\psi_j)} > t_{ij}.$$

The relation $\prec$ orders propositions according to relative strength of belief. For example, if $t_{ij} = 1$ and $\psi_i \prec \psi_j$ then ψ_i is more strongly believed than ψ_j. When $\prec$ is a partial order, the Shoham revision $A_P \star \theta$ is defined to be the disjunction of those of the answers ψ_i consistent with θ that are minimal (i.e., most strongly believed) with respect to $\prec$. This Shoham revision operator does satisfy the desiderata (Lin and Kelly, 2012, Theorem 3).

Since the agent faces a decision as to whether or not to endorse a given proposition, the Bayesian might be inclined towards a decision-theoretic account of endorsement, such as the decision-theoretic approach to judgement of §1 and the decision-theoretic approach to granting presented above. Such an account can induce a suitable Shoham revision operator, as we shall now see. Write $\psi_i^{\checkmark}$ for the option *endorse* ψ_i and $\psi_i^{\times}$ for *don't endorse* ψ_i. Suppose the utility of a type II error is independent of the answer to the question, $U(\psi_i^{\checkmark}|\neg\psi_i) = U(\psi_j^{\checkmark}|\neg\psi_j)$ for all i, j. As a convention, let that constant utility be zero (utilities can be translated by a constant without affecting which option maximises expected utility). Then let $\psi_i \prec \psi_j$ iff $\psi_i^{\checkmark}$ has higher expected utility than $\psi_j^{\checkmark}$. I.e., iff

$$P(\psi_i)U(\psi_i^{\checkmark}|\psi_i) + P(\neg\psi_i)U(\psi_i^{\checkmark}|\neg\psi_i) > P(\psi_j)U(\psi_j^{\checkmark}|\psi_j) + P(\neg\psi_j)U(\psi_j^{\checkmark}|\neg\psi_j).$$

I.e., iff,

$$\frac{P(\psi_i)}{P(\psi_j)} > \frac{U(\psi_j^{\checkmark}|\psi_j)}{U(\psi_i^{\checkmark}|\psi_i)}.$$

This is a partial order if $U(\psi_i^{\checkmark}|\psi_i) > 0$ for $i = 1, \ldots, k$. Such a partial order is of the form required to induce a Shoham revision operator that satisfies the desiderata outlined above. Note that the utility function U will depend on the endorsement attitude in question, i.e., on the particular interpretation of A_P. Note also that if more than one answer ψ_i is minimal (i.e., if there is more than one answer whose endorsement would have maximum expected utility) then one should endorse the disjunction of those answers.

In sum, then, a decision-theoretic account of endorsement can be used to generate a Shoham revision operator. This in turn yields an operator $\star$ for revising a set of endorsed propositions to accommodate a newly endorsed proposition θ.

Having isolated one or more appropriate revision operators, the question then arises as to which constraints *contraction* operators should satisfy if they are to be suited to the need of Bayesian epistemology. This question is typically answered indirectly by appealing to the *Harper identity*,

$$A_P - \theta = (A_P \star \neg\theta) \cap A_P.$$

The idea here is that a contraction operator is deemed appropriate just in case it is related to some appropriate revision operator via the Harper identity.

§9
Merging Evidence

Thus far we have argued for a particular notion of evidence—namely $E = G$, i.e., evidence is what is rationally granted—and have made some tentative suggestions as to norms on rational granting. In this section we shall return to the question posed in §1: how should the evidence of individual agents be merged for the purposes of public deliberation and judgement?

¶ *Norms for merging evidence.* *Belief merging* is a formal approach to the merging 'belief bases' or 'knowledge sets' (see, e.g., Konieczny and Pino Pérez, 2011). One might thus presume that this formalism would be appropriate for merging evidence, if only $E = B$ or $E = K$. Having argued instead that $E = G$, the question arises as to whether the belief merging formalism remains appropriate for merging evidence. In Williamson (2009) I suggested that it does remain the appropriate formalism. Here I would like to re-evaluate that claim. We will see that in fact the axioms of belief merging need to be altered somewhat if they are to be applied to merging evidence.

There are two sets of norms for belief merging in the literature: one for merging in the presence of *integrity constraints*, constraints that must be satisfied by the merged set of propositions, and another for merging in the absence of integrity constraints. For our purposes it will suffice to consider the latter, simpler set of norms for merging in the absence of integrity constraints.

Consider a multiset of evidence bases $\mathscr{E} = \{E_1, \ldots, E_n\}$. Each E_i is a consistent set of propositions (represented in, say, propositional logic) that constitutes the evidence base of individual i. The notation E_i is customarily used to denote either this set of propositions or the conjunction of those propositions. Denote by $\mathscr{E} \sqcup \mathscr{F}$ the multiset $\{E_1, \ldots, E_n, F_1, \ldots, F_m\}$. Two such multisets are *equivalent*, $\mathscr{E} \equiv \mathscr{F}$, if there is a bijection f from $\mathscr{E}$ to $\mathscr{F}$ such that the corresponding evidence bases are logically equivalent, $E_i \equiv f(E_i)$. We will use the symbol Δ to denote a merging operator, i.e., a mapping from a multiset of individual evidence bases to their merger, the group evidence base (again, thought of as either a set of propositions or the conjunction of these propositions). Konieczny and Pino Pérez (1998) proposed the following norms of merging:

M1. $\Delta\mathscr{E}$ is consistent.

M2. If $\bigwedge\mathscr{E}$ is consistent then $\Delta\mathscr{E} = \bigwedge\mathscr{E}$.

M3. If $\mathscr{E} \equiv \mathscr{F}$ then $\Delta\mathscr{E} \equiv \Delta\mathscr{F}$.

M4. If $E_i \wedge E_j$ is inconsistent then $\Delta\{E_i, E_j\} \not\models E_i$.

M5. $\Delta\mathscr{E} \wedge \Delta\mathscr{F} \models \Delta(\mathscr{E} \sqcup \mathscr{F})$.

M6. If $\Delta\mathscr{E} \wedge \Delta\mathscr{F}$ is consistent then $\Delta(\mathscr{E} \sqcup \mathscr{F}) \models \Delta\mathscr{E} \wedge \Delta\mathscr{F}$.

Let us consider these norms in turn. We shall assume a fixed context of enquiry. In particular, we shall assume that the context of enquiry of the group is the same as that of each of its individuals. Otherwise, the evidence of certain individuals may be irrelevant to that of the group.

M1 requires that the group evidence base be consistent. If evidence is what is rationally granted, this is reasonable—indeed, norm G1 on rational granting is itself a consistency norm.

M2 says that if the individual evidence bases are mutually consistent then the merged evidence base should simply conjoin the individual evidence bases. This is too strong: a group is not rationally compelled to take for granted anything that anyone grants. While it may be the case that an individual has observed something that the other individuals in the group were not in a position to observe, and in such a case it is reasonable for the group evidence to include that item of evidence, in other situations the individuals may disagree as to whether a particular proposition qualifies as evidence. Thus if $\theta \in E_i$ but $\theta \notin E_j$, that may be because individual j has considered θ and decided not to grant it, and for good reason. This renders θ open to deliberation. Thus there is no normative imperative for the group to simply take θ for granted.

Although M2 is not plausible if $E = G$, certain weakenings of M2 are plausible. Thus we might replace M2 with:

M2a. If $E_i \models \theta$ for all i then $\Delta\mathscr{E} \models \theta$.

M2b. $\bigwedge\mathscr{E} \models \Delta\mathscr{E}$.

M2a says that the group should grant anything that everyone grants, while M2b says that the group evidence base should not imply anything that no individual grants. Arguably the group should grant everything that everyone grants whether or not the individual evidence bases are mutually consistent: since each individual evidence base is consistent, all such propositions are uncontroversial. For a set G of sentences, let $[G]$ be a canonical finite evidence base for G (e.g., take $[G]$ to contain only sentences in disjunctive normal form). Then, M2a and M2b are equivalent to:

M2.* $\bigwedge\mathscr{E} \models \Delta\mathscr{E} \models [\bigcap_{i=1}^{n} Cn(E_i)]$.

One might worry that the problem of spurious unanimity (Mongin, 2005) is a problem for M2*. In this sort of situation, everyone grants θ but for different and mutually inconsistent reasons. M2* would require that θ be granted by the group as a whole, even though the group fails to grant any of the grounds for θ, which are controversial. Is this consequence of M2* undesirable? It would only seem so, if one were to maintain that a group should grant the grounds for any proposition that the group grants. But it would be hard to maintain such a principle, for two reasons. First, it introduces a problem of justificatory regress. There would be no finite evidence base that could be adopted by the group, unless these propositions were themselves self-justifying or mutually justifying. While the foundationalist might accept the first option, justificatory autonomy, and the coherentist might condone the second, justificatory circularity, neither option would appeal to the infinitist who holds that chains of justification never terminate. Second, it seems too demanding to impose a principle on the group that none of its constituent individuals is likely to satisfy. If the evidence bases of the individuals are not closed under justification, but their merger is to be closed under justification, then the justification must be plucked from thin air, as it were. Arguably, then, spurious unanimity does not provide a reason to reject M2*.

M3 says that if $\mathscr{E}$ and $\mathscr{F}$ are equivalent then $\Delta\mathscr{E}$ and $\Delta\mathscr{F}$ are logically equivalent; this does seem reasonable, as there is no further contextual information that distinguishes $\mathscr{E}$ and $\mathscr{F}$. Similarly for M4, which says that one individual should not trump another where they disagree. This seems plausible, not on the grounds of fairness to the individuals, as suggested by Konieczny and Pino Pérez (1998)—we are concerned with determining the best group evidence, not with treating individuals equitably—, but on the grounds that, since it is the evidence itself that is in question, there is nothing yet taken for granted that can tell us who should trump whom. Note that this condition does not imply that, where the evidence of two subgroups is inconsistent, the evidence of the group as a whole should not coincide with that of one of the subgroups. For example, if $\mathscr{E} = \{\{\theta\}\}$ and $\mathscr{F} = \{\{\theta,\varphi\},\{\theta,\neg\varphi\}\}$ then it is reasonable that $\Delta(\mathscr{E} \sqcup \mathscr{F}) \models \bigwedge\mathscr{E}$ although $\bigwedge\mathscr{E} \wedge \bigwedge\mathscr{F}$ is inconsistent. Indeed this conclusion is underwritten by M2a.

One might suggest that if a large group of individuals all take θ as evidence but for a single dissenter granting $\neg\theta$, then the majority should prevail and the group as a whole should grant θ. While such a suggestion is permitted by M4, it should arguably be rejected in any case. This is because this sort of disagreement is substantial enough to render θ contentious and thus not admissible for granting by the group. In order to rule out this sort of scenario, we need a strengthening of M4:

*M4**. If $E_i \wedge E_j$ is inconsistent then $\Delta\{E_i,\ldots,E_i,E_j\} \not\models E_i$.

(A weaker sort of majority rule is more plausible: if the majority grant θ but the minority do not grant θ, then θ may not be contentious and it can be appropriate for the group to grant θ. This form of majority rule remains compatible with M4*.)

M5 says that merging the evidence of two subgroups should not yield propositions that are not implied by the merged evidence of one or other subgroup. This seems fairly intuitive: it is hard to argue in the absence of further evidence that the group as a whole should take something for granted that is logically stronger than what the two subgroups grant individually, when what they grant individually is simply pooled. Note that as long as $\Delta\{E_i\} = E_i$ (which is implied by M2a), M5 implies M2b.

M6 says that the group as a whole should grant everything implied by what two subgroups grant individually, when pooled and consistent. This seems too strong for similar reasons to those invoked in the discussion of M2: a group is not compelled to grant everything granted by each subgroup, because the subgroups in question may disagree as to whether or not to grant some proposition, rendering that proposition contentious. Plausibly, though, M6 can be weakened to:

*M6**. If $\Delta\mathscr{E} \models \theta$ and $\Delta\mathscr{F} \models \theta$ then $\Delta(\mathscr{E} \sqcup \mathscr{F}) \models \theta$.

Note that as long as $\Delta\{E_i\} = E_i$, M6* implies M2a, as can be seen by induction on n.

Taking into account the suggested modifications, we are thus left with the following list of norms:

M1. $\Delta\mathscr{E}$ is consistent.

*M2**. $\bigwedge\mathscr{E} \models \Delta\mathscr{E} \models [\bigcap_{i=1}^{n} Cn(E_i)]$.

M3. If $\mathscr{E} \equiv \mathscr{F}$ then $\Delta\mathscr{E} \equiv \Delta\mathscr{F}$.

*M4**. If $E_i \wedge E_j$ is inconsistent then $\Delta\{E_i,\ldots,E_i,E_j\} \not\models E_i$.

M5. $\Delta\mathscr{E} \wedge \Delta\mathscr{F} \models \Delta(\mathscr{E} \sqcup \mathscr{F})$.

*M6**. If $\Delta\mathscr{E} \models \theta$ and $\Delta\mathscr{F} \models \theta$ then $\Delta(\mathscr{E} \sqcup \mathscr{F}) \models \theta$.

An example of a merging operator that satisfies these norms is $\Delta\mathscr{E} \stackrel{\text{df}}{=} [\bigcap_{i=1}^{n} Cn(E_i)]$ (see Proposition 1 in the Appendix). According to this merging operator, the group only grants everything that each member of the group grants.

Interestingly the following merging operator satisfies the first five norms but not M6* (Proposition 2, Appendix): $\Delta\mathscr{E} \stackrel{\text{df}}{=} \bigvee\{\bigwedge\mathscr{G} : \mathscr{G} \in \hat{\mathscr{P}}\mathscr{E}\}$, where $\hat{\mathscr{P}}\mathscr{E}$ is the set of maximal consistent sub-multisets of $\mathscr{E}$. According to this merging operator, the group grants the disjunction of what is granted by maximal consistent subgroups. Moreover, another plausible merging operator fails M5 as well as M6* (see below). In the light of this, one might take the view that M5 and M6* are too strong. We shall view the first four norms as the core norms of evidence merging, with the addition of M5 and M6* characterising what might be called *merging with sub-group conformity*.

¶ *Contentious propositions.* To say much more about norms for merging evidence, one would need more to go on than just the individual evidence bases $E_1, \ldots, E_n$. Let us suppose we also know the set $C_{\mathscr{E}}$ of propositions that are contentious for the group. (For example, if the agents all conform to the decision-theoretic account of granting outlined above, and one knows each individual's belief function P_i and utility function U_i then one can determine, for any particular proposition θ, how close the individual is to rationally granting θ. This last can be measured by the difference between the individual's expected utility of granting θ and her expected utility of not granting θ, i.e., $E_{P_i}U_i(\theta^{\checkmark}) - E_{P_i}U_i(\theta^{\times})$. If this is strongly negative for some individual, then that is a sense in which θ is contentious for that individual. Given some small negative threshold τ, say that θ is τ-*uncontentious* if $E_{P_i}U_i(\theta^{\checkmark}) - E_{P_i}U_i(\theta^{\times}) > \tau$ for $i = 1, \ldots, k$, and let $C_{\mathscr{E}} = C_{\mathscr{E}}^{\tau}$ be the set of propositions that are τ-contentious.)

Presumably $C_{\mathscr{E}}$ will satisfy the following conditions:

C1: If $E_i \models \neg\theta$ for some i then $\theta \in C_{\mathscr{E}}$.

C2: If $E_i \models \theta$ for all i then $\theta \notin C_{\mathscr{E}}$.

C3: If $\theta \equiv \varphi$ then $\theta \in C_{\mathscr{E}}$ iff $\varphi \in C_{\mathscr{E}}$.

C4: $C_{\mathscr{E}} \cup C_{\mathscr{F}} \subseteq C_{\mathscr{E} \sqcup \mathscr{F}}$.

The extra information embodied in $C_{\mathscr{E}}$ can be taken into account in the norms for merging. In particular, we can formulate:

M2c. If $E_i \models \theta$ for all i and $\theta \notin C_{\mathscr{E}}$, then $\Delta\mathscr{E} \models \theta$.

M6c. If $\Delta\mathscr{E} \models \theta, \Delta\mathscr{F} \models \theta$ and $\theta \notin C_{\mathscr{E} \sqcup \mathscr{F}}$, then $\Delta(\mathscr{E} \sqcup \mathscr{F}) \models \theta$.

Note that given C2, M2c is equivalent to M2a. Hence there is no need to replace M2*.

This extra information also motivates a further protocol for merging evidence. Let $\Delta\mathscr{E} \stackrel{\text{df}}{=} [\{\theta \notin C_{\mathscr{E}} : \text{some } E_i \models \theta\}]$. According to this merging operator, the group grants any uncontentious proposition that is granted by some individual in the group. This operator satisfies M1–5, as well as M6c (Proposition 3, Appendix).

¶ *The Social Entropy Process.* Wilmers (2010) offers the following alternative account of public deliberation by merging evidence. First, let $\mathbb{E}_i$ be the set of probability functions that satisfy constraints imposed by evidence base E_i, for $i = 1,\ldots,n$. Attention is restricted to the case in which all the $\mathbb{E}_i$ are closed and convex. Then simultaneously determine $P, P_1 \in \mathbb{E}_1, \ldots, P_n \in \mathbb{E}_n$ so as to minimise the sum

$$\sum_{i=1}^{n} \sum_{\omega} P(\omega) \log \frac{P(\omega)}{P_i(\omega)}$$

of the Kullback-Leibler divergences of P from the P_i. If P is not uniquely determined, but merely constrained to lie in a set $\mathbb{E}$ of probability functions, then choose $P \in \mathbb{E}$ that has maximum entropy. The resulting function P is the group's belief function. Wilmers (2010) justifies this procedure, called the ***social entropy process***, on the grounds that it satisfies a number of desiderata that, he maintains, a public deliberation procedure should satisfy.

There are two main points of difference between the social entropy process and the approach of this paper. First, the social entropy process was put forward as a way of generalising the maximum entropy principle (Jaynes, 1957), which is often applied in the context of individual deliberation, to public deliberation.[13] The maximum entropy principle is core to some versions of objective Bayesianism (e.g., Williamson, 2010), but is eschewed by other kinds of Bayesianism. In contrast, the methods of this paper are neutral regarding the underlying view of Bayesianism.

Second, the evidence bases E_i are not sets of propositions formulated in propositional logic. In fact as Wilmers acknowledges, the social entropy process does not handle propositional evidence very well: if one individual has θ as evidence—i.e., gives probability 0 to $\neg\theta$—then the group as a whole must take θ as evidence; hence the process cannot cope with the situation in which one individual has evidence θ but another individual has evidence $\neg\theta$. So the evidence bases must instead be thought of as attaching non-extreme probabilities to propositions, e.g., to contain statements of the form θ^X where X is a closed interval of probabilities or a single probability that is not 0 or 1. Wilmers (2010, §1) is clear that these probabilities are credences, so the view of evidence behind the social entropy process is closer to the $E = C$ view criticised in §5 than to the $E = G$ view advocated in §7, with the further restriction that the credences should not be the extreme values 0 or 1.

But there is an important sense in which the social entropy process does fit the general approach of this paper. The group belief function P is determined from the individuals' evidence bases E_i rather than their belief functions—in line with the discussion of public deliberation in §1. Moreover, the set $\mathbb{E}$ of probability functions obtained at the first step of the social entropy process can be thought of as representing the group evidence base. Thought of in that way, the social entropy process induces a merging operator that satisfies versions of the original merging norms M1–6 (Adamcik and Wilmers, 2014).

We saw above that some of the norms M1–6 are overly restrictive as norms of evidence merger, and need to be changed. Hence the social entropy process inherits these restrictive characteristics. Suppose, for instance, that in a large group, the evidence bases of all individuals but one attach probability 0.9 to θ. Then the evidence of the majority must hold sway in the group and will swamp the evidence of a dissenter who has a lower probability for θ. In such a case the social entropy

[13] Kern-Isberner and Rödder (2004) offer another approach with a similar goal.

process *forces* a group belief in θ that is close to 0.9. In contrast, the evidence-merger process advocated in this paper can deem substantial disagreement about θ enough to render θ contentious and open to deliberation, in which case there is no constraint imposed on the group belief function that forces a strong belief in θ. Thus one can argue that the evidence of dissenters can be given too little weight under the social entropy process.

On the other hand there are cases in which the evidence of a lone dissenter can be given too much weight. Suppose that in a large group the evidence bases of all individuals attach the interval $[0.3, 0.5]$ to some atomic proposition θ. Then the social entropy process would say that the group degree of belief in θ should be 0.5. Now the group is joined by a new individual who attaches probability $x < 0.3$ to θ. The presence of this outlier forces the group evidence base to give θ a fixed probability that is lower than 0.3. (Hence the group degree of belief in θ moves from 0.5 to below 0.3, however large the original group.) As in the previous example, this seems too restrictive: the evidence of the group is very tightly constrained when there is a lone dissenter. But in this case, the dissenter's opinion perhaps leads to too much movement away from the opinion of the majority. Indeed, by taking x to be small enough, the opinion of the group can be forced to be arbitrarily close to zero.

This sort of case can also arise when the individual evidence bases are mutually consistent. Suppose for instance that all individuals use data from a sample to determine a confidence interval estimate of the probability of θ. Almost all individuals operate at the 99% level to attach an interval of $[0.3, 0.5]$ to θ. But a lone dissenter is prepared to attach a 95% confidence interval of $[0.36, 0.44]$ to θ. Merger norm M2 (and thus the social entropy process) would take the group evidence to be characterised by the narrower interval, since that interval represents the conjunction of the individuals' mutually consistent evidence bases. But this would give too much weight to the lone dissenter: it is precisely the width of the confidence interval that is contentious here. Arguably one should take the wider interval as the group evidence in this case, as it is uncontentious to all that the probability of θ is within that interval.

Note that this last example is merely one instance of a very general problem. We often appeal to confidence or reliability considerations when deciding what to take as our evidence. Moreover, these confidence thresholds are often contentious. M2 would require that, when the individuals' evidence bases are mutually consistent, the group evidence base should be determined by the more specific evidence, which tends to be the evidence yielded by the confidence thresholds that are most lax. But this is perverse—arguably in such cases the group evidence should be the evidence that all the individuals are prepared to endorse.

In sum, while the social entropy process is an interesting proposal that fits with the general approach to public deliberation outlined in §1, it is aligned with the $E = C$ view of evidence and it conforms to the standard merger norms M1–6—two considerations that leave it open to criticism.

An alternative approach to handling evidence bases which can contain information of the form θ^X, where θ is a proposition and X is an interval set of probabilities attaching to θ, is to take the convex hull of those sets of probability functions determined by maximal consistent subsets of evidence bases (see, e.g., Williamson, 2010, §7.3). In this case $\mathbb{E} = \langle\{P : P \text{ satisfies some } \mathscr{G} \in \hat{\mathscr{P}}\mathscr{E}\}\rangle$, where $\langle\cdot\rangle$ is the convex hull operator, and $\Delta\mathscr{E} \stackrel{\text{df}}{=} \{\theta^X : P(\theta) \in X \text{ for all } P \in \mathbb{E}\}$. This merging operator satisfies the core norms for evidence merger outlined above (Proposition 4, Appendix). Moreover, this approach is quite compatible with $E = G$, particularly if one interprets the probabilities as non-epistemic—i.e., if the evidence is construed as evidence of chances, rather than as credences set by observation.

§10
Summary

We have seen in this paper that a natural Bayesian account of public deliberation and judgement demands that evidence be merged (§1). But before one can say how evidence should be merged, one needs a better understanding of what evidence is. Bayesianism places certain demands on evidence (§2), and evidence cannot be what is known (§3), what is fully believed (§4), observationally set credence (§5), or information (§6). This paper has argued instead that evidence is what is rationally taken for granted (§7), and has attempted to specify some norms of granting (§8). If this claim is correct then the standard axioms of merging should be altered if they are to be applied to merging evidence (§9).

Of course the question remains as to whether we can live up to the demands placed on us by such an account of public deliberation and judgement. The reader might be sceptical as to whether this sort of account is relevant to political situations, for instance, where an individual's main aim may be to thwart opponents and where evidence can be somewhat sidelined. However, where there is genuine common interest and where evidence can be made explicit—e.g., in certain scientific committees, drug approval committees, legal committees and in multi-agent systems in AI—there is clearly some scope to apply an account of public deliberation and judgement that proceeds by merging evidence.

Acknowledgements

This research was conducted as a part of the project *From objective Bayesian epistemology to inductive logic* supported by the UK Arts and Humanities Research Council and was a contribution to the *Rationality and decision* network supported by the NWO. I am grateful to Hanti Lin, Christian List, Kevin Kelly, Timothy Williamson, George Wilmers and an anonymous referee for very helpful comments.

Appendix: Examples of merging operators

Proposition 1. *Norms M1, M2*, M3, M4*, M5 and M6* are satisfied by the merging operator* $\Delta\mathscr{E} = [\bigcap_{i=1}^n Cn(E_i)]$.

Proof:

M1. $\Delta\mathscr{E}$ is consistent because each E_i is consistent.

M2.* This holds because $\bigwedge\mathscr{E} \models [\bigcap_{i=1}^n Cn(E_i)]$.

M3. If $\mathscr{E} \equiv \mathscr{F}$ then $\bigcap_{i=1}^n Cn(E_i) = \bigcap_{i=1}^n Cn(F_i)$ so $\Delta\mathscr{E} = \Delta\mathscr{F}$.

M4.* If $E_i \wedge E_j$ is inconsistent then $Cn(E_i) \cap \cdots \cap Cn(E_i) \cap Cn(E_j) = Cn(E_i) \cap Cn(E_j) \subset Cn(E_i)$. So there is a $\theta \in Cn(E_i)$ such that $\Delta\{E_i, \ldots, E_i, E_j\} \not\models \theta$.

M5. $\Delta\mathscr{E} \wedge \Delta\mathscr{F} = [\bigcap_{i=1}^n Cn(E_i)] \wedge [\bigcap_{i=1}^m Cn(F_i)] \models [\bigcap_{i=1}^n Cn(E_i)] \models [\bigcap_{i=1}^n Cn(E_i) \cap \bigcap_{i=1}^m Cn(F_i)] = \Delta(\mathscr{E} \sqcup \mathscr{F})$.

*M6**. If $\Delta\mathscr{E} \models \theta$ and $\Delta\mathscr{F} \models \theta$ then $\theta \in \bigcap_{i=1}^{n} Cn(E_i) \cap \bigcap_{i=1}^{m} Cn(F_i)$ so $\Delta(\mathscr{E} \sqcup \mathscr{F}) \models \theta$.

□

Proposition 2. *Norms M1, M2*, M3, M4*, M5 are satisfied by the merging operator* $\Delta\mathscr{E} = \bigvee\{\bigwedge\mathscr{G} : \mathscr{G} \in \hat{\mathscr{P}}\mathscr{E}\}$, *but M6* is not satisfied.*

Proof:

M1. A disjunction of consistent propositions is consistent.

*M2**. $\bigwedge\mathscr{E} \models \bigwedge\mathscr{G}$ for each $\mathscr{G} \in \hat{\mathscr{P}}\mathscr{E}$ so $\bigwedge\mathscr{E} \models \Delta\mathscr{E}$. Now each $E_i \models \bigcap_{i=1}^{n} Cn(E_i)$ so $\bigwedge\mathscr{G} \models \bigcap_{i=1}^{n} Cn(E_i)$ for each $\mathscr{G} \in \hat{\mathscr{P}}\mathscr{E}$ and $\Delta\mathscr{E} \models [\bigcap_{i=1}^{n} Cn(E_i)]$.

M3. If $\mathscr{E} \equiv \mathscr{F}$ then $\hat{\mathscr{P}}\mathscr{E} \equiv \hat{\mathscr{P}}\mathscr{F}$, so $\Delta\mathscr{E} \equiv \Delta\mathscr{F}$.

*M4**. If $E_i \wedge E_j$ is inconsistent then $\hat{\mathscr{P}}\{E_i, \ldots, E_i, E_j\} = \{\{E_i, \ldots, E_i\}, \{E_j\}\}$. Hence, $\Delta\{E_i, \ldots, E_i, E_j\} = E_i \vee E_j \not\models E_i$.

M5. If $\mathscr{G} \in \hat{\mathscr{P}}(\mathscr{E} \sqcup \mathscr{F})$ then $\mathscr{G} = \mathscr{E}' \sqcup \mathscr{F}'$ for some consistent sub-multisets $\mathscr{E}'$ of $\mathscr{E}$ and $\mathscr{F}'$ of $\mathscr{F}$, one or other of which might be empty. $\Delta(\mathscr{E} \sqcup \mathscr{F})$ is a disjunction of such $\bigwedge\mathscr{G}$. Now for each such $\mathscr{G}$ there are maximal consistent sub-multisets $\mathscr{E}''$ of $\mathscr{E}$ and $\mathscr{F}''$ of $\mathscr{F}$ such that $\bigwedge\mathscr{E}'' \models \bigwedge\mathscr{E}'$ and $\bigwedge\mathscr{F}'' \models \bigwedge\mathscr{F}'$. So each of the disjuncts of $\Delta(\mathscr{E} \sqcup \mathscr{F})$ is logically implied by a disjunct of $\bigvee\{\bigwedge\mathscr{E}'' \wedge \bigwedge\mathscr{F}'' : \mathscr{E}'' \in \hat{\mathscr{P}}\mathscr{E}, \mathscr{F}'' \in \hat{\mathscr{P}}\mathscr{F}\}$. But $\Delta\mathscr{E} \wedge \Delta\mathscr{F} \equiv \bigvee\{\bigwedge\mathscr{E}'' \wedge \bigwedge\mathscr{F}'' : \mathscr{E}'' \in \hat{\mathscr{P}}\mathscr{E}, \mathscr{F}'' \in \hat{\mathscr{P}}\mathscr{F}\}$, so $\Delta\mathscr{E} \wedge \Delta\mathscr{F} \models \Delta(\mathscr{E} \sqcup \mathscr{F})$.

To see that M6* is not satisfied, take $\mathscr{E} = \{\{\theta, \psi_1\}, \{\varphi, \psi_2\}\}$ and $\mathscr{F} = \{\{\neg\theta, \psi_2\}, \{\neg\varphi, \psi_1\}\}$; then $\hat{\mathscr{P}}\mathscr{E} = \{\mathscr{E}\}$ and $\hat{\mathscr{P}}\mathscr{F} = \{\mathscr{F}\}$ so $\Delta\mathscr{E} \models \psi_1 \wedge \psi_2$ and $\Delta\mathscr{F} \models \psi_1 \wedge \psi_2$, but $\Delta(\mathscr{E} \sqcup \mathscr{F}) \not\models \psi_1 \wedge \psi_2$ because $\hat{\mathscr{P}}(\mathscr{E} \sqcup \mathscr{F}) = \{\mathscr{E}, \mathscr{F}, \{\{\theta, \psi_1\}, \{\neg\varphi, \psi_1\}\}, \{\{\neg\theta, \psi_2\}, \{\varphi, \psi_2\}\}\}$ and the latter two sub-multisets do not entail $\psi_1 \wedge \psi_2$. □

Proposition 3. *Given C1–4, norms M1, M2*, M3, M4*, M5 and M6c are satisfied by the merging operator* $\Delta\mathscr{E} \stackrel{df}{=} [\{\theta \notin C_{\mathscr{E}} : \textit{some } E_i \models \theta\}]$.

Proof:

M1. C1 guarantees that $\Delta\mathscr{E}$ is consistent.

*M2**. $\bigwedge\mathscr{E} \models \Delta\mathscr{E}$ because for any $\theta \in \Delta\mathscr{E}$, some $E_i \models \theta$. $\Delta\mathscr{E} \models [\bigcap_{i=1}^{n} Cn(E_i)]$ by C2.

M3. C3 ensures that if $\mathscr{E} \equiv \mathscr{F}$ then $\Delta\mathscr{E} \equiv \Delta\mathscr{F}$.

*M4**. If $E_i \wedge E_j$ is inconsistent then there is some θ such that $E_i \models \theta$ but $E_j \models \neg\theta$. By C1, $\theta \in C_{\{E_i, E_j\}} \subseteq C_{\{E_i, \ldots, E_i, E_j\}}$ by C4. Hence $\Delta\{E_i, \ldots, E_i, E_j\} \not\models E_i$.

M5. If $\Delta(\mathscr{E} \sqcup \mathscr{F}) \models \theta$ then some $E_i \models \theta$ or some $F_j \models \theta$ and $\theta \notin C_{\mathscr{E} \sqcup \mathscr{F}}$ so by C4 $\theta \notin C_{\mathscr{E}}$ and $\theta \notin C_{\mathscr{F}}$. Hence $\Delta\mathscr{E} \wedge \Delta\mathscr{F} \models \theta$.

M6c. If $\Delta\mathscr{E} \models \theta$ then some $E_i \models \theta$. So if $\theta \notin C_{\mathscr{E} \sqcup \mathscr{F}}$ then $\Delta(\mathscr{E} \sqcup \mathscr{F}) \models \theta$.

□

Proposition 4. *Suppose all propositions are of the form θ^X where θ is a sentence of a propositional language and X is a convex set of probabilities. Let $\mathbb{E} = \langle\{P : P$ satisfies some $\mathscr{G} \in \hat{\mathscr{P}}\mathscr{E}\}\rangle$. Then $\Delta\mathscr{E} \stackrel{df}{=} \{\theta^X : P(\theta) \in X$ for all $P \in \mathbb{E}\}$ satisfies norms M1, M2*, M3, M4*.*

Proof: We write $P \models \theta^X$ for $P(\theta) \in X$. Note that satisfaction is closed under convex combinations: if $P \models \theta^X$ and $Q \models \theta^X$ then $\lambda P + (1-\lambda)Q \models \theta^X$ for $l \in [0,1]$. This can be seen as follows. $R = \lambda P + (1-\lambda)Q$ is defined by $R(\omega) = \lambda P(\omega) + (1-\lambda)Q(\omega)$ for each state (possible world) ω of the propositional language. $R(\theta) = \sum_{\omega\models\theta} \lambda P(\omega) + (1-\lambda)Q(\omega) = \lambda \sum_{\omega\models\theta} P(\omega) + (1-\lambda)\sum_{\omega\models\theta} Q(\omega) = \lambda P(\theta) + (1-\lambda)Q(\theta) \in X$ since X is convex.

M1. Note that $\mathbb{E}$ is a non-empty, convex set of probability functions. Hence if $\Delta\mathscr{E} \models \theta^X$ and $\Delta\mathscr{E} \models \theta^Y$ then $X \subseteq Y$ or vice versa.

M2.* If $\bigwedge\mathscr{E}$ is inconsistent then trivially $\bigwedge\mathscr{E} \models \Delta\mathscr{E}$. Otherwise, $\hat{\mathscr{P}}\mathscr{E} = \{\mathscr{E}\}$ and $\Delta\mathscr{E} \models \theta^X$ iff $P(\theta) \in X$ for all P satisfying $\bigwedge\mathscr{E}$, i.e., $\bigwedge\mathscr{E} \models \Delta\mathscr{E}$. Now if $E_i \models \theta^X$ for all i then $\bigwedge\mathscr{G} \models \theta^X$ for each $\mathscr{G} \in \hat{\mathscr{P}}\mathscr{E}$, so $P \models \theta^X$ for all $P \in \mathbb{E}$ (satisfaction being closed under convex combination), and $\Delta\mathscr{E} \models [\bigcap_{i=1}^n Cn(E_i)]$.

M3. If $\mathscr{E} \equiv \mathscr{F}$ then $\hat{\mathscr{P}}\mathscr{E} \equiv \hat{\mathscr{P}}\mathscr{F}$ and $\mathbb{E} = \mathbb{F}$, so $\Delta\mathscr{E} = \Delta\mathscr{F}$.

M4.* If $E_i \wedge E_j$ is inconsistent then $\hat{\mathscr{P}}\{E_i,\ldots,E_i,E_j\} = \{\{E_i,\ldots,E_i\},\{E_j\}\}$. So, $\mathbb{E} = \langle\{P : P \models E_i$ or $P \models E_j\}\rangle$. Since $E_i \wedge E_j$ is inconsistent, there is some $P \in \mathbb{E}$ such that $P \not\models E_i$. Hence $\Delta\{E_i,\ldots,E_i,E_j\} \not\models E_i$.

□

Bibliography

Adamcik, M. and Wilmers, G. (2014). Probabilistic merging operators. *Logique et Analyse*, in press.

Alchourrón, C. E., Gärdenfors, P., and Makinson, D. (1985). On the logic of theory change: partial meet functions for contraction and revision. *Journal of Symbolic Logic*, 50:510–530.

Aumann, R. J. (1976). Agreeing to disagree. *The Annals of Statistics*, 4(6):1326–1329.

Bird, A. (2007). Underdetermination and evidence. In Monton, B., editor, *Images of Empiricism: Essays on Science and Stances*, pages 62–82. Oxford University Press, Oxford.

Bratman, M. E. (1992). Practical reasoning and acceptance in a context. *Mind*, 101(401):1–15.

Corfield, D. (2001). Bayesianism in mathematics. In Corfield, D. and Williamson, J., editors, *Foundations of Bayesianism*, pages 175–201. Kluwer, Dordrecht.

Dawid, A. P. (1982). The well-calibrated Bayesian. *Journal of the American Statistical Association*, 77:604–613. With discussion.

Earman, J. (1992). *Bayes or bust? A critical examination of Bayesian confirmation theory.* MIT Press, Cambridge MA.

Foley, R. (1993). *Working without a net.* Oxford University Press, New York.

Gillies, D. (1991). Intersubjective probability and confirmation theory. *British Journal for the Philosophy of Science*, 42:513–533.

Hild, M., Jeffrey, R., and Risse, M. (2008). Preference aggregation after Harsanyi. In Fleurbaey, M., Weymark, J. A., and Salles, M., editors, *Justice, Political Liberalism, and Utilitarianism: Themes from Harsanyi and Rawls*, pages 198–218. Cambridge University Press, Cambridge.

Howson, C. (1997). Logic and probability. *British Journal for the Philosophy of Science*, 48:517–531.

Hylland, A. and Zeckhauser, R. (1979). The impossibility of Bayesian group decision making with separate aggregation of beliefs and values. *Econometrica*, 47(6):1321–1336.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *The Physical Review*, 106(4):620–630.

Jeffrey, R. (1968). Probable knowledge. In *Probability and the art of judgement*, chapter 3, pages 30–43. Cambridge University Press, Cambridge, 1982 edition.

Jeffrey, R. (2004). *Subjective probability: the real thing.* Cambridge University Press, Cambridge.

Kern-Isberner, G. and Rödder, W. (2004). Belief revision and information fusion on optimum entropy. *International Journal of Intelligent Systems*, 19:837–857.

Keynes, J. M. (1921). *A treatise on probability.* Macmillan, London, 1973 edition.

Konieczny, S. and Pino Pérez, R. (1998). On the logic of merging. In *Proceedings of the 6th International Conference on Principles of Knowledge Representation and Reasoning*, pages 488–498, Trento. Morgan Kaufmann.

Konieczny, S. and Pino Pérez, R. (2011). Logic based merging. *Journal of Philosophical Logic*, 40:239–270.

Kyburg Jr, H. E. (1991). Evidential probability. In Mylopoulos, J. and Reiter, R., editors, *Proceedings of the 12th International Joint Conference on Artificial Intelligence*,

pages 1196–1203, Sydney. Morgan Kaufmann.
Lewis, C. I. (1946). *An analysis of knowledge and valuation.* Open Court, La Salle, IL.
Lin, H. and Kelly, K. T. (2012). Propositional reasoning that tracks probabilistic reasoning. *Journal of Philosophical Logic*, 41:957–981.
List, C. (2012). The theory of judgment aggregation: an introductory review. *Synthese*, 187:179–207.
Littlejohn, C. (2011). Evidence and knowledge. *Erkenntnis*, 74:241–262.
Milne, P. (1991). A dilemma for subjective Bayesians—and how to resolve it. *Philosophical Studies*, 62:307–314.
Mongin, P. (2005). Spurious unanimity and the Pareto principle. LSE Choice Group working paper series, vol. 1, no. 5, Centre for Philosophy of Natural and Social Science (CPNSS), London School of Economics.
Quine, W. V. O. and Ullian, J. S. (1970). *The web of belief.* Random House, New York.
Rowbottom, D. P. (2014). Information versus knowledge in confirmation theory. *Logique et Analyse*, 57(226):137–149.
Shoham, Y. (1987). A semantical approach to nonmonotonic logics. In Ginsberg, M. L., editor, *Readings in nonmonotonic reasoning*, pages 227–250. Morgan Kaufmann, San Francisco, CA.
Skorupski, J. (2010). *The domain of reasons.* Oxford University Press, Oxford.
Williamson, J. (2009). Aggregating judgements by merging evidence. *Journal of Logic and Computation*, 19:461–473.
Williamson, J. (2010). *In defence of objective Bayesianism.* Oxford University Press, Oxford.
Williamson, J. (2011). Objective Bayesianism, Bayesian conditionalisation and voluntarism. *Synthese*, 178:67–85.
Williamson, J. (2013). From Bayesian epistemology to inductive logic. *Journal of Applied Logic*, 11(4):468–486.
Williamson, T. (2000). *Knowledge and its limits.* Oxford University Press, Oxford.
Williamson, T. (2007). *The philosophy of philosophy.* Blackwell, Oxford.
Wilmers, G. (2010). The social entropy process: Axiomatising the aggregation of probabilistic beliefs. In Hosni, H. and Montagna, F., editors, *Probability, Uncertainty and Rationality*, pages 87–104. Edizioni della Normale, Pisa.
Wittgenstein, L. (1969). *On certainty.* Harper, New York. G. E. M. Anscombe and Georg Henrik von Wright (eds).