Philos Stud (2009) 144:361–375
DOI 10.1007/s11098-008-9214-4

# Assertion, Moore, and Bayes

**Igor Douven**

Published online: 15 March 2008
© Springer Science+Business Media B.V. 2008

**Abstract** It is widely believed that the so-called knowledge account of assertion best explains why sentences such as "It's raining in Paris but I don't believe it" and "It's raining in Paris but I don't know it" appear odd to us. I argue that the rival rational credibility account of assertion explains that fact just as well. I do so by providing a broadly Bayesian analysis of the said type of sentences which shows that such sentences cannot express rationally held beliefs. As an interesting aside, it will be seen that these sentences also harbor a lesson for Bayesian epistemology itself.

**Keywords** Assertion · Moore · Bayesian epistemology

According to the knowledge account of assertion, the practice of assertion is governed by the rule that

One should assert only what one knows. (1)

Its many proponents point to various types of linguistic data and contend that the knowledge account explains these data better than any rival account of assertion.[1] Among these data are people's general reluctance to assert, previous to the drawing of a lottery, that a ticket of that lottery is a loser; the fact that the question "How do you know?" is a socially accepted response to an assertion; and the fact that assertions of instances of the "Moorean" schemas

[1] See, most notably, Williamson (2000), Adler (2002), and DeRose (2002).

I. Douven (✉)
Institute of Philosophy, University of Leuven, Leuven, Belgium
e-mail: igor.douven@hiw.kuleuven.be

$$\varphi\text{, but I don't believe }\varphi \tag{2}$$

and

$$\varphi\text{, but I don't know }\varphi \tag{3}$$

sound, or would sound, positively odd to anyone's ears.[2] That the knowledge account is able to explain all these data in a satisfactory way is beyond dispute. That it explains them best is not.

Elsewhere I have defended what I called the rational credibility account of assertion, according to which assertions are governed by the rule that

$$\text{One should assert only what is rationally credible to one,} \tag{4}$$

and argued that, on balance, this account provides a better explanation of the relevant types of phenomena than the knowledge account.[3] I say "on balance," because I was unable to show that it explains better, or even at least as well, *each* of the relevant types of phenomena. I showed that the rational credibility account has no difficulty explaining the data involving lottery tickets nor why it is socially acceptable to ask an asserter how she knows what she asserted, and that it furnishes an explanation of our pre-theoretic judgements about assertions of rationally credible falsehoods that is superior to the one offered by the proponents of the knowledge account.[4] But with regard to explaining the odd-soundingness of "Moorean sentences"—as I shall call the instances of (2) and (3)—I thought it had to be conceded that the knowledge account does markedly better than the rational credibility account.

The primary aim of the present paper is to argue that this concession was rash and that, on a more careful examination of the matter, the rational credibility account appears to do as good a job in explaining the odd-soundingness of those sentences as the knowledge account. A secondary aim is to give a broadly Bayesian analysis of Moorean sentences, which so far is still missing from the literature. This is surprising, given the current popularity of Bayesian epistemology; in fact it is doubly surprising, given that, as will be seen, the said sentences harbor a lesson for that epistemology itself. It also seems to be a serious omission, as there is still no consensus about exactly why Moorean sentences seem odd, and from a Bayesian perspective this can be explained on the basis of some seemingly quite innocuous assumptions.

1. The knowledge account explains the odd-soundingness of Moorean sentences as follows: By conceptual necessity, a person cannot both know $\varphi$ and know that she does not know $\varphi$, nor can she both know $\varphi$ and know that she does not believe $\varphi$.[5] Hence she cannot know any instance of either (2) or (3) and thus, given the

[2] Famously, Moore (1962) contains the first discussion in the literature of instances of (2) and (3).

[3] See Douven (2006).

[4] Or rather to the *ones* offered by those proponents. There are slight differences between Williamson's (2000, p. 256) and DeRose's (2002, 180 f) explanations of the designated judgements; see Douven (2006, Sect. 6).

[5] Supposing, in the latter case, that knowledge requires belief. See Lewis (1996, p. 227) for a (rare) dissenting opinion.

knowledge account, cannot have warrant for asserting any such instance. So their odd-soundingness is straightforwardly explained by their unassertability.[6]

This explanation makes only uncontroversial assumptions about knowledge: that it distributes across conjunctions, so that knowing an instance of either (2) or (3) requires knowing both conjuncts of the instance, and that it is factive, so that knowing that you do not believe $\varphi$ is impossible if you do believe it. Prima facie, it already seems that an explanation of why instances of (2) cannot be *rationally credible* cannot be this simple, given that rational belief is not factive, and that an explanation of why instances of (3) cannot be rationally credible must be even more involved: it is not immediately clear that it could never happen that, for some $\varphi$, $\varphi$ is rationally credible to a person and at the same time it is rationally credible to this person that she does not know $\varphi$.

But advocates of the rational credibility account of assertion are not committed to the unassertability of Moorean sentences. They must explain why these sentences sound odd to us, which they may then in turn use to explain why so many have thought that the sentences are unassertable. One way to try and explain the odd-soundingness of Moorean sentences that does not require their unassertability starts by noticing that the odd-soundingness of an expression may simply be due to a lack of exposure to that expression. So, instances of (2) and (3) may sound odd for no other reason than that we never hear any of them being asserted. Of course this raises the question *why* we never hear them, if they are not necessarily unassertable. The second part of the explanation now points to the fact that there may be good pragmatic reasons for not adding "but I don't believe it" or "but I don't know it" to anything we have asserted. Indeed, appending to an assertion any remark that reflects negatively on one's epistemic position vis-à-vis that assertion is likely to thwart one's attempt to convince the audience, which is typically the purpose of making an assertion. This is, in a nutshell, the explanation I offered in Douven (2006, Sect. 5).

I admitted there that this explanation is much less simple and elegant than the one provided by the knowledge account. It now seems to me, however, that a bigger problem with the former explanation is that it is empirically inadequate. For consider that there exist many sentences and even types of sentences to which we are never, or hardly ever, exposed and which yet do not sound odd, or at least not to the extent that, and not in the way in which, Moorean sentences sound odd. Here one may think of sentences which the world gives us little opportunity to use, like "John seeks a unicorn" for instance, or sentences that are grammatically correct but somehow awkward, like conjunctions with very many conjuncts or (multiply) nested conditionals. Though the term "odd-sounding" is a bit vague, it still seems clear enough that the odd-soundingness of Moorean sentences is of a very different category than that of other sentences to which we have little or no exposure.[7] As a

[6] See, for instance, Unger (1975, 263 f), Williamson (2000, 252 f), and Adler (2002, p. 36).

[7] It might be thought that this is so because Moorean sentences do not merely sound odd but are even heard as contradictory. However, I think what DeRose (1991, p. 597) says in relation to an instance of (3) holds generally for all of us: "I don't have a special feeling for inconsistencies; I can sense some kind of clash, but cannot distinguish my sensing of an inconsistency from my sensing of whatever it is that's wrong with the Moorean sentence." See on this also Douven (2006, 475 f).

result, the explanation I offered seems to fail: lack of exposure does not generally explain odd-soundingness, at least not the variety of odd-soundingness exhibited by Moorean sentences.

This does not sink the rational credibility account, for, as intimated earlier, I think it is still able to explain the odd-soundingness and, in fact, to explain it in a quite simple and natural way. In claiming the latter, I am assuming that you share my and several other philosophers' conviction that it is natural to think of our beliefs both in a qualitative and in a quantitative fashion, and that besides categorical or outright beliefs we also have graded beliefs.[8] I am further assuming that you are willing to grant some seemingly innocuous assumptions involving (and connecting) those notions. On these assumptions, it follows that no Moorean sentence can be rationally believed so that, given the rational credibility account of assertion, such sentences are unassertable. Hence, just as the proponent of the knowledge account is able to do, the advocate of the rational credibility account of assertion can explain the odd-soundingness of Moorean sentences by pointing to their unassertability, and thus need not resort to the above (or any other) pragmatic explanation of it.

So far as I can see, the new explanation to be offered is no less simple than the explanation, cited in the first paragraph of this section, based on the knowledge account. It is worth stressing, however, that whether this is really the case is not so important for my purposes. On the score on which the knowledge account seemed to offer the patently best explanation, we are now able to offer a competing explanation based on the rational credibility account which is decisively simpler than our earlier pragmatic explanation. So even if the new explanation is not quite as simple as the one based on the knowledge account, we will still be able to claim—and with more right now—that, *on balance*, the rational credibility account does a better job than its rival in explaining the data from the practice of assertion.

Before going into the details of the explanation, it should briefly be mentioned that previous attempts have been made to show that Moorean sentences cannot be rationally believed.[9] Apart from the fact that most of these are attempts only to explain why instances of (2) cannot be rationally believed, they either seem inadequate to me or to rely on assumptions that are too strong or at least stronger than is required for our present purposes.[10] (Besides, none of them is Bayesian or can obviously be extended to a Bayesian analysis of Moorean sentences; as intimated earlier, giving such an analysis of these sentences is of independent interest.)

For instance, Williams (1996, p. 137) argues that instances of (2) fail to be rationally credible because a person's belief in any of them is necessarily false. But

[8] See, for instance, Foley (1992) for a compelling defense of this idea.

[9] Some authors even claim to show that Moorean sentences cannot be believed *tout court*; see for instance Hintikka (1962, p. 67) and Tennant (1997, 251 f). It would seem that that claim just cannot be true, for, we may suppose, the madman can believe anything. On closer inspection, however, it appears that these authors typically make certain assumptions about the notion of belief that rule out mad belief, and in fact require some notion of rational belief.

[10] Or they are a bit vague. For instance, Shoemaker's (1995) analysis relies on a notion of "availability" of beliefs which he himself confesses to stand in need of further elucidation (p. 227).

this cannot be correct, for necessary falsehoods can be rationally credible. It is necessarily false that arithmetic is complete, and yet in the years preceding the publication of Gödel's famous incompleteness results, most mathematicians who were interested in the issue believed arithmetic to be complete, and it would seem that they did so rationally on any pre-theoretically plausible standard of rationality.[11] Or, if you find this example unconvincing, just think of the many false identity and non-identity statements which once were rationally credible to us because we had strong, though misleading, evidence for them (a well-worn example: "Hesperus and Phosphorus are distinct heavenly bodies").

Tennant (1997, 251 f) also argues that at least instances of (2) are not rationally credible, but in doing so he makes a number of stark idealizations. Among others, he assumes that it is inconsistent to suppose that someone believes a necessary falsehood, whether or not any actual reasoner is capable of recognizing it as such. He further assumes that if it is inconsistent to suppose that a person believes $\varphi$, then it is inconsistent to suppose that she asserts it, too. But of course this seems tenable only on the assumption that the person is rational and would never assert anything that would commit her to an inconsistency and would, in addition, notice whenever the supposition that she believes a given proposition at a given time is inconsistent. While the idealizations involved in these assumptions may be acceptable within the confines of the discussion of semantic antirealism, in which Tennant's analysis of Moorean sentences plays a crucial dialectical role,[12] we clearly cannot get a satisfactory explanation of why those sentences sound odd to *us*, non-ideal mortals, by arguing from premises that are true only of idealized beings.[13]

2. To begin my own explanation, I first make some minimal assumptions about the notion of rational (categorical) belief. First, I assume that rational belief is closed under the rule of conjunction elimination, that is,

$$\mathrm{RB}_i(\varphi \wedge \psi) \Rightarrow \mathrm{RB}_i(\varphi), \tag{5}$$

where "$\mathrm{RB}_i(\varphi)$" means that a given person $i$ rationally believes $\varphi$ and the symbol $\Rightarrow$ stands for entailment. This seems to be a conceptual truth. In any event, it should be uncontroversial, given that we unhesitatingly report, or would report, that someone rationally believes $\varphi$ once we know that she rationally believes the conjunction of $\varphi$ and $\psi$ (for any $\varphi$ and $\psi$).

[11] In 1928, Presburger had proved the completeness of arithmetic without multiplication. This made it seem as if a completeness proof for Peano arithmetic was just around the corner.

[12] See Douven (2005, Sect. 4) for discussion.

[13] Hintikka's (1962) analysis of instances of (2) faces essentially the same objection. As he admits (p. 36), his "results are not directly applicable to what is true or false in the actual world of ours" because of the various idealizations involved in the doxastic logic he develops and deploys. On Williamson's (2000) and Adler's (2002) analyses of Moorean sentences, these come out as not being rationally credible as well. However, for them this follows from the fact that these sentences cannot be known, together with the principle that one ought to believe only what one knows (Williamson 2000, 255 f), or that one properly believes something only if one knows it [in Adler's (2006, p. 284) formulation]. This principle in turn follows from their assumption that, loosely speaking, assertion and belief are two sides of the same coin (Williamson 2000, p. 255; Adler 2002, p. 74), in conjunction with the knowledge account of assertion. I share the former assumption (Douven 2006, Sect. 1), but, of course, not the latter.

The second assumption concerns a very weak relationship between rational and graded belief. Let $\mathrm{Cr}_i$ be the degrees-of-belief function of person $i$; it is left open, at least for now, whether $\mathrm{Cr}_i$ obeys the axioms of probability theory or is rational in any other sense. Then the assumption is that

$$\mathrm{RB}_i(\varphi) \Rightarrow \mathrm{Cr}_i(\varphi) > \mathrm{Cr}_i(\neg\varphi), \tag{6}$$

that is, rationally believing something requires believing the thing to be more likely than not.[14] This seems to be another conceptual truth about rational belief. It may in effect already hold on the basis of the concept of (categorical) belief alone, for it would seem absurd to say that one could believe something even though one deems it at least as likely *not* to be the case.[15]

Note that (6) is consistent with the assumption that believing something, or at least believing it rationally, requires believing the thing to a *much* greater degree than its negation. Also note that it is consistent with (6) that the degree of belief required for (categorical) belief, or at least rational belief, is context-dependent as long as there is no context in which one qualifies as rationally believing a sentence which one believes to the same or a lower degree than the sentence's negation.[16] Finally, note that it is not assumed that finding a thing to be more likely than not is sufficient for rational belief. That assumption would seem wrong, if only because we are not assuming degrees of belief to be rational. Even if they *were* assumed to be rational, the assumption would still be problematical. It may be rational for me to find it much more likely that my lottery ticket is a loser than that it is a winner, yet—most philosophers would say—the lottery paradox shows that it is not rational for me to believe outright that it is a loser.[17]

While I expect (6) to be *largely* uncontroversial, I should still note a possible objection to it that may be extracted from Maher's (1993, 137 ff) discussion of the relation between assertion and high (subjective) probability. According to Maher,

[14] If degrees of belief were assumed to be probabilities, then (6) would of course be equivalent to this: $\mathrm{RB}_i(\varphi) \Rightarrow \mathrm{Cr}_i(\varphi) > 0.5$. Absent that assumption, it is at least as far as (6) goes possible for a person rationally to believe something that she believes to a degree of 0.3, say, provided she believes its negation to a lower degree.

[15] It does hold on the accounts of belief to be found in Mellor (1993, p. 233), van Fraassen (1995, Sect. 4), Hunter (1996, p. 87), Schiffer (2005, p. 278), and Weatherson (2005, p. 424), among others.

[16] For the suggestion of context-dependence, see for instance Schiffer (2005) and Weatherson (2005).

[17] Some, like Kyburg (1970) and Foley (1992), think this *is* rational to believe. They avoid the lottery paradox by rejecting what is sometimes called "the conjunction principle," according to which rational belief is closed under the rule of conjunction introduction. But at least a wholesale rejection of this principle seems hard to defend; see Douven (2002, Sect. 2) and Douven and Williamson (2006) for more on this. (Incidentally, there are approaches to the lottery paradox that maintain the conjunction principle, and yet permit that it *can* be rational to believe of one's ticket that it is a loser—though on these approaches it cannot be rational to believe this of every ticket in the lottery; see, e.g., Harman (1986, p. 71), Johnson (2004, 134 ff), and Douven (2008a).) On the other hand, it is well known that if one endorses (6) and one does not want to exclude that degrees of belief can be probabilities—and, as will become apparent shortly, for us it is essential *not* to exclude this—then one will need some response to Makinson's (1965) preface paradox. See Douven and Uffink (2003) and Weatherson (2005) for different solutions to this paradox which both involve placing a relatively weak restriction on the conjunction principle; Douven (2008b) makes the suggestion that the preface paradox may not be telling in the least against the conjunction principle, but may at bottom be a confirmation-theoretic problem stemming from the fact that currently we have less than full clarity about the notion of relevant evidence.

"anyone reflecting on the history of science ought to give a low probability ... to any given significant current theory being literally correct" (p. 137). And yet, he thinks, "scientists continue to sincerely assert significant scientific theories" (ibid.). Hence, if assertion requires rational credibility, as it does according to the rational credibility account, then, it seems, (6) can be maintained only by assuming—implausibly—that scientists are typically far too confident about their theories.[18]

By way of response, first let me remark that, in view of much recent work done in the scientific realism debate, Maher's remark that we should assign low probabilities to all current scientific theories is (mildly put) contentious. Especially noteworthy here are Kitcher's (1993, Chap. 5) and Psillos's (1999, Chap. 5) critiques of Laudan's (1981) so-called pessimistic meta-induction, which show that considerations from the history of science are hardly able to undermine the epistemic status of current scientific theories in the general way required to warrant Maher's claim. In addition to this, it is far from implausible to assume that assertions of scientific theories conversationally implicate that the asserter takes the theory she asserts to be *approximately* true. After all, it is common knowledge, at least among scientists, that theories tend to hold under various idealizations only.[19] So, even if it should be true that scientists typically assign low probabilities to their theories, they may still think it highly probable that these theories are approximately true, which may be all that is conveyed by asserting such a theory. As Maher provides no textual evidence whatsoever that suggests otherwise—let alone that he provides evidence suggesting that scientists typically, or at least sometimes, rationally believe theories they deem no more likely to be true than false—the foregoing remarks should suffice to put to rest the worry that data from the practice of science militate against (6).

The next assumption is that rational belief is weakly transparent in the sense that if a person rationally believes something, then her degree of belief that she categorically—but not necessarily rationally—believes that thing is at least as great as her degree of belief that she does *not* categorically believe it. That is to say,

$$\mathrm{RB}_i(\varphi) \Rightarrow \mathrm{Cr}_i\big(\mathrm{B}_i(\varphi)\big) \geqslant \mathrm{Cr}_i\big(\neg \mathrm{B}_i(\varphi)\big), \tag{7}$$

where "$\mathrm{B}_i(\varphi)$" means that person $i$ categorically believes $\varphi$. I have three comments on this principle.

First, standard logics of categorical belief, most notably $\mathsf{K45}$ and $\mathsf{KD45}$, already assume that such belief is transparent in the sense that the (categorical) belief

[18] It may be noticed that if Maher is right about what follows from the history of science, then the fact that scientists seem to assert outright their theories on conferences and in publications is not just bad news for (6), but also for the knowledge account of assertion: even if those are right who think that knowledge does not require probability 1 (see, e.g., DeRose 1996, p. 568, 577 f), it is beyond serious dispute that knowledge is incompatible with *low* probability.

[19] The supposition that scientists' assertions of theories typically have the said kind of implicature is quite consistent with these assertions themselves' ("what is said," in Grice's (1989) terminology) being unqualified and categorical. Maher seems to overlook this when he rejects the designated supposition by saying that "Einstein is categorically asserting [the General Theory of Relativity], not merely that the theory is approximately correct" (p. 138).

operator satisfies the S4 modal principle $\Box\varphi \to \Box\Box\varphi$.[20] Clearly, (7) is weaker than that in two respects: It does not apply as soon as one believes something; the belief must be rationally held. And, it does not require that a person believe of each of those things that she rationally believes that she believes them, but just that if one rationally believes something, then one does not think it more likely that one does *not* believe it than that one does believe it, which, for all we are supposing here, is compatible with not believing it.

(In his formal analysis of instances of (2), Hintikka (1962, p. 69) also makes use of the assumption that the belief operator (or rather the rational belief operator–cf. note 9) obeys the S4 principle (he makes use of the corresponding assumption for the knowledge operator in his analysis of instances of (3)). He claims, though, that the principle is not really needed (p. 70). His claim is based on the assumption that in order to show that instances of (2) cannot be believed it suffices to show that someone who asserts such an instance "could not have added 'and I believe what I just said' without making the resulting pair of statements obviously indefensible" (ibid.). This is strange, given that what we are after is an explanation of why the instance of (2) without the said addition appears to be "obviously indefensible," and we surely do not automatically get that explanation by explaining why the instance *with* the addition is indefensible. In the same context, Shoemaker (1995, p. 220) commits himself to the principle that "a rational person who believes that *P* at least tacitly believes that she believes that *P*." Given his understanding of (rational) tacit belief this is weaker than the assumption that the rational belief operator obeys the S4 principle, but it is still stronger than (7).)

Second, it must be emphasized that for the person to whom $\mathrm{RB}_i$, $\mathrm{B}_i$, and $\mathrm{Cr}_i$ refer, the right-hand side of (7) is supposed to express an attitude *de se* vis-à-vis $\varphi$: what the person finds at least as likely as not is that *she herself* (categorically) believes $\varphi$, not that the person called "$i$" believes $\varphi$, where it may be hidden from her that she is the person called "$i$." So, for instance, if the operators contain a reference to me, then the right-hand side means that I find it at least as likely as not that *I* believe $\varphi$.

And third, even though it is weaker than the principle that believing $\varphi$ entails believing that one believes $\varphi$, one might still have qualms about (7). For—one might worry—could a belief, even one held on good grounds, not be hidden so well in the back corners of a person's memory (perhaps because its content is painful to her) that it might in fact be unlikely to her that it is among her beliefs? Besides, beliefs may be held implicitly; many people will rationally believe that, to use Searle's (2002, p. 196) example, "people tend to vote near the surface of the earth," without having an explicit belief with that content in their belief boxes (to speak figuratively). Might one not easily fail to realize that the said belief is among one's

[20] See Meyer and van der Hoek (1995, Chap. 2). It is worth noting that nothing said so far implies that the principle $\mathrm{RB}_i(\varphi) \Rightarrow \mathrm{RB}_i\big(\mathrm{RB}_i(\varphi)\big)$ is indefensible. However, if—as some might want to argue—rational belief requires a margin for error in the sense of Williamson (2000, Chap. 5), then this principle would seem to run afoul of the type of argument Williamson (2000, Chap. 4) deploys to support his claim that knowledge is anti-luminous (but also see Mendola 2007). Note that (7) is not endangered by such an argument, if only because it would be implausible to think that with the notion of finding something more likely than not to be the case, there would be associated a margin for error.

beliefs, and therefore believe that it is among one's beliefs only to a very low degree, or even a degree of 0?

There are at least two possible responses to these questions. The first is to note that we can simply dodge them by restricting (7) and the other principles presented in this section to beliefs that are conscious (or conscious and explicit, but I presume that consciousness requires explicitness). After all, it seems that for purely psychological reasons asserting $\varphi$ is incompatible with $\varphi$'s being buried deeply in one's mind, or with merely implicitly believing it. However unconscious an asserter *might* otherwise have been of her believing $\varphi$, or however implicit the belief *might* have been, we may suppose that in contemplating asserting $\varphi$—at the latest—the asserter will become consciously aware of her having the belief. So, we will have explained why Moorean sentences are unassertable—and hence sound odd—if we have shown that they cannot be consciously believed, on pain of irrationality. And, as will be apparent from the arguments to be given below, restricting our principles in the suggested way would still allow us to show that.

The second response starts by recalling that degrees of belief are generally assumed to be measurable by betting rates.[21,22] More exactly, a person's degree of belief in $\varphi$ is thought to be measured by the highest price she is willing to pay for a bet that pays one monetary unit if $\varphi$ turns out true, and nothing otherwise. Now suppose that you believe only to a low degree that you categorically believe that people tend to vote near the surface of the earth. Given the just-mentioned assumption this would mean that the maximum price you are willing to pay for a bet that you do have that categorical belief will be low relative to what the bet pays in case you have the belief. But it would seem that a proposal to bet that you categorically believe that people tend to vote near the surface of the earth will, by mentioning the very belief in the course of making the proposal, be enough to make you aware that it *is* among your beliefs (plausibly supposing that it is). Consequently, it would seem absurd to think that the maximum price you are willing to pay for the bet is low relative to what the bet pays. It thus seems hard to maintain that you could believe to a low degree that you categorically believe that people tend to vote near the surface of the earth, however well hidden or implicit that belief may be.

We need one more doxastic principle:[23]

$$\begin{array}{l}\text{A person believes } \varphi \text{ rationally only if it does not readily follow}\\ \text{strictly on the basis of the assumption of her rationally believing } \varphi \text{ plus}\\ \text{principles (5)–(7) that her degrees of belief are not probabilities.}\end{array} \tag{8}$$

[21] According to operationalism, which still has its defenders in the social sciences (see Gillies 2000, Chap. 9), a person's degrees of belief can even be *identified* with the rates at which she would bet under certain conditions (cf. Gillies 2000, p. 200). The exact metaphysics of graded belief need not detain us here, however.

[22] I should note that not all who have written on degrees of belief accept this assumption; see, for instance, Christensen (2004, 113 f). I hope that for the dissenters the first response suffices.

[23] It will be noticed that in the statement of this principle the phrase "her degrees of belief" can only be meant as a nonrigid designator. For else it would follow from any set of premises that the person's degrees of belief are not probabilities if in fact they are not (for that would be a necessary truth then).

Here, "readily follow" may be taken to imply that hardly any real logical skills are needed to recognize the consequence. This principle does not assume, note, that rational belief, at least insofar as we are specifying it, requires that one's degrees-of-belief function *is* a probability function. For all we have said, a person may, for instance, believe $\varphi$ rationally even if she fails to believe many tautologies to a degree of 1; that she fails to do so, or that she violates probability theory in any other way, just must not readily follow from the designated assumption and principles. It is quite important, of course, that I do not have to take on board the usual idealizations of Bayesian epistemology, for else my analysis of Moorean sentences would be vulnerable to basically the same objection that I raised against Tennant's; it might then help to explain why these sentences appear odd to perfect Bayesian reasoners (supposing they do), but not why *we* find them odd. If an ordinary mortal, possibly lacking any special logical capacities, can easily recognize that from the assumption that her belief in $\varphi$ is rational it follows, given certain simple and intuitively evident principles, that her degrees of belief violate the axioms of probability, then surely we are not setting the standards for rationality too high by declaring that her belief in $\varphi$ fails to be rational.

The above principles are not meant as jointly constituting a definition of rational belief; they are only supposed to state conditions necessary for rational belief. This is enough for our purposes: if instances of (2) and (3) cannot be rationally believed on the grounds that belief in them would fail to satisfy the said conditions, then of course they cannot be rationally believed on any more inclusive account of rational belief.

In order to see that these instances cannot be rationally believed, first consider the Moorean sentence

$$\text{It's raining in Paris, but I don't believe that it is raining in Paris,} \tag{9}$$

and assume, toward a *reductio*, that a given person $i$ does rationally believe it. By principle (8) it should then at least not readily follow from this assumption and principles (5)–(7) that there exists no probability function Pr such that $\mathrm{Cr}_i = \mathrm{Pr}$ (where, recall, $\mathrm{Cr}_i$ is the person's degrees-of-belief function). But suppose that $\mathrm{Cr}_i = \mathrm{Pr}$ for some Pr, suppose, that is, that $i$'s degrees of belief *are* probabilities. Then from (6) we have

$$\mathrm{Cr}_i\,(\text{It's raining in Paris, but I don't believe that it is raining in Paris}) > 0.5 \tag{10}$$

and hence also, by probability theory,

$$\mathrm{Cr}_i(\text{I don't believe that it's raining in Paris}) > 0.5. \tag{11}$$

Furthermore, from the assumption that the person rationally believes (9) it follows by the conjunction elimination principle (5) that she rationally believes that it is raining in Paris. By (7) and again the assumption that her degrees of belief are probabilities, it then follows that

$$\mathrm{Cr}_i(\text{I believe that it's raining in Paris}) \geqslant 0.5. \tag{12}$$

However, from (11) and (12) it follows that it cannot possibly be that $\mathrm{Cr}_i = \mathrm{Pr}$ for some probability function Pr. For "I believe that it's raining in Paris" and "I don't believe that it's raining in Paris" are contradictories and so, by the axioms of probability theory, their probabilities must sum to 1, whereas $\mathrm{Cr}_i$(I believe that it's raining in Paris) + $\mathrm{Cr}_i$(I don't believe that it's raining in Paris) > 1. Thus, from our supposition that (9) is rationally believed by person $i$ and principles (5)–(7) we have derived that $i$'s degrees-of-belief function is not a probability function, where moreover the derivation consists of no more than a few elementary deductive steps, so that the consequence can be seen to follow readily on any reasonable interpretation of "readily." Hence, from the said assumptions we have derived that (9) is not rationally believed by her. Consequently, (9) cannot be rationally believed. Nor, clearly, can any other instance of (2).[24]

3. We cannot show in quite the same way that instances of (3), like

$$\text{It's raining in Paris, but I don't know that it is raining in Paris,} \tag{13}$$

cannot be rationally believed either. By the foregoing reasoning, we can derive from the assumption that (13) is rationally believed by a person $i$ both

$$\mathrm{Cr}_i(\text{I don't know that it's raining in Paris}) > 0.5 \tag{14}$$

and (12) again, but from (12) and (14) it does not follow that $\mathrm{Cr}_i$ cannot be a probability function. For all we have said thus far, one may find it quite likely that one does not know that it is raining in Paris, but not quite so likely that one does not believe it; perhaps one believes that one's justification for the belief that it is raining in Paris is somehow deficient, so that the belief does not amount to knowledge.

The following additional principle will help us out with (13):

$$\mathrm{RB}_i(\varphi) \Rightarrow \mathrm{Cr}_i\big(\mathrm{K}_i(\varphi)\big) \geqslant \mathrm{Cr}_i\big(\neg\mathrm{K}_i(\varphi)\big), \tag{15}$$

with "$\mathrm{K}_i(\varphi)$" meaning that person $i$ knows $\varphi$. It is easy to see how this helps to derive the conclusion that (13) cannot be rationally believed. Supposing that the person rationally believes (13), it follows by (5) that she rationally believes that it is raining in Paris. Thus, by (15) and the assumption that her degrees of belief are probabilities,

$$\mathrm{Cr}_i(\text{I know that it's raining in Paris}) \geqslant 0.5. \tag{16}$$

And of course together with (14) this does show that the person's degrees of belief cannot be probabilities, contradicting our supposition that (13) was rationally believed. Hence the supposition that (13) is rationally believed leads to a contradiction, and quite readily so. Therefore, (13) cannot be rationally believed. Again the conclusion generalizes swiftly to all other instances of (3).

[24] The referee worried that this is a relatively shallow explanation of why we cannot believe instances of (2) and that a more thorough one might well require norms stronger than (5)–(8). First, however, I do not share the referee's intuition that the present explanation is shallow; it is simple, to be sure, but I do not think of this as being a disadvantage. Second, invoking stronger norms would almost certainly go at the expense of the generality of the present explanation, and I want the explanation to be as general as possible in order not to tie the rational credibility account of assertion to any specific theory of rationality.

The last principle, (15), may appear to be not as compelling as the earlier ones. However, first note that we assumed that rationally believing something implies believing it to be more likely than not. If, as some think, rational belief, or even belief *tout court*, requires a degree of belief close to subjective certainty,[25] then (15) can be supplanted by something like the much weaker principle that a person cannot rationally believe something and at the same time be close to subjectively certain that she does not know it. For example, if rational belief requires a degree of belief of 0.9, then for the above argument to go through it is enough to assume that if one rationally believes something, then one's degree of belief that one knows it is no lower than 0.1. That seems to be quite undemanding.

Second, (15) is still weaker than what Adler (2002, 36 ff) (tentatively) argues for, to wit, that we cannot but regard our beliefs as being known to us. Saying that we regard our beliefs as being known to us manifestly requires more than that we think of our *rationally held* beliefs that it is *at least as likely as not* that they are being known to us.

But we can do better than making these defensive remarks, for note, finally, that it seems possible to argue for (15) by means of an inference to the best explanation, as follows. Run through some batch of your rationally held beliefs. Then—supposing I am not somehow exceptional in this respect—you will encounter none that you do not take to be very likely to be known by yourself, let alone that there would be ones which you think more likely not to be known than to be known by you.[26,27] While this *might* be a mere coincidence, a far better explanation for it—in fact, the best, so far as I can see—is to assume that (15) holds true, that is, that it is because the concepts of rational belief, graded belief, and knowledge are connected in ways which render (15) true that we seem to find it at least as likely as not of our rationally held beliefs that they are known by us. In short, (15) is justified because it is the best explanation for (what we may assume are) the empirical data.[28,29]

4. In the above, we have made use of part of the apparatus of Bayesian epistemology in order to show that Moorean sentences cannot be rationally believed and hence, given the rational credibility account of assertion, that they are not

[25] See, for instance, Mellor (1993, p. 233); on van Fraassen's (1995, Sect. 4) account, belief even requires certainty.

[26] In view of such arguments as are to be found in, among others, Pollock (1986, Chap. 5), Wedgwood (1999), and BonJour's contribution to BonJour and Sosa (2003), the assumption that you can determine of your rationally held beliefs that they are rationally held seems utterly plausible to me. For present purposes, however, it suffices to assume that you can determine of *enough* of your rational beliefs that they are rationally held to make a case for (15).

[27] It in effect appears to me that I am *fully* convinced of any of my rationally held beliefs that I know them. So, it may be possible to give an abductive argument even for the principle that rationally believing something entails subjective certainty that one knows the thing (which, of course, we do not need).

[28] There may be a more direct argument for (15), one revealing the exact conceptual interconnections between rational belief, graded belief, and knowledge in virtue of which (15) holds true (supposing it does hold true). But for now we can rest content with pointing to the abductive argument for (15) given here.

[29] It is worth noting in this connection that, as Jonathan Adler pointed out to me (in personal correspondence), to think of $\varphi$ as being known by oneself does not require one to think of $\varphi$ under a description of it as satisfying the conditions for knowledge. Indeed, for all (15) requires, one need not have the faintest idea of what those conditions are.

warrantedly assertable. But these sentences also show something of importance to the Bayesian position itself. According to this position, the key notion of doxastic evaluation is that of *probabilistic coherence* (or simply *coherence*): degrees of belief are rational iff they are coherent iff they are representable by a probability function. Now, as far as probability theory goes, the only sentences that cannot possibly be assigned perfect probability are ones that are self-contradictory.[30] Since, as has been frequently noted in the literature, instances of both (2) and (3) are consistent provided what is substituted for $\varphi$ is consistent, by the foregoing it should be possible to rationally believe, for instance, (9) to a degree of 1. However, if we accept what again appear to be conceptual or near-to-conceptual connections between the notions of categorical and graded belief, then we are barred from ever rationally believing that sentence, or similar ones, to such a degree.

Of course Bayesians are not typically concerned with the notion of categorical belief. But even if they are right that this notion should play no central role in a serious epistemology, they cannot deny that it figures prominently in our ordinary ways of talking and thinking. Nor, it would seem, can they deny the following connection between categorical and coherent graded belief:

$$\mathrm{Pr}_i(\varphi) = 1 \Rightarrow \mathrm{B}_i(\varphi), \tag{17}$$

where the use of "$\mathrm{Pr}_i$" instead of "$\mathrm{Cr}_i$" is meant to indicate that person $i$'s degrees of belief are assumed to be probabilities. It may be that no coherent degree of belief smaller than 1 suffices for categorical belief, but if one coherently believes something to the highest possible degree, then one believes it categorically. Furthermore, a principle like (5) will also hold for categorical belief; one cannot believe a conjunction without thereby also believing the conjuncts, that is,

$$\mathrm{B}_i(\varphi \wedge \psi) \Rightarrow \mathrm{B}_i(\varphi). \tag{18}$$

And it was said that (6) appears to hold already on the basis of the concept of categorical belief. Thus, for any person $i$ with coherent degrees of belief we will have the following:

$$\mathrm{B}_i(\varphi) \Rightarrow \mathrm{Pr}_i(\varphi) > 0.5. \tag{19}$$

Finally, categorical belief may be assumed to be transparent at least in the exceedingly weak sense that if you categorically believe something, then you are not 100% sure that you do *not* believe it. We may therefore assume that for any person $i$ with coherent degrees of belief it holds that

$$\mathrm{B}_i(\varphi) \Rightarrow \mathrm{Pr}_i\big(\mathrm{B}_i(\varphi)\big) > 0. \tag{20}$$

For present purposes, it is unimportant whether principles (17)–(20) are best regarded as further necessary conditions for rational graded belief, next to coherence, or whether they are more sensibly thought of as preconditions for rationality, in the sense that a person's not respecting them exhibits an insufficient

[30] This is false if we adopt the requirement of strict coherence, which impels us to reserve probability 1 for tautologies only (see Kemeny 1955 and Jeffreys 1961). However, most Bayesian epistemologists nowadays think that the requirement of strict coherence is untenable; see Howson (2000) and Hájek (2003) for some strong objections against it.

grasp of concepts whose possession, even if perhaps only implicit or to some extent, is fundamental to being capable of having rational degrees of belief in the first place. What *is* important is that these principles should all seem unimpeachable from a Bayesian perspective.

Now suppose, without loss of generality, that

$$\text{It's raining in Paris, but I don't believe that it is raining in Paris} \tag{9}$$

is believed to a degree of 1 by a person $i$, and that this person has rational degrees of belief. Then, by (17),

$$\mathrm{B}_i(\text{It's raining in Paris, but I don't believe that it is raining in Paris}) \tag{21}$$

and thus, by (18),

$$\mathrm{B}_i(\text{It's raining in Paris}), \tag{22}$$

whence, by (20),

$$\mathrm{Pr}_i(\text{I believe that it's raining in Paris}) > 0. \tag{23}$$

But from the assumption that person $i$ believes (9) to a degree of 1, and given that her degrees of belief are supposed to be probabilities, it also follows that

$$\mathrm{Pr}_i(\text{I don't believe that it's raining in Paris}) = 1 \tag{24}$$

and hence that

$$\mathrm{Pr}_i(\text{I believe that it's raining in Paris}) = 0, \tag{25}$$

which together with (23) contradicts our assumption that the person has rational degrees of belief. So a person respecting principles (17)–(20) cannot believe (9) and sentences of the same form to a perfect degree without violating the axioms of probability.[31] Hence, there exist sentences that are consistent yet cannot possibly be rationally believed to a perfect degree, contrary to Bayesian orthodoxy.[32,33]

## References

Adler, J. (2002). *Belief's own ethics*. Cambridge, MA: MIT Press.
Adler, J. (2006). Withdrawal and contextualism. *Analysis, 66*, 280–285.
BonJour, L., & Sosa, E. (2003). *Epistemic justification*. Malden, MA: Blackwell Publishing.
Christensen, D. (2004). *Putting logic in its place*. Oxford: Oxford University Press.

[31] To derive the same conclusion for sentences of the form of (13), one needs instead of (20) the principle $\mathrm{B}_i(\varphi) \Rightarrow \mathrm{Pr}_i\big(\mathrm{K}_i(\varphi)\big) > 0$. The proof is straightforward and left to the reader. Of course this does not add anything to the general point we are making.

[32] It is certainly true, as the referee remarked, that if I am in excruciating pain, then I cannot rationally have a degree of belief of 1 that I am in no pain at all, even though this is consistent, and that Bayesians must already acknowledge as much. But it is still *possible* rationally to believe to a degree of 1 that I am in no pain at all (namely, when I am not in pain). In the case of sentences like (9), by contrast, a conceptual or near-to-conceptual impossibility bars us from ever believing them rationally to such a degree—which is the point of the present section.

[33] I am greatly indebted to Jonathan Adler and to an anonymous referee for extremely helpful comments on earlier versions of the paper. Thanks also to Paolo Casalegno and Leon Horsten for useful feedback.

DeRose, K. (1991). Epistemic possibilities. *Philosophical Review, 100*, 581–605.
DeRose, K. (1996). Knowledge, assertion, and lotteries. *Australasian Journal of Philosophy, 74*, 568–580.
DeRose, K. (2002). Assertion, knowledge, and context. *Philosophical Review, 111*, 167–203.
Douven, I. (2002). A new solution to the paradoxes of rational acceptability. *British Journal for the Philosophy of Science, 53*, 391–410.
Douven, I. (2005). A principled solution to Fitch's paradox. *Erkenntnis, 62*, 47–69.
Douven, I. (2006). Assertion, Knowledge, and rational credibility. *Philosophical Review, 115*, 449–485.
Douven, I. (2008a). The lottery paradox and our epistemic goal. *Pacific Philosophical Quarterly*, in press.
Douven, I. (2008b). Review of Christensen (2004). *Philosophical Review, 117*, 123–126.
Douven, I., & Uffink, J. (2003). The preface paradox revisited. *Erkenntnis, 59*, 389–420.
Douven, I., & Williamson, T. (2006). Generalizing the lottery paradox. *British Journal for the Philosophy of Science, 57*, 755–779.
Foley, R. (1992). The epistemology of belief and the epistemology of degrees of belief. *American Philosophical Quarterly, 29*, 111–124.
Gillies, D. (2000). *Philosophical theories of probability*. London: Routledge.
Grice, H. P. (1989). Logic and conversation, in his *Studies in the Way of Words* (pp. 22–40). Cambridge MA: Harvard University Press.
Hájek, A. (2003). What conditional probability could not be. *Synthese, 137*, 273–323.
Harman, G. (1986). *Change in view*, Cambridge, MA: MIT Press.
Hintikka, J. (1962). *Knowledge and belief*, Ithaca, NY: Cornell University Press.
Howson, C. (2000). *Hume's problem: Induction and the justification of belief*. Oxford: Clarendon Press.
Hunter, D. (1996). On the relation between categorical and probabilistic belief. *Noûs, 30*, 75–98.
Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford: Clarendon Press.
Johnson, D. (2004). *Truth without paradox*, Lanham, MD: Rowman and Littlefield.
Kemeny, J. (1955). Fair bets and inductive probabilities. *Journal of Symbolic Logic, 20*, 263–273.
Kitcher, P. (1993). *The advancement of science*. Oxford: Oxford University Press.
Kyburg, H. (1970). Conjunctivitis. In: M. Swain (Ed.), *Induction, acceptance and rational belief* (pp. 55–82). Dordrecht: Reidel.
Laudan, L. (1981). A confutation of convergent realism. *Philosophy of Science, 48*, 19–49.
Lewis, D. (1996). Elusive knowledge. *Australasian Journal of Philosophy, 74*, 549–567. (Reprinted in K. DeRose & T. Warfield (Eds.) (1999), *Skepticism* (pp. 220–239). Oxford: Oxford University Press; the page reference is to the reprint.)
Maher, P. (1993). *Betting on theories*. Cambridge: Cambridge University Press.
Makinson, D. (1965). The paradox of the preface. *Analysis, 25*, 205–207.
Mellor, D. H. (1993). How to believe a conditional. *Journal of Philosophy, 90*, 233–248.
Mendola J. (2007). Knowledge and evidence. *Journal of Philosophy, 104*, 157–160.
Meyer, J.-J., & van der Hoek, W. (1995). *Epistemic logic for AI and computer science*. Cambridge: Cambridge University Press.
Moore, G. E. (1962). *Commonplace book: 1919–1953*. London: Allen and Unwin.
Pollock, J. (1986). *Contemporary theories of knowledge*. Totowa, NJ: Rowman and Littlefield.
Psillos, S. (1999). *Scientific realism: How science tracks truth*. London: Routledge.
Schiffer, S. (2005). Paradox and the a priori. In T. Szabó Gendler & J. Hawthorne (Eds.), *Oxford studies in Epistemology* (Vol. I, pp. 273–310). Oxford: Oxford University Press.
Searle, J. (2002). *Consciousness and language*. Cambridge: Cambridge University Press.
Shoemaker, S. (1995). Moore's paradox and self-knowledge. *Philosophical Studies, 77*, 211–228.
Tennant, N. (1997). *The taming of the true*. Oxford: Oxford University Press.
Unger, P. (1975). *Ignorance: A case for scepticism*. Oxford: Clarendon Press.
van Fraassen, B. (1995). Fine-grained opinion, probability, and the logic of full belief. *Journal of Philosophical Logic, 24*, 349–377.
Weatherson, B. (2005). Can we do without pragmatic encroachment? *Philosophical Perspectives, 19*, 417–443.
Wedgwood, R. (1999). The a priori rules of rationality. *Philosophy and Phenomenological Research, 59*, 113–131.
Williams, J. (1996). Moorean absurdities and the nature of assertion. *Australasian Journal of Philosophy, 74*, 133–149.
Williamson, T. (2000). *Knowledge and its limits*. Oxford: Oxford University Press.