

# *Moore-paradoxical belief, conscious belief and the epistemic Ramsey test*

**John N. Williams**

## **Synthese**

An International Journal for  
Epistemology, Methodology and  
Philosophy of Science

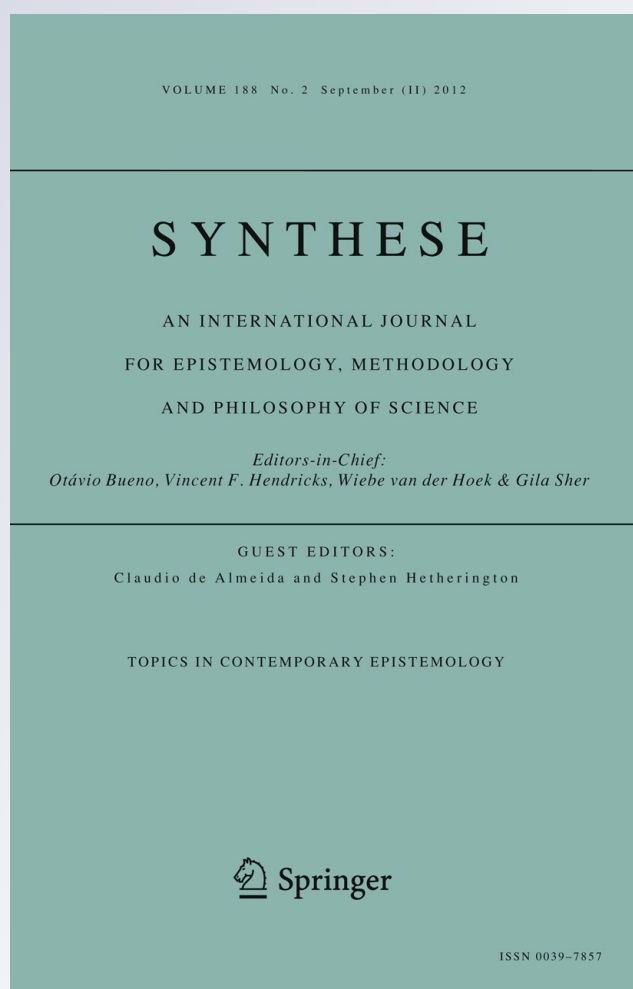
ISSN 0039-7857

Volume 188

Number 2

Synthese (2012) 188:231-246

DOI 10.1007/s11229-011-9925-5



**Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media B.V.. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.**

# Moore-paradoxical belief, conscious belief and the epistemic Ramsey test

John N. Williams

Received: 11 March 2011 / Accepted: 11 March 2011 / Published online: 23 March 2011  
© Springer Science+Business Media B.V. 2011

**Abstract** Chalmers and Hájek argue that on an epistemic reading of Ramsey's test for the rational acceptability of conditionals, it is faulty. They claim that applying the test to each of a certain pair of conditionals requires one to think that one is omniscient or infallible, unless one forms irrational Moore-paradoxical beliefs. I show that this claim is false. The epistemic Ramsey test is indeed faulty. Applying it requires that one think of anyone as all-believing and if one is rational, to think of anyone as infallible-if-rational. But this is not because of Moore-paradoxical beliefs. Rather it is because applying the test requires a certain supposition about conscious belief. It is important to understand the nature of this supposition.

**Keywords** Ramsey · Ramsey test · Moore · Chalmers · Hájek · Conditionals · Paradox · Belief · Conscious belief · Infallibility · Omniscience · Irrationality

## 1 Introduction

Chalmers and Hájek (2007) argue that on an epistemic reading—one that seems reasonable—of Ramsey's test for the rational acceptability of conditionals, it is faulty. They argue that applying the test to each of certain pair of conditionals requires one to think that one is omniscient or infallible, unless one forms irrational Moore-paradoxical beliefs. I show that this claim is false. The epistemic Ramsey test is indeed faulty. Applying it requires that one think of anyone as all-believing and if one is rational, to think of anyone as infallible-if-rational. But this is not because of Moore-paradoxical

---

J. N. Williams (✉)  
School of Social Sciences, Singapore Management University,  
90 Stamford Road, Singapore 178903, Singapore  
e-mail: johnwilliams@smu.edu.sg

beliefs. Rather it is because applying the test requires a certain supposition about conscious belief. It is important to understand the nature of this supposition.

## 2 Chalmers and Hájek on the epistemic Ramsey test

Frank Ramsey (1994, 155 fn) made the pregnant remark

If two people are arguing “If  $p$ , will  $q$ ?” and are both in doubt as to  $p$ , they are adding  $p$  hypothetically to their stock of knowledge, and arguing on that basis about  $q$ ; ... they are fixing their degrees of belief in  $q$  given  $p$ .

The part after the ellipsis suggests that you should accept that if  $p$  then  $q$  iff your subjective conditional probability  $\Pr(q|p)$  is high.<sup>1</sup> But the part before the ellipsis suggests an epistemic test. The entirety of Chalmers and Hájek’s paper is as follows.

Let us take the first sentence the way it is often taken, as proposing the following test for the acceptability of an indicative conditional:

‘if  $p$  then  $q$ ’ is acceptable to a subject  $S$  iff, were  $S$  to accept  $p$  and consider  $q$ ,  $S$  would accept  $q$ .

Now consider an indicative conditional of the form

(1) If  $p$ , then I believe  $p$ .

Suppose that you accept  $p$  and consider ‘I believe  $p$ ’. To accept  $p$  while rejecting ‘I believe  $p$ ’ is tantamount to accepting the Moore-paradoxical sentence ‘ $p$  and I do not believe  $p$ ’, and so is irrational. To accept  $p$  while suspending judgment about ‘I believe  $p$ ’ is irrational for similar reasons. So rationality requires that if you accept  $p$  and consider ‘I believe  $p$ ’, you accept ‘I believe  $p$ ’.

Consider also an indicative conditional of the form

(2) If I believe  $p$ , then  $p$ .

Suppose that you accept ‘I believe  $p$ ’ and consider  $p$ . To accept ‘I believe  $p$ ’ while rejecting  $p$  is tantamount to accepting the Moore-paradoxical sentence ‘Not- $p$  and I believe  $p$ ’, and so is irrational. To accept ‘I believe  $p$ ’ while suspending judgment about  $p$  is irrational for similar reasons. So rationality requires that if you accept ‘I believe  $p$ ’ and consider  $p$ , you accept  $p$ .

According to the Ramsey test, understood as above, it follows that all instances of (1) and (2) are acceptable to a rational subject. But if one accepts all instances of (1), one should accept that one is omniscient. And if one accepts all instances of (2), one should accept that one is infallible. So Ramseyan and Moorean principles entail that rational subjects should accept that they have the epistemic powers of a god.

<sup>1</sup> Many agree that this is a good test. See Adams (1975), Hansson (1992, p. 552) and McGee (1989, p. 487, fn 2). But for dissent, see Morton (2004).

### 3 Preliminary remarks on the nature of the epistemic Ramsey test

Although there are no quotation marks in the right-hand side of Chalmers and Hájek's epistemic reading of Ramsey's test, they sometimes talk of accepting or considering sentences and at other times talk of accepting or considering propositions. In order to avoid objections to this, we may read the test entirely in terms of propositions as

(ET) *If  $p$  then  $q$*  is acceptable to a subject  $S$  iff, were  $S$  to accept that  $p$  and consider whether  $q$ ,  $S$  would accept  $q$

where  $S$  is a rational subject.

This seems like a reasonable reading.<sup>2</sup> (ET) is not a test of the truth of a conditional. Rather it is a test of whether it is rational to accept that the conditional is true. That these may come apart may be seen from a case in which knowing that a certain conditional is false, I have a sensible way of deciding that it is rational for you to accept that it is true, given the background of your actual beliefs, which differ from mine. For example, suppose that I know that it is false that if it is raining then the streets are wet, because I see rain falling outside my window onto a canopy over arching dry streets. However, I also know that you have no view of the outside world and that you believe that there is nothing to prevent rain from wetting the streets. In that case I might sensibly decide that it is rational for you to accept that if it is raining then the streets are wet, because I know that you would assign high subjective probability that the streets are wet, conditionalized upon it raining under such circumstances.

In my own case, deciding that a proposition is true should make me believe that it is true—but not necessarily conversely, because I could sensibly decide that it is worthy of belief without fully endorsing its truth. For example, suppose that you ask me if it will rain soon. Observing heavy dark clouds gathering nearby, I reply, "I can't say it definitely will, but I *think* so". This reply might be sensible if I am aware that my evidence fails to meet standards required for a 'flat-out' assertion, yet justifies belief that falls short of total conviction.

(ET) is not a test of whether one *would* accept that a conditional is true, but a test of whether it is *rational* for one to accept that it is true. A test of whether one would accept that a conditional is true might give a causal role to one's hypothetical acceptance of the antecedent given various backgrounds of actual beliefs. Since the predicted acceptance might not be rational, such a test would not be interesting from the point of view of what one or others *should* accept, although it would be useful if we are interested in predicting what conditionals others do in fact accept given various backgrounds of actual beliefs. The nature of such a causal role of one's hypothetical acceptance of the antecedent, as well as the extent to which determinism enters the picture of such a test, is beyond the scope of this paper.

According to (ET), in order to decide whether *If  $p$  then  $q$*  is acceptable to me as a rational subject, I must suppose that I accept that  $p$ —or as Ramsey might have said, hypothesize that I add the proposition that  $p$  to my existing stock of beliefs, including

<sup>2</sup> For dissent, see Barnett (2008). Barnett's dissent is based upon his assumption that supposing for the sake of argument that  $p$  involves a degree of confidence that  $p$ , under this supposition. I argue in §6 that this assumption is false. See note 16.

those that constitute what I know—while considering whether  $q$ . If I decide that I would accept that  $q$  under this supposition, I should accept *If  $p$  then  $q$* , otherwise I should not accept this conditional.

Another way to read the test is to take Ramsey's mention of adding the antecedent hypothetically to one's stock of *knowledge* as suggesting that I begin the test by supposing that I come to know it. I will return to this suggestion later in §9.

(ET) seems to give a correct verdict on the acceptability of most conditionals. For example, suppose that I apply it to

(3) If I release this ball, then it will fall

while holding the ball in my hand. I start by supposing that I accept that I release the ball, while considering whether it will fall. In order to make what I have hypothetically accepted consistent with my existing background knowledge and beliefs (including my beliefs about gravity and the absence of unusual circumstances), it seems that upon consideration, I must accept that the ball will fall. Thus I should accept that if I drop the ball, then it will fall. This seems to be the correct verdict. However, as we will see in §5 and §6, there must be more to the test than the characterization of it that I have just given.

#### 4 The failure of Chalmers and Hájek's objection to the test

Before examining Chalmers and Hájek's objection to the test, it is worth noting that they employ an idiosyncratic sense of "omniscient". They hold that if one accepts all instances of

(1) If  $p$ , then I believe that  $p$

then one should accept that one is omniscient. This is presumably because they think that someone is omniscient if she believes all truths. But she may believe all truths while also believing some falsehoods. In that case we would not say that she is omniscient, if we think that omniscience is "knowledge of all things". Then we must agree that omniscience includes knowledge of all truths.<sup>3</sup> But now suppose that someone knows all truths while also believing some falsehoods. Whatever she knows, she believes. So for some truth that  $p$ , she believes that  $p$  and believes that not- $p$ . It is plausible that this alone means that she does not know that  $p$ .<sup>4</sup>

Since Chalmers and Hájek take (ET) as a test to be used by a rational agent, they are committed to saying that one must accept that one is both omniscient and rational. It is easy to see that an omniscient and rational being cannot believe a falsehood. In that case, for some truth that  $p$ , she believes that  $p$  and believes that not- $p$ , yet she is

<sup>3</sup> Here I ignore the modal dimension of omniscience. One might say that someone is omniscient just in case it is necessary in some sense or other that she knows any truth. This does not seem to introduce any complication into the following discussion that cannot, for my purposes, be set aside.

<sup>4</sup> Suppose that I know that  $p$ . Then I believe that  $p$ . Now suppose that I also believe that not- $p$ . Whether or not I have *prima facie* justification for that belief, it is an overrider of any *ultima facie* justification I might have for my belief that  $p$ , with the result that I do not know that  $p$  after all. I owe this important point to Claudio de Almeida.

rational in both beliefs. This is impossible, since any justification she has for the one belief will undercut that for the other.

To avoid these problems, let us stipulate that a person is *all-knowing* just in case she knows all truths and *all-believing* just in case she believes all truths. Thus if a person is all-knowing then she is all-believing, but not necessarily conversely. Let us also say that she is *infallible* just in case all her beliefs are true.<sup>5</sup> Any all-believing rational subject is infallible. To see this, suppose for *reductio* that I am rational and I hold at least one mistaken belief. Then for some proposition  $p$ , it is true both that  $p$  and that I believe that not- $p$ . So it is true that  $p$ . But since I am all-believing, I believe all truths, so I believe that  $p$ . Thus I believe that  $p$  and that I believe that not- $p$ . Therefore I am not rational after all, in virtue of holding contradictory beliefs, because any non-overridden evidence for the one belief is non-overridden counterevidence for the other. A second argument is that since I am all-believing, I believe the truth that  $p$  and I believe that not- $p$ . But if I am rational, this is impossible, as I explain in §5. Since any all-knowing being is all-believing, any all-knowing rational being is also infallible.

Given this preliminary clarification, Chalmers and Hájek can be seen charitably as arguing that unless you form irrational Moore-paradoxical beliefs, (ET) compels you to think that you are both all-believing and infallible.

We are now in a position to diagnose the flaw in Chalmers and Hájek's argument; when applied to (1) or (2), (ET) need not saddle a rational agent with Moore-paradoxical beliefs. Suppose that I accept that  $p$  and consider whether I believe  $p$ . Chalmers and Hájek claim that to accept that  $p$  while rejecting *I believe that p* is "tantamount" to accepting the Moore-paradoxical,  $p$  and *I do not believe that p*. This claim is plausible. I accept that  $p$  at  $t$  only if I have formed the belief that  $p$  at  $t$ . I reject the proposition that  $p$  at  $t$  only if I have formed the belief that not- $p$  at  $t$ . So after I have accepted that  $p$  and rejected the proposition that *I believe that p*, I believe that  $p$  and I also believe that I don't believe that  $p$ . Chalmers and Hájek now need

*Belief-collection:* If I believe that  $p$  and I also believe that  $q$ , then I believe that both ( $p$  &  $q$ )

to deliver the result that I have the irrational omissive Moore-paradoxical belief that  $p$  & *I don't believe that p*.<sup>6</sup> There is reason to doubt that belief-collection always holds, even for a rational subject. For one thing, *Searle's Principle* is plausible: my belief that  $p$  requires my ability to think the thought that  $p$  (1992, pp. 155–162). Surely I cannot think the 'fat' thought that conjoins the content of *all* the beliefs that I hold about the world.<sup>7</sup> Nonetheless belief-collection appears to hold in the context of applying the

<sup>5</sup> This ignores the modal nature of infallibility. One might say that a person is infallible just in case it is necessary in some sense that whatever she believes is true. I set this complication aside for present purposes.

<sup>6</sup> "Omissive" because it reports the omission of a specific true belief. This useful term is coined by Sorensen (1988, p. 16).

<sup>7</sup> One reason for this is the fact that my beliefs are uncountably many; I believe, at least unconsciously, that I live at least ten miles away from Nelson's Column. I also believe, at least unconsciously, that I live at least eleven miles away from Nelson's Column ... I have the ability to think each thought in an infinite series, but I do not have the ability to think the thought of their conjunction, for that would be a thought that I could never finish thinking. See Williams (2006a, §4).

epistemic test; I am *considering* whether  $q$  under the *supposition* that I accept that  $p$ , so I am *ipso facto* thinking the thought that  $p$  while also thinking the thought that  $q$ . This point about consciousness will become important in §6.

But another way for a rational subject to avoid accepting (1)'s consequent is to suspend judgement about it. Chalmers and Hájek claim that to accept  $p$  while suspending judgement about *I believe  $p$*  is “irrational for similar reasons”. This claim is false. By suspending judgement about a proposition, I *withhold belief* that it is true (and also withhold belief that it is false). So after I have accepted that  $p$  and suspended judgement about *I believe  $p$* , I believe that  $p$ . In addition, I do not believe that I believe that  $p$  (and I do not believe that I don't believe that  $p$ ). Because this additional fact is about my *lack* of belief, there is no way for belief-collection to saddle me with the omissive Moore-paradoxical belief, because there is no pair of beliefs to collect.<sup>8</sup> All I am left with is the belief that  $p$ , which might be perfectly rational.

Now consider (2). Suppose that I accept that *I believe  $p$*  and consider whether  $p$ . Chalmers and Hájek claim that to accept *I believe  $p$*  while rejecting  $p$  is tantamount to accepting the Moore-paradoxical, *not- $p$  and I believe that  $p$* . This is also plausible. After I have rejected  $p$  and accepted *I believe  $p$* , I believe that not- $p$  and I also believe that I believe that  $p$ . So given belief-collection, I hold the irrational commissive Moore-paradoxical belief that (not- $p$  & I believe that  $p$ ).<sup>9</sup>

But to accept *I believe  $p$*  while suspending judgement about  $p$  is *not* similarly irrational, contrary to Chalmers and Hájek's claim. After I have suspended judgement about  $p$  and accepted *I believe  $p$* , I do not believe that  $p$  (and I also do not believe that not- $p$ ). In addition, I believe that I believe that  $p$ . Because this *first* fact is about my *lack* of belief, there is no way for belief-collection to saddle me with the commissive Moore-paradoxical belief because once again there is no pair of beliefs to collect. All I am left with is my belief that I believe that  $p$ , which again might be perfectly rational.

In the case of (1) it might be objected that as a rational subject I cannot believe that  $p$  while not believing that I believe that  $p$ , because as a rational subject I must obey

*Belief-Introduction*: If I believe that  $p$  then I believe that I believe that  $p$ .

To deal with (2) in the same way the objector would have to claim that as a rational subject I cannot believe that I believe that  $p$  while not believing that  $p$ , because as a rational subject I must obey the converse of belief-introduction, namely

*Belief-Elimination*: If I believe that I believe that  $p$  then I believe that  $p$ .

Chalmers and Hájek do not make this objection. Even if it provides a satisfactory argument for the irrationality of accepting the antecedent of (1) or (2) while suspending

<sup>8</sup> This is consistent with the fact that to make the conjunctive *assertion* that “ $p$  & I do not believe that  $p$ ” or “ $p$  & I have no beliefs either way about whether or not  $p$ ” is to make an omissive Moore-paradoxical assertion. Chalmers and Hájek do not even mention assertion, let alone try to explain the failure of (ET) in terms of Moore-paradoxical assertion. I will follow them in avoiding the notion of assertion. It is difficult to see how assertion may be relevant to applying (ET)—which itself eschews the notion—especially when I apply it in silent solitude. Moreover the notion is messy, because the relation of assertion to belief is far from clear. See (Williams 2006b, §4 and §8).

<sup>9</sup> “Commissive” because it reports the commission of a specific mistake in belief. Again, this term originates from Sorensen (1988, p. 16).



belief in its consequent, the irrationality is not that of Moore-paradoxical beliefs. So for Chalmers and Hájek to endorse the objection would be for them to abandon their argument that unless you form irrational Moore-paradoxical beliefs, (ET) compels you to think that you are both omniscient and infallible.

Moreover it is controversial whether either belief-introduction or belief-elimination is true even of a rational subject. Belief-introduction entails an infinite regress of beliefs; in believing that  $p$  I must believe that I believe that  $p$ . So I must believe that I believe that I believe that  $p$  ... and *ad infinitum*. This series will eventually contain a member too complex to be thought of by any human subject. So by Searle's Principle, no human subject may believe this member. But this does not seem to be a failure of rationality, but rather a psychological limit of knowledge.

Belief-introduction is a principle of all-believing, restricted to one's beliefs, while belief-elimination is a principle of infallibility applied to one's beliefs. It is unclear whether failing either principle counts as human epistemic irrationality. Against belief-introduction, suppose that I have the repressed belief that my mother was an adulteress, yet when asked if I hold this belief, reply sincerely that I don't have this belief. Given that I do not hold contradictory beliefs about whether I hold the repressed belief, I don't believe that I do hold it, thus falsifying belief-introduction. This makes me ignorant, and thus less than ideal.

It might be said that epistemic rationality is a matter of what beliefs I form on the basis of what else I know and rationally believe, not a matter of how much knowledge I have, or my infallibility in acquiring rational beliefs that help make up this basis. Those that follow Richard Foley in separating ignorance from irrationality in belief might say that ignorance alone does not count as epistemic irrationality. It might be said in support of this dichotomy that someone who has irrational beliefs must be blameworthy in some way, but someone who is ignorant may not be blameworthy in any way.

However, this is consistent with blameworthy ignorance enabling irrational belief. Suppose that I am culpably ignorant of the fact that whether a roulette ball lands on black or red is causally uninfluenced by past outcomes. This might at least partly explain why I form the irrational belief that the longer an unbroken series of red outcomes become, the more likely it becomes that the next outcome will be black.

But on the other hand again, my ignorance of my repressed belief does not seem to be necessarily culpable.

Against belief-elimination, take a different case of repressed belief. Suppose that I have taken myself to believe that my ex-lover is unworthy of respect. Yet faced with someone who criticizes her, I find myself defending her, and in doing so, I realize that I did not believe that she is unworthy of respect. It is unclear whether my epistemic rationality was compromised before this realization.<sup>10</sup>

Clearly much more needs to be said about the nature of epistemic rationality and its relation to ignorance. This brief discussion serves only to cast doubt on whether belief-introduction and belief-elimination must be true of rational subjects, a question that is, strictly speaking, peripheral to my purposes. A more germane question is whether

<sup>10</sup> I owe this example to T. Brian Mooney.

a rational agent must know or rationally believe what actual beliefs and knowledge she has. If not, then it is entirely possible that a rational agent cannot apply the test, because she cannot hypothetically adjust her supposed acceptance of the antecedent of the conditional for consistency with her existing background knowledge and beliefs if she does not know or rationally believe what this background is. Let us now put these questions aside and instead take a closer look at the test.

## 5 More on the nature of the epistemic Ramsey test

In §3 I said that there must be more to my characterization of how (ET) is applied to

(3) If I release this ball, then it will fall

when holding the ball in my hand. Since I am supposing myself to be a rational agent, I must suppose, hypothetically, that I accept that I release the ball while considering whether it will fall, and then decide whether rationality would oblige me to accept that it will fall. If I do accept that it will fall, then I do so, as Ramsey puts it, “on the basis” of accepting that I release it. What does “on the basis” mean? We have already discounted a causal basis as uninteresting. The answer seems to be that after supposing that I accept that I release it, I hypothetically revise my supposed-plus-actual beliefs and knowledge in a rational manner for coherence.<sup>11</sup> It is difficult to see what else it could mean, because (ET) asks what *effect* coming to believe that I release the ball would have on my actual beliefs—in particular whether coherence would oblige me to believe that it will fall. The nature of epistemic coherence is both complex and controversial, but it clearly includes consistency. A set of beliefs is coherent only if the propositional objects of the beliefs are consistent. It will turn out that adjusting for consistency is enough to diagnose problems with the test.

Ramsey suggests a test that is to be applied when one is in doubt as to whether the antecedent is true. But a better test should also be applicable when one knows that the antecedent is false. Even if I know that I will not drop the ball, I should still accept that under normal conditions, (3) is true. However, as I have characterized the test, there is a snag. I start by supposing that I accept that I release the ball. I then hypothetically add this belief to my actual beliefs and knowledge. This includes my knowledge that I do not release the ball. Next I adjust my actual-plus-hypothetical beliefs for consistency. My hypothetical belief that I release the ball is at odds with my actual knowledge that I do not release it. I am now faced with a choice. Do I hypothetically give up my

<sup>11</sup> Here another problem with (ET) emerges. Suppose that you apply it to

If Oswald did not kill Kennedy then someone else did

and that you actually believe that Kennedy was killed and that nobody other than Oswald killed Kennedy. You now hypothetically add the belief that Oswald did not kill Kennedy to your actual beliefs. You have hypothesized a set of inconsistent beliefs. One way to resolve the inconsistency is to hypothetically replace your belief that nobody other than Kennedy killed Oswald with the belief that someone other than Oswald killed Kennedy. So you should accept the conditional. The problem however is that another way to resolve the inconsistency is to hypothetically give up your belief that Kennedy was killed. Since there are typically a variety of ways in which one can adjust one's beliefs to render them consistent, this leaves the test highly indeterminate.

belief that I release the ball? Or do I hypothetically give up my belief—and with it my knowledge—that I do not release it?

This snag is minor. The answer must be that I should suppose that I do not know that I do not release the ball, because the alternative is to give up the very belief in the antecedent that I have supposed myself to have. That would be pointless, because I have supposed myself to have it in order to see what effect it would have upon my actual beliefs and knowledge, not *visa versa*.

A genuine difficulty emerges for conditionals with inconsistent antecedents such as

(4) If  $2 + 2 = 5$  then Singapore is an island.

I should see that any such conditional is only vacuously true. Or to put it differently, I cannot say that any such conditional is false. This suggests that it would be prudent, at the start of the test, to check whether the antecedent I have supposed myself to accept is consistent. But consider

(5) If there is a highest prime number then there is no highest prime number.

I might sensibly judge that this is true after being shown a *reductio* of the antecedent. The fact that unlike (4), (5) may be seen as having the form *if p then not-p* suggests a reason for continuing with the test, because depending upon their content, such conditionals might be testable for acceptability via *reductio*. So one might expect that the test would show me, as a rational agent, that (5) is not acceptable. Not so. As before, I start by supposing that I accept that there is a highest prime number. I then hypothetically add this belief to my actual beliefs and knowledge. I now adjust my actual beliefs and knowledge for consistency with this belief.

But now a real problem emerges. Suppose that I decide that this adjustment would leave me with the belief that there is no highest prime number. Then the test tells me that I should accept (5). The trouble is that I should now see that the antecedent is necessarily false. If I were performing *reductio* as a test of the truth of (5) then this would be no problem, because then I may simply discharge the supposition that its antecedent is true. But as a test of the acceptability of (5), I am still supposing myself to accept that (5)'s antecedent is true—which is irrational. So all that I am entitled to say is that I 'should' accept (5) if I am irrational. In other words rationality compels me to accept (5) if I am irrational enough to accept its antecedent. One might wonder whether this verdict is even coherent.

On the other hand, if I reject the consequent of (5) then I must reject (5)—but again only under the supposition that I irrationally accept its antecedent. This time, all I am entitled to say is that I 'should' reject the conditional if I am irrational. So the test provides no guidance to me as a rational agent.

The same problem arises with conditionals with Moore-paradoxical antecedents such as

(6) If I fail to believe the truth that it is raining, then I do not believe that it is raining.

This conditional is true. A rational subject should accept that it is true. But although the antecedent is consistent, I cannot rationally accept it. It might indeed be true that

I fail to believe the truth that it is raining. And I might accept it. But it cannot remain true once I accept it. Accepting the antecedent means accepting that it is raining, but then it becomes false that I fail to believe it—because whatever I accept I believe, although not conversely. The episode of acceptance I have supposed myself to have is *self-falsifying*; what I accept might be true and I might accept it, but it cannot be true once I accept it, a fact that I may be reasonably expected to see.<sup>12</sup>

As a rational agent I am bound by the norm of not forming—or continuing to have—specific beliefs that I can be reasonably expected to see are false.<sup>13</sup> I contravene this norm in having an omissive Moore-paradoxical belief. I have shot myself in the foot, purely as a result of my act of belief, something I am in a position to see.

de Almeida (2001, p. 43; 2007, p. 56) makes the important objection that seeing this is not easy. Symbolising ‘I believe that  $p$ ’ as ‘ $Bp$ ’, I have to use belief-distribution on  $B(p \ \& \ \sim Bp)$  to derive  $Bp \ \& \ B\sim Bp$ , followed by  $\&$ -elimination to derive  $Bp$ , followed by the rule that a conjunction is false if one of its conjuncts is, to derive  $\sim(p \ \& \ \sim Bp)$ —no mean feat.

I reply that although the self-falsifying nature of the belief is not obvious, we should not make too much of this fact. To see this, consider the case in which I believe that it is both raining and not raining. To see that what I believe is necessarily false I have to reason “If it is true both that it is raining and not raining then it is true that it is raining. But then it is false that it is not raining. And that means that it is false that it is both raining and not raining. So if it is true both that it is raining and not raining then it is false both that it is raining and not raining. So it *must* be false both that it is raining and not raining”. In other words I must use  $\&$ -elimination, followed by double-negation, followed by the rule that a conjunction is false if one of its conjuncts is, followed by the rule that anything that is false if true is necessarily false. If the first case is no mean feat, so is this.

Following Chan (2010, p. 219, fn 17) I might also add that what is required is merely that I can answer the question of whether my belief is self-falsifying if I were to consider it.<sup>14</sup>

Having supposed myself to accept the antecedent of (6) I must now continue with the test. Having hypothetically adjusted my actual knowledge and beliefs for consistency with my Moore-paradoxical belief, I now decide whether this adjustment would leave me with the belief that I do not believe that it is raining. Again it makes no difference what the answer is. If I would accept that I do not believe that it is raining then I must accept (6)—but only under the supposition that I irrationally accept its

<sup>12</sup> Chan (2010) argues that one is not irrational in having an omissive belief that ( $p$  and I do not believe that  $p$ ) unless one could be reasonably expected to see that believing it falsifies its content.

<sup>13</sup> The reference to *specific* beliefs avoids the objection that rationality does not require *all* one’s beliefs to be true. For example, some claim the preface paradox to be a case in which rationality demands that one has inconsistent beliefs.

<sup>14</sup> In the commissive case, my belief that ( $p$  & I believe that not- $p$ ) is self-falsifying unless I both believe that  $p$  and also believe that not- $p$ . It is reasonable to expect me to recognize this, at least in the sense that should someone explain this truth to me correctly, I should accept it. As a rational agent I am not only bound by the norm of avoiding specific self-falsifying beliefs. I am also bound by the norm of not forming—or continuing to hold—a pair of overtly contradictory beliefs. I must contravene one norm or other. I have shot myself in one of two feet, purely because of my act of belief.

antecedent. All I am entitled to say is that I ‘should’ accept (6) if I am irrational. On the other hand, if I would reject the consequent of (6) then I must reject (6)—but again only under the supposition that I irrationally accept its antecedent. Again, all I am entitled to say is that I ‘should’ reject the conditional if I am irrational. Either way I have no guidance.

Perhaps we might be prepared to live with this result, holding the sanguine attitude that no guidance in a few strange cases is a price worth paying for guidance in others. But worse is to come from (1) and (2). To see this however, we have to return to the nature of the test a second time.

## 6 Why applying the test involves supposing conscious belief

In fact (ET) compels a subject, *rational or not*, to accept any conditional of the form of (1). However this is not because of Moore-paradoxicality, but because of the role of supposing conscious belief in applying the test. Imagine that I am ready to apply the test to (1). I might have no actual beliefs one way or the other about whether the antecedent is true. I might even know that the antecedent is false. So I must start with a supposition about the antecedent.

There are various senses of “suppose”. You ask your friend if Biden will become the next President of the U.S.A. She frowns a bit and then says in an uncertain tone of voice “I suppose not”. She is making a hesitant assertion that Biden will not become the next President of the U.S.A., one that signals her unconfident acceptance. This is not the sense of “suppose” that is relevant to our purposes. We have in mind “suppose for the sake of argument” or more accurately still “suppose as a hypothesis”.

This sense of “suppose” is slippery. It is easy to confuse supposing that  $p$  with supposing that one accepts that  $p$ . To see that there is a difference, consider the surprise exam paradox. A teacher announces that there will be an exam exactly one day of next week, but that the student will not know the day before the exam that the exam is the next day. The student objects that supposing that the announcement is true, he would know at the close of Thursday that the exam is the next day. Having ruled out Friday, he reiterates this argument to rule out all days, thus apparently showing that the announcement must be false. But this *reductio* is easily seen to fail, because the student could have been asleep when the announcement was made. So what he supposes in supposing that the announcement is true, is consistent with the possibility that he has no mental attitudes whatsoever towards this supposed truth. But isn’t that too easy a response to the paradox? Surely to get his argument off the ground, the student has to suppose that he *accepts* the truth of the announcement.<sup>15</sup>

Take another example. You ask your friend, “Suppose that Obama will be assassinated. Then do you accept that Biden will become the next President of the U.S.?” It is entirely natural of her, given this context, to take you as asking her to suppose that she knows, or has learned, or has accepted, that Obama will be assassinated. But this natural move is not, strictly speaking, entailed by supposing that Obama will be assassinated. For you could have asked her instead “Suppose that Obama will be

<sup>15</sup> See Williams (2007, p. 77) and Sainsbury (1995, p. 93). Quine (1953, p. 66) makes a similar point.

assassinated but you will never be aware of this. Then will you be confident that Biden will become the next President of the U.S.A.?" To which she would reply "Of course not".

Or take another case again. You ask her "Suppose that ice is heavier than water. Then will ice sink in water?" To which she would answer "Yes". Then you ask her "Suppose that ice is heavier than water, but you are unaware of this fact. Then will ice sink in water?" To which she would again answer "Yes". Now you ask her "Suppose that you accept that ice is heavier than water. Then will you accept that ice will sink in water?" She would yet again answer "Yes". But if you ask her "Suppose that ice is heavier than water but you don't accept that it is. Will you accept that ice will sink in water?", then she would answer "No".

The upshot is that supposing that  $p$  does not essentially involve supposing that one accepts that  $p$ .<sup>16</sup> In applying the test to (1) I cannot merely suppose that the antecedent is true. That would be the start of a test of the truth of the conditional, not the start of a test of whether I should accept that it is true. As a test of the acceptability of a conditional (ET) cannot start with

*If  $p$  then  $q$  is acceptable to a rational subject  $S$  iff, were it the case that  $p$  ....*

This because the supposition that  $p$ —for example, that ice is heavier than water—is consistent with the possibility that  $S$  does not believe that  $p$ . In that case the truth that  $p$  would have no effect at all on  $S$ 's actual beliefs. Yet whether there are such effects, and what they are, is precisely what we must envisage. Thus in applying the test to (1) I must suppose, hypothetically, that I *accept* that  $p$  while considering whether  $q$ , and then decide whether rationality would oblige me to accept that  $q$ . But surely "...were I to accept  $p$ " cannot be read as "...were I to form the *unconscious* belief that  $p$ ". My belief that  $p$  could hardly be a basis for me as a rational subject to revise my actual beliefs for consistency with the belief that  $p$ , were I unaware of holding that belief.

Thus in applying the test I must suppose, hypothetically, that I *consciously believe* that  $p$  against the background of my actual beliefs while considering whether  $q$ , and then decide whether rationality would oblige me to accept that  $q$ . So it is no accident that I must suppose that I *accept* that  $p$ , because acceptance is an act of judgement and is therefore a conscious belief.<sup>17</sup>

<sup>16</sup> Barnett (2008) assumes that supposing for the sake of argument that  $p$  involves a degree of confidence that  $p$ , under this supposition. But imagine that you ask your friend "Suppose that ice is heavier than water but you are unaware that it is. Are you at all confident that ice is heavier than water?" She would answer "No". Barnett must say that the supposition that it is raining but you are not at all confident that it is, is incoherent. Plainly it is not.

<sup>17</sup> It is important to note that "conscious" is not meant as an antonym of "dispositional". Essentially, a dispositional belief is one that you will form when prompted. No belief, while only dispositional, has been formed. Since there is no belief to be aware of, dispositional beliefs are unconscious. But not all unconscious beliefs are dispositional. One may be unaware of beliefs that one has just formed, which one might call "occurrent" or that one has had for some time. For example, you may have formed rapidly changing perceptual or repressed beliefs of which you are unaware.

## 7 The nature of conscious belief

I have used “aware” as a synonym of “conscious”. This seems unobjectionable. Where  $N$  is a noun, whether it denotes an object, such as a coin in my pocket, or a mental state, surely I am conscious of having  $N$  just in case I am aware of having  $N$ . To say that I am conscious of a belief, fear, suspicion or toothache (or a coin in my pocket) is just to say that I am aware of having it. Moreover, I cannot be aware that I have  $N$  unless I believe that I have  $N$ . Where  $N$  is the belief that  $p$ , what follows is

*Rosenthal's Principle:* If I consciously believe that  $p$  then I both believe that  $p$  and I also believe that I believe that  $p$ .

As Rosenthal (1997) observes, I am conscious of my belief that  $p$  only if I have a “suitable” thought about that belief. Since my mere supposition that I hold a belief would not make me aware of a belief that I really do hold, the suitable second-order thought in question must be a belief.

## 8 Back to the epistemic Ramsey test again

We may now return to the epistemic test. It follows from §6 that in applying the test, I must start by supposing that I come to consciously believe the antecedent. I hypothetically fit my actual knowledge and belief to this supposed belief for coherence, minimally understood as consistency. I consider whether I would accept the consequent. If I decide that I would, I accept the conditional, and if I decide that I would reject it, then I reject the conditional.

This fuller understanding of the test does not affect its verdict on (4), (5) or (6). In the case of (6), supposing myself to consciously believe its antecedent amounts to supposing myself to consciously hold a Moore-paradoxical belief, which is to suppose myself to be irrational (Williams 2010, §5). However it does make a difference to the test's verdicts on (1) and (2).

To apply it to (1) I must suppose that I accept that  $p$  while considering whether I believe that  $p$ . We saw in §6 that this involves not only supposing that I have come to believe that  $p$  but also involves supposing that I have come to form the conscious belief that  $p$ . By Rosenthal's principle, this involves supposing that I have come to believe that I believe that  $p$ . Since I am supposing that I have come to believe that I believe that  $p$  while *considering* whether I believe that  $p$ , I am *ipso facto* supposing that I have accepted that I believe that  $p$ . So my supposition that I accept the antecedent of (1) *already includes* the supposition that I accept its consequent. Thus I must accept (1).

This remains the case whether or not I am a rational agent, since my acceptance of (1) is entailed by no more than my beginning to apply the test. The question of consistency does not get raised. The problem is not just that (ET) delivers the wrong verdict in *one* case. Since it tells us to accept *all* conditionals of the form of (1), it is committed to the absurd prescription that we think of ourselves as all-believing.<sup>18</sup>

<sup>18</sup> Or at least that we think of ourselves as *potentially* all-believing, because we must accept *any* conditional of the form of (1) that we *might* consider.

It might be objected that (ET) escapes this fate when I apply it to someone else. Suppose that I want to decide whether *if p then q* is acceptable to you as a rational agent. In that case I do not need to suppose that you have conscious beliefs.

I have two replies. Firstly, even if this claim is correct, (ET) still compels me to think of myself as all-believing. In fact it compels anyone to think of herself as all-believing. That is surely enough to sink the test. Secondly, I simply deny the claim. In applying the test to you I must suppose, hypothetically, that you accept that *p* against the background of your actual beliefs, while you consider whether *q*, and then decide whether rationality would oblige you to accept that *q*. But your belief that *p* could hardly be your basis for revising your actual beliefs for consistency with it, were you unaware of holding that belief. I am in a position to see this. So in applying the test to you I must suppose, hypothetically, that you *consciously believe* that *p* against the background of your actual beliefs while considering whether *q*, and then decide whether rationality would oblige you to accept that *q*. When the conditional in question is

(1') If *p* then you believe that *p*

then rationality drops out of the picture. I have already supposed that you consciously believe that *p* while you consider whether you believe that you believe that *p*. So I have *ipso facto* already supposed that you accept (1')'s consequent. Thus (ET) compels me to think of you as all-believing, even if I know that you are irrational. In fact it compels anyone to think this of anyone.

It would be a theoretical nicety if we could use a parallel argument that (ET) requires us to accept all instances of (2). Then we could claim that (ET) requires any subject, however irrational, to think of herself as infallible. However, no such argument may be given. On the supposition that I come to form the conscious belief that I believe that *p*, all that Rosenthal's principle shows is that I have supposed that I believe that I believe that *p*, not that I have supposed that I believe that *p*.

However, we have already seen in §2 that any rational all-believing subject is infallible. Since this is a fact that a rational subject may see, and since (ET) compels one to think of anyone as all-believing, it also compels any rational subject to think of anyone as infallible-if-rational.

## 9 The test in terms of knowledge

In §3 we noted the suggestion that one begins the test by supposing that one comes to know the antecedent. Then one hypothetically adjusts one's actual beliefs and knowledge for coherence—minimally understood as consistency—with this supposed knowledge. If one decides that one would be left with a belief in the consequent, then one should accept the conditional, and if one decides that one would be left with the belief in the negation of the antecedent, one should reject it.

This version of the test escapes none of the difficulties that afflict its predecessor. In the case of (4) or (5) I must start by supposing something impossible, since I cannot know that  $2 + 2 = 5$  or that there is a highest prime number. There can be no knowledge to which I hypothetically adjust my actual beliefs and knowledge for consistency. Thus the test delivers no verdict on the rational acceptability of these conditionals. In the



case of (6) I have to suppose that I know an ommissive Moore-paradoxical proposition. This too is impossible, despite the fact that such a proposition might be true. If I know that I fail to believe the truth that it is raining, then I know that it is raining. But I also know that I do not believe that it is raining, so I do not believe that it is raining. But since whatever I know, I believe, it follows that I do not know that it is raining. So I do and do not know that it is raining. Contradiction. Again there can be no knowledge to which I hypothetically adjust my actual beliefs and knowledge for consistency. Thus the test delivers no verdict on the rational acceptability of (6).

In the case of (1), “...were I to accept  $p$ ” cannot be read as “...were I to come to *unconsciously* know that  $p$ ”. My supposed knowledge that  $p$  could hardly be a basis for me as a rational subject to revise my actual beliefs for consistency with it, were I unaware of possessing that knowledge. Thus in applying the test I must suppose that I *consciously know* that  $p$  against the background of my actual beliefs while considering whether I believe that  $p$ . But whatever I am aware of knowing, I am aware of believing. So I have already supposed myself to have come to believe that I believe that  $p$  while considering whether I believe that  $p$ . Thus I am *ipso facto* supposing that I have accepted that I believe that  $p$ . So my supposition that I know the antecedent of (1) already includes the supposition that I accept its consequent. Thus I must accept (1).

I conclude for new reasons that the epistemic Ramsey test is faulty. It should not require us to conceive of ourselves in this way.

**Acknowledgments** I am grateful to Stephen Hetherington, Claudio de Almeida and Mark Nowacki for many helpful suggestions on previous drafts.

## References

- Adams, E. (1975). *The logic of conditionals*. Dordrecht: Reidel.
- Barnett, D. (2008). Ramsey + Moore  $\neq$  God. *Analysis*, 68, 168–174.
- Chan, T. (2010). Moore’s paradox is not just another pragmatic paradox. *Synthese*, 173, 211–229.
- Chalmers, D., & Hájek, A. (2007). Ramsey + Moore = God. *Analysis*, 67, 170–172.
- de Almeida, C. (2001). What Moore’s paradox is about. *Philosophy and Phenomenological Research*, 62, 33–58.
- de Almeida, C. (2007). Moorean absurdity: An epistemological analysis. In M. S. Green & J. N. Williams (Eds.), *Moore’s paradox: New essays on belief, rationality, and the first person* (pp. 53–76). Oxford: Oxford University Press.
- Hansson, S. (1992). In defense of the Ramsey test. *Journal of Philosophy*, 89, 522–540.
- McGee, V. (1989). Conditional probability and compounds of conditionals. *Philosophical Review*, 98, 485–541.
- Morton, A. (2004). Against the Ramsey test. *Analysis*, 64, 294–299.
- Quine, W. V. O. (1953). On a so-called paradox. *Mind*, 62, 65–67.
- Ramsey, F. P. (1994). General propositions and causality. In D. H. Mellor (Ed.), *Philosophical papers by F.P. Ramsey* (pp. 145–163). Cambridge: Cambridge University Press.
- Rosenthal, D. M. (1997). A theory of consciousness. In N. J. Block, O. Flanagan & G. Güzeldere (Eds.), *The nature of consciousness: Philosophical debates* (pp. 729–753). Cambridge, MA: MIT Press and Bradford Books.
- Sainsbury, M. (1995). *Paradoxes* (2nd ed.). Cambridge: Cambridge University Press.
- Searle, J. (1992). *The rediscovery of the mind*. Cambridge, MA: MIT Press.
- Sorensen, R. A. (1988). *Blindspots*. Oxford: Clarendon Press.
- Williams, J. N. (2006a). Moore’s paradoxes and conscious belief. *Philosophical Studies*, 127, 383–414.

- Williams, J. N. (2006b). Wittgenstein, Moorean absurdity and its disappearance from speech. *Synthese*, 149, 225–254.
- Williams, J. N. (2007). The surprise exam paradox: Disentangling two *reductios*. *Journal of Philosophical Research*, 32, 67–95.
- Williams, J. N. (2010). Moore's paradox, defective interpretation, justified belief and conscious belief. *Theoria*, 76, 211–248.