

NOTES AND COMMENTS

SELF-DECEPTION, RELIGIOUS BELIEF, AND THE FALSE BELIEF CONDITION

In a recent paper in this journal, Juha Räikkä¹ addressed the question of whether religious beliefs are typically self-deceptive, taking the belief in life after death as a representative example. He claims to want to ‘argue against the claim that religious beliefs are always self-deceptive (religions being forms of collective self-deception) and also against the moderate view that some religious beliefs are based on self-deception, while others are not’ (pp.513–514). His argument is based on a conceptual point about self-deception, namely, that self-deception must involve false belief. I will try to show that this argument can be seen to fail once we acknowledge the distinction between being self-deceived *in* believing that *p*, and deceiving oneself *into* believing that *p*.

Before we examine Räikkä’s argument, let me briefly outline his views on both self-deception and religious belief. Räikkä mentions two basic elements of self-deception to be as follows: ‘When a person deceives herself, she (1) believes something that she wants to be true, but (2) the object of belief is not properly supported by evidence’ (p.515). By (2) he presumably means that the belief is not supported by the evidence that she was acquainted with or that was available to her. Räikkä holds this true of self-deception at least in its ‘straight’ variety.² I take these points to be roughly correct, though some philosophers would contend that her evidence must not just fail to support the believed proposition but that it must actually support the *contrary* proposition being true, otherwise we may only have wishful thinking.³ It is also important that her desire causes her, in an appropriate way, to have this unwarranted belief, by for instance leading her to ‘[interpret] . . . information in a motivationally biased way’ (p.519).

Furthermore, Räikkä approvingly quotes Patrick Gardner, who says that ‘it is characteristic of the self-deceiver (. . .) that his beliefs and opinions concerning certain matters should seem to be at variance with what the available facts would lead an objective observer to conclude’ (quoted on p.520). This does not represent a further criterion, but is just a consequence of the second criterion he mentions. That is to say, given that the belief is ‘not properly supported by the evidence’, an

objective observer (who by definition is supposed to be one who would draw the warranted conclusion), acquainted with that same evidence, should not draw the same conclusion of the self-deceiver.

Räikkä thinks that there is *prima facie* plausibility to the idea that religious beliefs qualify as being self-deceptive under his understanding of self-deception. This plausibility derives, he thinks, from the fact that religious beliefs often meet the two criteria mentioned. For instance, the proposition that we will live after death is for many a most welcome prospect, and also the available evidence does not, as Räikkä suggests at numerous points, warrant unreserved confidence in this proposition. Given that many advocates of this idea may give it more credence than their evidence warrants, and given that they want it to be true, it may seem likely that they have deceived themselves on this matter.

However, Räikkä thinks that such people cannot be charged with self-deception. His reason is that a certain necessary condition for self-deception is not met. He claims that ‘[i]t is conceptually necessary for self-deception that self-deceptive beliefs be false’ (p.522). Let’s call this the False Belief Condition. Objective observers, however, who would assess the issue impartially, would not conclude that the proposition at hand is false. The reasonable position, Räikkä suggests, would be an acceptance that the available evidence doesn’t really warrant any definite conclusions either way. So ‘the belief in question cannot be said to be strictly false, whatever the epistemic weaknesses of that belief might be’ (p.520).

Initially Räikkä gives the impression that he is building up to the claim that religious beliefs cannot be classed as self-deceptive, but in the end, it seems that his final claim is much weaker. He believes that objective observers would mostly conclude that the evidence at hand conclusively warrants neither belief in the afterlife, nor a rejection of the possibility. We just can’t be sure about it. But the religious believe in the afterlife. So at least a good number of them probably believe unwarrantedly, because their desire has

biased their judgment.⁴ Yet for all anyone knows, there *might* be no afterlife. If that is the case, then the False Belief Condition will be met, and so these people will be self-deceived after all (unknown to us). They will satisfy conditions (1), (2), and the False Belief Condition. Therefore, ‘the conclusion must be that we *cannot know* whether religious beliefs are sometimes self-deceptive’ (p.522, my emphasis). And this weaker claim does indeed seem like all Rääkkä is entitled to conclude from his assumptions. So in other words, it is not that the False Belief Condition is not met in these cases, but that we don’t (and can’t) know whether it is. We can’t know if these religious believers are self-deceived.

One might think that Rääkkä’s defense of religious belief against the charge of self-deception would give cold comfort to the religious believer. After all, he seems to believe that the typical believer in the afterlife believes unwarrantedly, and that the explanation for this has something to do with the fact that she wants it to be true. Where she gets off the hook is on what Rääkkä, in note 32, refers to as a ‘technicality’. This is the idea that one is not, strictly speaking, self-deceived with respect to one’s belief that *p* unless *p* is false, and *p* is not ‘demonstrably false’ (p.522, my emphasis). I do not want to dispute Rääkkä’s general picture of self-deception here, nor his view on the warrantedness of belief in the afterlife, but it is on the issue of this ‘technicality’; the False Belief Condition, that I think Rääkkä leaves himself most open to objection.

Rääkkä thinks that for purely ‘lexical’ reasons, the self-deceptive belief must be false. In support, he quotes Alfred Mele: ‘[a] person is, by definition, *deceived in* believing that *p* only if *p* is false; the same is true of being *self-deceived in* believing that *p*’ (quoted on p.521). This point seems to be accurate, as far as it goes. It seems that one can only be deceived in believing that *p* if *p* is false. Being deceived in believing that *p* in this respect entails being *mistaken in* believing that *p*. However, Rääkkä neglects to recognize that there is another grammatical construction for attributing self-deception that does not carry this false belief entailment. This construction is of the form ‘*S* deceived herself into believing that *p*’. Being self-deceived *in* believing that *p*, and having deceived yourself *into* believing that *p*, should be distinguished⁵, and I suggest that where it is contradictory to say that *S* was self-deceived in believing something true, it is not contradictory to say that *S* deceived herself into believing something true.

Support for this idea can be found by considering cases of interpersonal deception. Say that I know that *p* is true and I know you will only believe that *p* if you hear it said from Jones, who is the only man you trust on this matter. I then lie to you by saying that I was in touch with Jones, who told me that *p*, and you then believe that *p*. In this situation, you are not *deceived in* believing that *p*, since this is a true belief. However, we can say here that you were *deceived into* believing this true proposition by me. This is like saying you were *tricked into* believing that *p*, which one can be even where *p* is true. Though *p* was true, I exercised deceptive means in order to get you to believe it.

It is sensible that our language would provide for us a way of making the charge of self-deception that gets around Rääkkä’s technicality, and this is what the ‘deceived into’ construction does. The crucial thing about self-deception, then, seems to lie in believing against good evidence because of a desire or fear, and whether the belief is true or false is not a critical consideration (though beliefs unwarrantedly by the evidence will generally be false, of course). Therefore, although Rääkkä’s believer in the afterlife may not be self-deceived in believing this (though we can’t be sure), she may still have deceived herself into believing in an afterlife, even if there is one.

Notes

1 Juha Rääkkä, ‘Self-Deception and Religious Beliefs’, *The Heythrop Journal* 48 (2007), pp. 513–526.

2 ‘Straight’ contrasts with ‘twisted’ self-deception. In twisted self-deception, someone unwarrantedly believes something she fears, as opposed to desires, to be true.

3 See George Graham, ‘Russell’s Deceptive Desires’, *The Philosophical Quarterly* 36 (1986), pp. 223–229. Alexander Bird, ‘Rationality and the Structure of Self-Deception’, in *European Review of Philosophy* 1 (Stanford: CSLI Publications, 1994) pp. 19–38, esp. pp. 37–38. Alfred R. Mele, ‘Real Self-Deception’, *Behavioral and Brain Sciences* 20 (1997), 91–102, esp. p. 100.

4 This does not entail that they believe *irrationally*. Whether it is rational to believe something may have to do with more than whether it is warranted. Considerations to do with whether the belief would ease anxiety may also be relevant, as may the considerations that inform Pascal’s Wager, for instance.

5 Perhaps the first discussion of this distinction was in B.P. McLaughlin, ‘Exploring the Possibility of Self-Deception in Belief’ in B.P. McLaughlin and A.O. Rorty (eds.), *Perspectives on Self-Deception* (Berkeley etc.: University of California Press, 1988), pp. 29–62.