SAFETY AND KNOWLEDGE OF GOD

T. J. MAWSON

St Peter's College, University of Oxford

Abstract. In recent 'secular' Epistemology, much attention has been paid to formulating an 'anti-luck' or 'safety' condition; it is now widely held that such a condition is an essential part of any satisfactory post-Gettier reflection on the nature of knowledge. In this paper, I explain the safety condition as it has emerged and then explore some implications of and for it arising from considering the God issue. It looks at the outset as if safety might be 'good news' for a view characteristic of Reformed Epistemology, *viz.* the view that if Theism is true, many philosophically unsophisticated believers probably know that it's true. A (tentatively-drawn) sub-conclusion of my paper though suggests that as safety does not by itself turn true belief into knowledge, the recent focus on it is not *quite* such good news for Reformed Epistemologists as they may have hoped: it's not that safety provides a new route by which they can reach this sort of conclusion. But safety *is* still good news for their view at least in the sense that there is no reason arising from considering it to count these philosophically unsophisticated believers as not knowing that there's a God. I conclude by reflecting that good news for Reformed Epistemology is perhaps bad news for the discipline of Philosophy of Religion more generally, as there's a possible 'reflection destroys knowledge'-implication to be drawn. Those who have been led to their religious beliefs in at least some philosophically unsophisticated ways seem to enjoy much safer religious beliefs than those who have been led to their religious beliefs by philosophical reflection, so the discipline as a whole will be adversely affected if safety is eventually accorded the role of a necessary condition for knowledge.

I. INTRODUCTION

For a given subject, *S*, to know a given proposition, *P*, it is uncontroversial that *S* must believe that *P* and that *P* must be true. That something more than this is needed is also uncontroversial. We all rebel at attributing

knowledge of the time to a man who glances at a stopped clock without realizing it's stopped, presumes that it's working, and hence comes to the belief that it's noon on the basis of what it reads. And we all do this even if we are told that the man in question just happens to be going through this process at noon and hence just happens to come via it to a true belief about the time. Cases of genuine knowledge, we all suppose, cannot be as is the truth of this person's belief – a matter of luck. Thus it is that all who are interested in Epistemology have an interest in understanding what it is for a true belief to be other than luckily true.

In recent Epistemology, a family of closely-related views has been developed in response to this interest. Duncan Pritchard, Ernest Sosa, and Timothy Williamson have all defended variants of what has come to be known as the 'safety' condition for knowledge and is widely considered to be an advance on the previously popular 'sensitivity' condition, most famously articulated and defended by Robert Nozick as in itself sufficient to generate knowledge when added to true belief.[1] Pritchard and Sosa have contributed as participants in the traditional enterprise of offering a reductive analysis of knowledge. Williamson, believing no such analysis will be successful, offers his account of safety by way of articulating an aspect of knowledge; he may remain untroubled by, indeed welcome, the claim that in determining the presence of this aspect one needs ultimately to draw on judgments of which beliefs count as knowledge. As well as the different ways in which these authors intend to use the notion of safety, the notion is itself differently expressed by each (and even by any one of these authors in different places). But many of these differences do not make a difference to the issues that I'm going to be concerning myself with in this paper, *viz.* the implications of and for the safety condition (and of and for the safety condition when it is allied with a cognitive ability/'epistemic virtue' condition – 'safety 'plus'', as I shall call it) arising from a consideration of God's existence or non-existence. That being so, I shall not spend much time going into the differences between these authors' views.[2] Rather, I shall proceed as follows.

I shall start at a general level, discussing the apparent problems for safety theorists (as I shall call those who wish to suggest that safety is

---

[1] I have previously discussed problems for Nozick's view arising from considering the case of God in my 'How a single personal revelation might not be a source of knowledge', *Religious Studies*, 39 (2003), 347-357.

[2] Nor shall I go into all the changes in the views that these authors have each undergone.

sufficient additive to make true belief into knowledge) presented by knowledge of necessary truths *per se*. I shall then move on to consider the more particular apparent problems presented by knowledge of the necessary truth that is the truth concerning God's existence or non-existence, that is to say the necessary truth that must be asserted either with the sentence 'God exists' or with the sentence 'God does not exist'. As comments made in the previous paragraph and my dropping in the word 'apparent' immediately before 'problems' just now may already have indicated, I certainly do not wish to suggest that the issues I diagnose need to be taken as fatal for those who wish to maintain that safety is *a part* of what must be added to true belief for knowledge, the part perhaps that gets past at least some Gettier cases; indeed, to reveal my own hand, I think it *is* an important part and it *does* get past at least some Gettier cases. And I do not wish to suggest that 'supplementing' (if supplementing it be) safety with a cognitive ability or epistemic virtue condition – safety 'plus' – isn't important or won't get past more. But the main sub-conclusion of my argument will be that consideration of the God issue suggests that it is at least somewhat implausible to maintain that safety – or even safety 'plus' – is the only thing that needs to be added to true belief for knowledge. That though may be considered something of a side-effect of my argument; my main point, of more relevance to the interests that are likely to be held by readers of this journal, will be that belief in God looks as if it is going to be safe (and safe 'plus') for just the sort of philosophically unsophisticated believer who Reformed Epistemologists are typically interested in claiming may know that there's a God. In short then, the recent focus on safety in 'secular' Epistemology is good news for Reformed Epistemology in the Philosophy of Religion. If, as many believe (myself included), safety (or safety 'plus') is an important part of the best post-Gettier understanding of knowledge that we have available to us, the Reformed Epistemologist has reason to be glad.

II. SAFETY AND KNOWLEDGE OF NECESSITIES

I shall assume in what follows that readers are familiar with the outlines of the Reformed Epistemology position, roughly, that if Theism is true, many philosophically unsophisticated believers probably know that it's

true.[3] But because I do not wish to assume that readers of this journal will be familiar with the literature on safety, let me start by giving an overview of that position.[4]

A good starting point for understanding the safety condition for knowledge is thinking of it as an attempt to capture our intuition that for *S* to know that *P* it has to be the case that *S* could not easily have believed that *P* and yet *P* have been false. That seems like what's gone wrong in our first example – the man who's glanced at a stopped clock without realizing it's stopped but, luckily, done so at just the hour the clock says it is. This formulation makes it appear though that there are two factors which go up to determine the safety of a particular person's particular belief, the *S* and the *P*. Is this right?

Well, of these two, the factor on which attention has most focused in the literature is the *S*. If *S* is the sort of person who, at least on the subject matter that *P* concerns, employs methods of coming to beliefs that are very reliable in yielding true beliefs and in avoiding yielding false beliefs, it's uncontroversial that that in itself helps *S* along the way to satisfying the safety condition. And that seems intuitively plausible too. To invigorate the intuition, we might remember the example of the man glancing at a stopped clock and imagine by way of contrast a person who instead only comes to judge of the time after consulting several clocks and each for long enough to establish that it is ticking over.

What is given less attention in the literature and is perhaps not intuitively plausible is the fact that it appears from the safety condition as stated that if *P* is the sort of thing which could not easily have been false, that too will help *S* along the way to satisfying the safety condition and the limiting cases of things that are not easily false are those things which are necessarily true. The only extended discussion of safety's problems with necessary truths that I know of occurs in a paper that I came upon only after having completed the main argument of this paper – Roland

---

[3] Or perhaps better, and more minimally, that there is no *de jure* objection against such a view that is not more ultimately a *de facto* objection against the truth of Theism. See Peter Forrest's 'The Epistemology of Religion', in the *Stanford Encyclopedia of Philosophy* for an introduction to the topic and its literature. Available at: <http://plato.stanford.edu/entries/religion-epistemology/#Ref> [accessed 30/04/2014].

[4] See also D. Rabinowitz, 'The Safety Condition for Knowledge', in *The Internet Encyclopedia of Philosophy*. Available at: <http://www.iep.utm.edu/safety-c/> [accessed 30/04/2014].

and Cogburn's 'Anti-luck Epistemology and Necessary Truths'.[5] The arguments of that paper are different from, but largely compatible with, those of this paper; they focus on establishing that safety doesn't do any better than sensitivity in dealing with necessary truths *per se*. I report though that, personally, I don't think that a crucial example of theirs – one that involves a demon who is determined to feed only true results to a broken calculator – can do the work that they require of it. The examples I am about to give bypass the issues they need this problematic example for – amongst others, meeting the challenge of those who would carve up methods of belief formation such that their example of someone coming to a belief as a result of a determinedly-truthful-demon/calculator counts as using a different method from someone who comes to a belief as a result of a whimsical-when-it-comes-to-truth-demon/calculator combination.

So, is satisfying the safety condition for knowledge of necessary truths always going to be, *ceteris paribus*, easier than for knowledge of contingent truths? I said a moment ago that it would 'perhaps' not be intuitively plausible to think so, but there is a fine tradition – of which Descartes is perhaps the most obvious exponent – of holding that at least some necessary truths are in fact easier to know than any contingent ones and that they are easier to know in part in virtue of their necessity. In any case – straight out of the gate, as it were – safety theorists appear to have a problem dealing with cases of knowledge of necessary truths, for it is true of *everyone* who comes to believe *any* necessary truth that it is not the case that they could very easily have believed that particular necessary truth and yet that particular necessary truth have been false. And however taken we are with the Cartesian thought that knowledge of at least some necessary truths might be easier for a given S than knowledge of any contingent truths, we surely won't wish to say that everyone knows every necessary truth they believe. I don't suppose that intuition needs much invigorating, but consider the following case:

Unbeknownst to me, I have been hypnotized so that I will believe whatever it is that I next read. I have then been placed at a desk with two pieces of paper on it, one to my left; the other, to my right. Each piece of

paper is blank on the uppermost side, but the one to my left has written on its other side '2 + 2 = 4', whilst the one to my right has written on its other side '2 + 2 = 5'. I decide to pick up one of these bits of paper and read anything that's on the other side of it, and the decision of which one of the bits of paper I'll pick up depends on the results of the toss of a fair coin; heads, it'll be the piece to my left; tails, it'll be the piece to my right. I toss the coin; it comes up heads; I thus end up believing that 2 + 2 = 4. Surely, our intuitions suggest, I don't know this.

I have constructed this thought experiment whilst having in view a passage from Pritchard which I am about to quote, but my having this passage in view has meant that I have used his examples of beliefs and they are ones that are perhaps not in themselves ideal for my purposes as it can reasonably be supposed that I had the belief that 2 + 2 = 4 prior to getting myself into the situation described in the thought experiment and that reasonable supposition threatens the stability of the intuition that I don't know that 2 + 2 = 4 in the situation imagined. But I hope it can be seen that an adaptation of the thought experiment could easily sidestep this problem, if problem it be. Instead of the simple and probably-already-believed truth that 2 + 2 = 4 and the simple and obvious falsehood that 2 + 2 = 5, imagine that on the piece of paper to my left is the true statement of a sum that I hadn't previously got any beliefs about and on the piece of paper to my right a false statement of that sum. We wouldn't be at all inclined to think that I'd know the true sum then.

According to Pritchard, 'it is pretty easy to see how one might go about extending the account of safety to these propositions, even if the details might be tricky. After all, all we need to do is to talk of the doxastic result of the target belief-forming process, whatever that might be, and not focus solely on belief in the target proposition. For example, if one forms one's belief that 2 + 2 = 4 by tossing a coin, then while there are no near-by possible worlds where *that* belief is false, there is a wide class of near-by possible worlds where that belief-forming process brings about a doxastic result which is false (e.g., a possible world in which one in this way forms the belief that 2 + 2 = 5). The focus on fully contingent propositions is thus simply a way of simplifying the account; it does not represent an admission that the account only applies to a restricted class of propositions.'[6]

---

[6] D. Pritchard, 'Safety-Based Epistemology: Whither now?', *Journal of Philosophical Research*, 34 (2009), 33-45 (p. 34). Williamson in effect proposes the same move by

So it is that we are now in a position to see that, according to the safety theorist properly understood, when thinking about *S*'s knowing that *P*, we should *not* in fact think of the safety of *S*'s belief that *P* as a function of two factors, the *S* and *P*; the 'rigidity', as we may put it, of *P*'s truth over nearby or relevant possible worlds is entirely moot in determining the safety of *S*'s belief that *P*. Rather, the onus for generating safety falls entirely on *S* and in particular the reliability of his or her 'target belief-forming process', as Pritchard puts it – that is to say the process by which *S* has in fact come to the belief that *P*. If that particular belief-forming process is one that could easily have led to a different and false belief, *Q*, then that fact undermines *S*'s safety in believing that *P* even if (as will of course be the case when *P* is a necessary truth) there is no possible world in which that process leads to the belief that *P* and yet *that* belief is false. If the particular belief-forming process does not have this feature, then, presumably, *S*'s safety in believing that *P* is not undermined. The necessity or otherwise of *P* can thus be said to have nothing to do with the safety of the truth of the belief that *P*. So, once we do that which Pritchard assures us is 'pretty easy', *viz.* extend the safety account from contingent propositions to necessary ones, for – after all – the 'focus on ... contingent propositions ... is ... a way of simplifying the account; it does not represent an admission that the account only applies to a restricted class of propositions', we can see that he is indeed right to say this; the 'rigidity', as I have put it, of the truth of *P* over possible worlds and thus the modal status of *P* drop out of the picture entirely and safety-based epistemology may be seen for what it is, a form of 'belief-forming process' reliabilism.

Well, I say that safety-based epistemology may be seen as a form of reliabilism, but it is worth noting that Pritchard himself doesn't see it this way. Indeed in one work he says, 'There is no inherent reason ... why a safety-based account of knowledge should mention the reliability of processes at all.'[7] But his views have changed since he wrote that and in

---

the manner in which he justifies counting a similar coin-tossing case as not yielding knowledge, in his *Knowledge and Its Limits* (Oxford: OUP 2000), p. 182 - the method, he points out, could easily have yielded a false belief. And Pritchard returns to a similar example, consulting a broken calculator, in his most recent piece, 'Anti-Luck Virtue Epistemology', *The Journal of Philosophy* (forthcoming) and proposes the same move once more in response to it.

    [7] D. Pritchard, 'Virtue Epistemology and Epistemic Luck', *Metaphilosophy*, 34 (2003), 106-130 (p. 119).

later work it seems to me that he (almost) sees it this way, in virtue of his claim that in order to know, *S* must satisfy 'an ability condition of some sort – i.e., a condition to the effect that the true belief was gained via the employment of the agent's reliable cognitive abilities'.[8] I say 'almost', as in contexts where he speaks of this, he states that a belief's being safe and its satisfying an ability condition are its doing *two different* things,[9] whereas it seems to me – given the considerations already presented – that if the account of knowledge is genuinely to apply to necessary truths as well as contingent ones, then the safety of a belief must be understood just as the belief being produced by a reliable process or cognitive ability.[10] (There is still room perhaps for a contrast with a Goldman-style reliabilism; Pritchard tells me in email correspondence that he thinks of his view as offering 'a modal rendering of reliability ... [If so], the contrast with standard reliability views, which are probabilistic, would still hold'.) Nevertheless, wishing to bypass the issue of whether safety in a belief just reduces to the reliability of method/ability employed in arriving at it, I shall spend some time when articulating the relevant examples in making it obvious that they concern people whose belief is safe 'and' satisfies any plausible 'ability condition'. I intend it to be plausible that the 'belief-forming process', as Pritchard puts it in some places, or the 'cognitive ability', as he puts it in other places, which has been utilized by a certain character that I'm about to introduce in a thought experiment – *viz.* whichever of Theo Theophilus and Atheo Theophobus has the true belief about God – is unerringly reliable; in every possible world in which this character exists, he or she uses this process/ability and it gets to the truth. I also intend it to be plausible that either in the thought experiments as they stand or in more fine-tuned 'iterations' of them, the ability is as well-integrated with (by being fundamental to) their cognitive abilities and what have you, and its exercise is as virtuous as any epistemic virtue ever gets. (Or at least, we have no reason to think it's not virtuous until we have the intuition that the subject exercising it doesn't thereby know.)

[8] D. Pritchard, 'Safety-Based Epistemology: Whither now?', *Journal of Philosophical Research*, 34 (2009), 33-45 (p. 41).

[9] Indeed he has a whole paper devoted to arguing that they're different things, his 'Anti-Luck Virtue Epistemology', *The Journal of Philosophy* (forthcoming).

[10] Another point of disagreement, for the record: it also seems to me that the character of 'Temp', from one of Pritchard's thought experiments in his 'Anti-Luck Virtue Epistemology', *does* – *pace* Pritchard - know the temperature of the room.

At a general level, the safety theorist (or 'belief-forming process reliabilist', if that's indeed what he or she is) seems vulnerable on the necessary truth front if we can construct a case along the following lines. For a given necessary truth, *P*, we keep *S*'s belief in *P* fixed across close/relevant possible worlds and, for reasons just sketched, keep fixed the belief-forming process/cognitive ability the exercise of which has led to it (rather than allow of *S* that he or she could easily have ended up believing in some false *Q* by that same process/ability or allow of *S* that he or she could easily have deviated from that particular belief-forming process/ability and deployed another). Ideally then, we'd have a case where there's *no* possible world in which *S* exists yet fails to employ that same process/ability; there's no possible world in which *S* employs that same process/ability and yet comes to another belief, a *Q*, as a result of it; and there's not even a possible world where *S* employs the same process/ability and yet merely fails to come to believe the *P* that he or she actually comes to believe. And (to take in safety 'plus' views) we'd have the thought-experimental situations we thereby constructed capable of iterations such that the relevant ability could be integrated or what have you (without limit) with the subject's belief-forming mechanisms and in general made to count as a virtue however the notion of epistemic virtue is to be unpacked, or at least count as such until we have the intuition that the subject using that ability/method doesn't know thereby. If we can construct such an example and it is one where we don't intuitively wish to say of *S* that he or she knows *P*, then the safety- (or even safety 'plus'-) theorist will be resultantly troubled; remember, we are using the term to refer to those who wish to maintain that safety (or safety 'plus') is *all* that needs to be added to true belief in order for there to be knowledge. It seems to me that we can construct an inconclusive example of this sort by considering the case of God. It is inconclusive as it is not clear in which direction we should 'run' the example. Tentatively, I suggest that we should run it as indeed showing that safety (or even safety 'plus') is insufficient additive to make knowledge out of true belief; the safety theorist is wrong. But an alternative, which I do not wish to rule out unequivocally, would take it as showing that knowledge of God's existence, if He exists (or of His non-existence, if He doesn't) is easier than might have been supposed. In any case, as it seems to me that safety (or safety 'plus') is – despite my tentative willingness to run the examples to show its insufficiency to meet all Gettier problems – an important feature of knowledge (it is not the philosopher's stone as

it were – *the* additive, but rather *an* additive, that needs to be present for true belief to become knowledge), this is reassuring news to Reformed Epistemologists: the focus on safety in 'secular' Epistemology certainly does nothing to undermine the knowledge-status of the sorts of religious beliefs had by many philosophically unsophisticated believers.

III. SAFETY AND KNOWLEDGE OF GOD

On then to the examples. The first example I'll consider won't do anything much to advance the case, but I want to put it on the table, as the parallel for the first case we considered in engaging with Pritchard, prior to moving on.

1. *The Coin-tossers*

Theo and Atheo Coin-tosser are brother and sister. They find themselves sequentially visiting a hypnotist who can (for the purposes of the thought experiment, unfailingly) make them believe with absolute conviction either that God does exist or that He does not. Which way this hypnotist hypnotizes them will be entirely dependent on what they ask him to do. Both Theo and Atheo Coin-tosser choose to let which of two beliefs – 'There's a God' or 'There's not a God' – they ask the hypnotist to induce in them depend on the toss of a fair coin. 'If it's heads, I'll ask for (and thus get) belief in God; if it's tails, I'll ask for (and thus get) belief that there's no God', they both say. Naturally, given some necessities about the nature of belief to do with the fact that one cannot regard a mental occurrence as a belief if one believes of it that one has no truth-tracking reason for that mental occurrence rather than its negation, they both also determine to get the hypnotist to make them instantly forget how it is their belief that there's a God or their belief that there's not a God originated, but that will prove no difficulty either; the hypnotist provides this service as standard. Theo Coin-tosser goes first; he tosses his coin and it lands heads-side up; he thus ends up believing that there is a God. Atheo goes next; she tosses her coin and it lands tails; she thus ends up believing that there is no God.

The method that the Coin-tosser siblings employ can be seen to be unreliable; it's in essence the same as the method we talked about when first engaging with Pritchard. And it's uncontroversially unreliable. Both

theist and atheist can see that in the actual world it's led to a false belief in one of the Coin-tosser siblings and that in nearby worlds it leads to false beliefs quite a lot of the time for the 'alethically' lucky one in the actual world, 50% of the time indeed (by stipulation – otherwise it wouldn't be a fair coin). Of course, which of the Coin-tosser siblings is the one who has been alethically lucky in the actual world will be an issue on which theist and atheist will have diametrically opposed views, but that one of the siblings is such is something on which they will agree.

2. The Enthusiasts

Theo Theophilus and Atheo Theophobus are not brother and sister, but rather two people of entirely different lineages, at least as far back as can make no difference. We'll go back as far as their parents, Mr and Mrs Theophilus and Mr and Mrs Theophobus.

Mr and Mrs Theophilus met at church when they were very young (as had their parents and their parents before them); their mutual religion drew them together like nothing else could have done; in particular, they loved quoting to one another passages from the Bible which were premised on Theism's being true. Little did they know, because little did they reflect, that Theism is something which, if true, is a necessary truth. Being unreflective as to its modal status, they were nevertheless in agreement on the fact that Theism (even if they didn't call it that) was one of the most attractive features of their religion; had the only religions they been exposed to been atheistic versions of Buddhism, for example, they'd never have met. As it was, Mr and Mrs Theophilus married and turned their minds to having a child. They agreed that they'd rather have no child at all than have one who didn't accept without question that God exists. Perhaps this was because they believed, as so many who believe in God believe, that God is likely to send those who don't believe in Him to Hell for all eternity. In any case, they set up a superfluity of safeguards – by way of pre-commitments, trust funds, executors, guardians, and so forth – to ensure that – regardless of what happened to them after their child's birth – any child of theirs would follow a rigorously-stipulated programme of education which would ensure that he or she would believe this truth. Again, they were in absolute and unwavering agreement that, had it proven impossible to set all this up, they wouldn't have risked having a child at all. That all being in place however, they

conceived a child, a boy who they named Theo. Theo is now an adult who has believed from his youth that God exists. [11]

    Mr and Mrs Theophobus met at a local rally of secular humanists and free-thinkers when they were very young (as had their parents before them, *et cetera*); their mutual antipathy towards theistic religion drew them together like nothing else could have done. In particular, they loved quoting to one another passages that directly implied Atheism taken from the writings of those they regarded as great thinkers. Little did they know, because little did they reflect, that Atheism is something which, if true, is a necessary truth. Being unreflective as to its modal status, they were nevertheless in agreement on the fact that Atheism was one of the most attractive features of their shared worldview; had either been tempted to think that belief in God might, after all, have something going for it, they'd never have met. As it was, Mr and Mrs Theophobus married and turned their minds to having a child. They agreed that they'd rather have no child at all than have one who didn't accept without question that God does not exist. Perhaps this was because they believed, as so many who believe that there is no God believe, that belief that there is a God is likely to lead to the person who holds that belief significantly wasting the one and only life that he or she has. In any case, they set up a superfluity of safeguards – by way of pre-commitments, trust funds, executors, guardians, and so forth – to ensure that – regardless of what happened to them after their child's birth – any child of theirs would follow a rigorously-stipulated programme of education which would ensure that he or she would believe this truth. Again, they were in absolute and unwavering agreement that, had it proven impossible to set all this up, they wouldn't have risked having a child at all. That all being in place, they conceived a child, a girl who they named Atheo. Atheo is

---

[11] If the global reliability of the methods of belief acquisition/cognitive abilities to which this educative process gives rise is called into question, we can augment the details of the thought experiment to establish it definitively by adding other truths into the pot which, for ease of exposition, just has 'God exists' in it at the moment. And we can, if needs be, tweak the thought experiment so as to integrate the relevant cognitive ability in a way which makes it virtuous; we can, if needs be, do it in a 'cheating' way - by adding to the pot the truth that the beliefs in the pot are arrived at in a virtuous way. Inserting the belief that the beliefs in the set are epistemically virtuous is not at all *ad hoc* in the case of belief in God; the idea that one believes as a result of faith and that faith is a virtue, one of the cardinal virtues indeed, is mainstream. I shall return to clarify and expand on these points in the main text later.

now an adult who has believed from her youth that God does not exist. [12]

If we reflect on the situation of whichever of the Enthusiasts we suppose is right, we will think that we have in this example an example of a given truth, a *P* – *viz.* whichever of Theism or Atheism is right – that is necessary and thus true in all possible worlds; it's true then in all close/relevant worlds however we end up determining closeness and relevance. We have a given subject, a *S* – *viz.* whichever of the Enthusiasts is right – who believes this truth across all close/relevant possible worlds. Again this result is secured against the vagaries of how we judge of worlds that they're close/relevant by my having constructed a story whereby this particular *S*'s coming into existence depends on the shared enthusiasm of his or her parents for this particular truth. Of course, I need to suppose not just what I've explicitly stated, that these two parents would never have conceived a child had they not shared their enthusiasm, but also something about the necessity of origins and that in itself is a questionable metaphysical supposition. Certainly if granted, the facts are then in place to prevent this particular *S* existing in any world in which he/she comes to a belief in some false *Q*, rather than this particular *P* by the method that leads to *P* in the actual world. This particular *S* then could not easily have ended up believing in some false *Q* by the same process that has actually led him/her to believe this true *P*; nor could this particular *S* have easily deviated from that particular belief-forming process and deployed another.

If God exists, then Theo Theophilus's belief that He exists is safe. If God does not exist, then Atheo Theophobus's belief that He does not exist is safe. If safety was the missing piece for knowledge, we could say then that one of Theo or Atheo knows the truth about whether or not there is a God. And, if we didn't class ourselves as agnostics, we'd be able to pick out which one it was. And if safety plus global reliability of method/cognitive ability were the 'composite' missing piece, we'd be able to do the same for whichever of the Enthusiasts is right; his or her method/ability is globally reliable. And, if safety 'plus' virtue were the missing piece, then, at least for an iteration of the thought experiment that has beliefs about the nature of virtue being had in this way, we'd be

---

[12] Again, if needs be, one can supplement the number of truths to which they took this attitude, so as to sidestep the generality problem in establishing global reliability of method. One can repeat the 'cheat' too, if one wants. Again, I'll return to clarify and expand on these points later.

able to do the same. Now, I myself am somewhat tentative in drawing this conclusion from these thought experiments, but it seems to me intuitively implausible to say that either knows. And it is not as if this is a hypothetical scenario entirely divorced from the realities of everyday religious or irreligious belief. Quite a few actual couples are (or certainly, historically, *have been*) as committed to their Theism as Mr and Mrs Theophilus. In the present day, one might consider the Amish. Whilst smaller in number (both in absolute and relative terms) and a more recent development than Theophilus-like fans of Theism, there are also Theophobus-like fans of Atheism. In the recent past (and even, to a lesser extent admittedly, the present day), we might consider ideologically-pure devotees of Communism. Just by way of anecdotal evidence, I report that I know several Christians who are relevantly exactly like Mr and Mrs Theophilus and one Communist who is relevantly exactly like Mr and Mrs Theophobus. I know of scores of theists who are relevantly very like Mr and Mrs Theophilus and at least half a dozen atheists who are relevantly very like Mr and Mrs Theopobus. So, the situations of Theo Theophilus and Atheo Theophobus, whilst probably more exteme than any you or I have come across, are not so bizarre that we should not trust our resultant intuitions at all. The Theism and Atheism of quite a few theists and atheists in the actual world are (and certainly have been) in relevant respects as are the Theism and Atheism of Theo Theophilus and Atheo Theophobus. The Enthusiasts thought experiment provides then, I suggest, at least some reason at least for thinking that knowledge is not simply safe true belief or even safe-'plus' true belief. Safety may be, indeed I think it is, an important piece of the post-Gettier jigsaw, but it is not *the* missing piece.

In short, because Theo Coin-tosser could easily have been (in respect of his belief *vis-à-vis* the God issue) qualitatively the same as Atheo Coin-tosser, the belief of Theo Coin-tosser that there is a God and the belief of Atheo Coin-tosser that there's not a God don't satisfy the safety condition and hence, if safety is at least a necessary component of knowledge, neither can be knowledge. And that seems right. But because Theo Theophilus could not easily have been (in respect of his belief *vis-à-vis* the God issue) qualitatively the same as Atheo Theophobus, whichever of them believes the truth about the God issue will count as satisfying the safety condition for knowledge and any plausible unpacking of a reliability of method/ability condition; and thus, proportionate to our reflective reluctance to count whichever of them is right as knowing, we have reason to think that

safety is not sufficient additive to make knowledge out of true belief and nor is safety 'plus'. Inserting a method/ability/virtue condition does not seem capable of changing things, for (a) there seems no non-question-begging way to justify counting these characters as non-virtuous (to me, even as they stand, without any iterations) and (b) one can in any case iterate the thought experiment, putting the belief that these beliefs are virtuous into the relevant set of truths.[13] The bottom line: along with truth and belief, safety or ('and', if you will) reliability of method/ability may be necessary for knowledge, but – it seems – it is not (they are not, if you will) sufficient. Nor are safety 'plus' virtue.

Now, in unpacking the Enthusiasts thought experiment so that it can do the work we require of it, there are certainly some issues we need to address. We need to be careful about how we specify the belief-forming process, method, or cognitive ability that Theo/Atheo has utilized to get to the relevant belief. As perhaps already indicated, the intention is to portray a *S* who believes a particular *P* in an entirely 'basic' way. The believer has just, as far back as he or she can remember, believed it. It's important to keep to this line as, for the example to do its work (that is generate an interesting result), we need to keep the reliability of the method/ability that the believer employs high and yet its intuitive plausibility as knowledge-generating low and, to that end, we don't want to end up inadvertently adding justification into the mix. And I concede at the outset that holding to this line will make it harder (though not, I think, quite impossible) to hold to the line that I'm also taking, that an appeal to epistemic virtue (in the manner of the most recent Pritchard) doesn't fundamentally affect things, that safety 'plus' doesn't do any better than safety *simpliciter*.

As foreshadowed in some footnotes, I can imagine it not being immediately obvious to some from the case as hitherto presented that the process/ability Theo/Atheo employs in coming to his/her belief is, as it is sometimes put, 'globally reliable'. It would be tempting to reply to a challenge along these lines by expanding on the details of the example, giving details of the education that their parents were so committed to providing such that it would be obvious that coming to believe a thing

---

[13] This then will be to insert a contingency into the set of truths, so they won't now all be necessary, but the rigidity of the truth over nearby worlds seems pretty robust: if I truly believe that I'm virtuous in this regard, then the exercise of this cognitive ability will be a virtue in all nearby worlds and indeed quite a few far-out ones.

as a result of that education is employing a globally reliable belief-forming process or cognitive ability/set of cognitive abilities; in essence, one could easily stipulate that they secure for their child a first-rate education. However, in meeting the challenge in this way, one would start to run the risk of undermining the intuition that this S doesn't know this P as a result; certainly, the education that the parents would, by such expansion, then be being depicted as providing for this S would give this S knowledge of a whole host of other issues; so it'd be more tempting to think that it might be giving knowledge of this one too. If, in order to meet this problem, one improved the general level of education, but 'compartmentalized off' the processes/abilities that led to the belief Theism/Atheism, the objector would have more cause to raise the 'generality problem' and to characterize his or her education and the methods/abilities to which it gives rise as a curate's egg – good in parts. It is, they might then maintain, not generally good in those parts that have to do with the methods/abilities of discerning metaphysical necessities. Or perhaps such a compartmentalized mind fails plausible standards for virtue.

In order to establish the reliability of the method/ability whilst keeping stable the intuition that whichever of Theo or Atheo is right doesn't yet have knowledge and in order to do all this whilst sidestepping the generality problem, it seems to me that the best way to fine-tune the thought experiment would be by increasing the number of metaphysically necessary truths towards which, in the un-iterated thought experiment, the parents take the attitude that they take towards the necessary truth of Theism/Atheism. And it seems to me that by adding details of this sort, one can render plausible the suggestion that Theo/Atheo's parents are instilling in their child an epistemic virtue too, however that is ultimately to be unpacked. Or rather, that one can render it plausible that this is what they are doing until one weighs in the balance the intuition that Theo/Atheo, whichever is right, doesn't by utilizing it thereby know. For example, one might specify that the parents took a similarly strong liking to 'Nothing comes from nothing, nothing ever did' (having heard a song with that lyric sung by the lovely Julie Andrews in the film *The Sound of Music*); for 'Que sera, sera. Whatever will be, will be' (having heard that from the lovely Doris Day in *The Man Who Knew Too Much*); and so on. They'd each only have a child with someone who believed each of these things in the same manner as in the thought experiment as it stands they'd only have a child with someone who believed that there's a God/

no God. Throw enough of these into the pot and eventually it will have to be conceded that the parents' upbringing provides their child with a method/cognitive ability for arriving at a whole series of metaphysical truths (and no falsehoods) and thus is globally reliable however fine-grained one gets in method individuation. The 'belief-acquisition process' or 'cognitive ability', if one wishes to call it that, that whichever of Theo/Atheo is right on the God issue exercises whenever he/she trusts his/her brute metaphysical instincts unfailingly yields a wide variety of truths and never any falsehoods. Of course one might say that it doesn't deserve to be counted as an epistemic virtue nonetheless, but it is not obvious on what basis such a charge could be made without begging the question, i.e. noticing that it's not knowledge-providing. As mentioned in a previous note, one could even throw into the pot – though this might be thought 'cheating' – the belief that the beliefs in the pot are arrived at virtuously; as long as virtues are rigid (as I put it) over nearby worlds, that will be a belief which – even if not a necessary truth – will be very safe if true.[14] With all these maneuvers, one might hope to make the case against the safety theorist more definitive, as one could show that despite the then-proven superb reliability of the method/ability the relevant Enthusiast employs, our intuition remains that whichever of Theo/Atheo is right, doesn't yet know.[15] But, the longer the story, the further it takes us away from the religious beliefs of actual people and the less secure any intuitive response we have to it becomes. I confess to losing confidence in my intuitions as we iterate in these ways. And

---

[14] Of course, this is to assume something about the contingency with which (at least some) abilities get to be epistemic virtues as well as the slow 'rate of change' in what counts as virtues as one moves through logical space.

[15] Of course this could all be happily accepted by Williamson, who is not in the business of presenting safety and so on as parts of a non-circular reductive analysis of knowledge. In this context, it's worth considering Pritchard's use of the character Alvin, someone who is characterised as having a brain lesion that reliably causes him to believe he has a brain lesion ('Anti-Luck Virtue Epistemology', *The Journal of Philosophy* (forthcoming), p. 13, though the example goes back to Plantinga). Pritchard hopes that it will be obvious that whilst Alvin's resultant belief is safe, Alvin fails an ability condition, and (thereby) fails to know. But a lot of work seems to me to be being done in making this case by the word 'lesion' and by the fact that the lesion causes just the one belief; it's not globally reliable. Were we to replace the word 'lesion' with the more neutral 'condition' and specify that the brain condition gives Alvin a wide range of safe beliefs about a certain area of the world, that would – it seems to me – be for it to be very plausibly the neurological base of a cognitive ability, one the exercise of which would be epistemically virtuous.

thus I only tentatively conclude that safety (or safety 'plus') is not *the* missing piece (that when added to true belief makes knowledge). But even if not the missing piece, it seems to be a piece – consider again the thought experiments concerning stopped clocks with which we started – and thus it will be comforting for those inclined to Reformed Epistemology that if Theism is true, the safety condition seems likely to be met by the sorts of believers they typically wish to portray as knowing that there's a God, those who believe in God not for any philosophically sophisticated reason, but rather just because they've been brought up that way.

IV. CONCLUSION

We may conclude a number of things. Firstly, presenting the safety condition as sufficient, when added to true belief, for knowledge of the truth-value of 'God exists' is intuitively implausible. From this we may conclude that whilst safety may be a part of the jigsaw that is knowledge, we have reason proportionate to our confidence in that intuition to think that it is not *the* missing piece. In addition, I hope that my thought experiments have rendered plausible the suggestion that 'adding' (if adding it be – and a focus on the case of necessary truths seems to me to indicate that it is not) a 'reliability of method' or 'cognitive ability' condition will not substantively affect the issue; safety 'plus' reliability of method/ability isn't the missing piece either. And I hope to have rendered plausible the suggestion that adding a virtue condition won't fundamentally affect things; there's no non-question-begging way to count as non-virtuous the exercise of the globally-reliable (if one iterates the thought-experiments properly) methods employed by the protagonists in the relevant thought experiments. New iterations can be constructed, if needs be bringing the belief that these beliefs are arrived at virtuously into the set of beliefs in question, into 'the pot', as I call it. So, is all lost for safety (and safety 'plus')? I think not. The first reason is the most obvious: these thought experiments are hardly conclusive and the iterations that I have projected make them even less so. The second is that it seems to me that safety (or safety 'plus') does well in 'de-Gettierizing' some areas of knowledge. As I think through my intuitions about what I would say were one to tweak the examples I have given by adding justification to the cases of whichever of Theo Theophilus and

Atheo Theophobus is right on the God issue, I find that I'd say that they *did* know the relevant *P*. This is so even if I think of justification in a very weak sense, merely as the ability to provide defeaters to defeaters to their belief, e.g. perhaps a schematic solution to the Problem of Evil for Theo. That in itself might tempt me to think that knowledge is safe (or safe 'plus') justified true belief, that safety (or safety 'plus') de-Gettierizes everything. However, this temptation should be resisted. The arguments of this paper show that safety (or even safety 'plus') won't deal with the following sort of case. I shall give the case in two forms, so as to cater for both theist and atheist readers. First, for the theists: Theo Theophilus glances at the stopped clock and comes to the unsafe but justified and true belief that it's noon. He then comes to the safe (and safe 'plus') and justified true belief that either God exists or it's noon. We don't want to count Theo's safe (and safe 'plus'), justified, and true belief in the disjunction as an item of knowledge. Second, for the atheists: Atheo Theophobus glances at the stopped clock and comes to the unsafe but justified and true belief that it's noon. She then comes to the safe (and safe 'plus') and justified true belief that either God does not exist or it's noon. We don't want to count Atheo's safe (and safe 'plus'), justified and true belief in the disjunction as an item of knowledge. But, that having been said, safety (or safety 'plus') just does seem to me to be the right way to get past some Gettier cases – for example, simple beliefs about the time formed from looking at stopped clocks and simple beliefs about there being a barn in a field formed whilst unwittingly travelling through 'fake barn country' – and, if it is, then that is not something that we should discard. If that's right, then safety (safety 'plus') is an essential part of any post-Gettier reflection on the nature of knowledge; it's here to stay. And thus philosophical reflection on whether or not we can know that God exists has to take it into account. If the (tentatively-drawn) sub-conclusion that is the side-effect of my paper is right (safety/safety 'plus' is not 'the philosopher's stone' that by itself turns true belief into knowledge), then this is not quite such good news for Reformed Epistemologists as they may have hoped for. It's not that safety (safety 'plus') provides a new route by which they can reach their preferred conclusion, *viz.* (roughly)[16] that, if Theism is true, many philosophically unsophisticated believers probably know it's true. But it *is* still good news in at least the sense that there'll be no reason arising from considering safety (or safety 'plus') to

---

[16] See earlier note.

count these philosophically unsophisticated believers as not knowing that there's a God. Good news for Reformed Epistemology is perhaps bad news for the discipline of Philosophy of Religion more generally, as there's a possible 'reflection destroys knowledge'-implication to be drawn too. Those who have been led to the religious beliefs that they have in manners significantly like Theo Theophilus seem to enjoy much safer religious beliefs than those who have been led to their religious beliefs by philosophical reflection, so the discipline as a whole will be adversely affected if safety (or safety 'plus') is eventually accorded the role of a necessary condition for knowledge. That of course is only on the assumption that Sophos requires knowledge.