EUROPEAN JOURNAL
FOR PHILOSOPHY OF RELIGION

VOLUME 6 NUMBER 2 SUMMER 2014

BOOK SYMPOSIUM

Richard Swinburne's *Mind, Brain, and Free Will*

Richard SWINBURNE
Précis of Mind, Brain, and Free Will 1

Lynne RUDDER BAKER
Swinburne on Substance Dualism 5

William JAWORSKI
Swinburne on Substances, Properties, and Structures 17

James K. DEW, Jr.
Swinburne's New Soul: A Response to Mind, Brain, and Free Will 29

David P. HUNT
Swinburne on the Conditions for Free Will and Moral Responsibility 39

Richard SWINBURNE
Response to Reviewers 51

ARTICLES

Alexander R. PRUSS
Nested Modes, 'Qua' and the Incarnation 65

T. J. MAWSON
Safety and Knowledge of God 81

Ebrahim AZADEGAN
Divine Love and the Argument from Divine Hiddenness 101

Ross PARKER
A Critical Evaluation of Rea's Response to the Problem of Divine Hiddenness 117

Paul THOM
Inherence and Denomination in the Trinity 139

Jane DURAN
Hildegard of Bingen: A Feminist Ontology 155

Jason MEGILL
Hume, Causation and Two Arguments Concerning God 169

Michael R. SLATER
Kant's Neglected Objection to the Ontological Argument 179

Imran AIJAZ
Traditional Islamic Exclusivism – A Critique 185

BOOK SYMPOSIUM

PRÉCIS OF

MIND, BRAIN, AND FREE WILL

by

RICHARD SWINBURNE

University of Oxford

This book is a defence of substance dualism (that we human beings consist of two parts, body – a contingent part, and soul -the one essential part), and of libertarian free will (that we are not always fully caused to form our intentions), from which it follows that we are often morally responsible for our actions. But in order to establish these points I need to introduce some crucial terminology, and to defend some general theories of metaphysics and epistemology.

So in chapter 1 I distinguish three kinds of thing: substances (the constituents of the world such as electrons, planets, and houses), their properties (such as weighing 1000kg, or being spherical), and events (occurrences at particular times, which consist in substances having or changing their properties); and I argue that the history of the world (in an objective sense) just is all the events that happen. In order to tell that history we must pick out substances, etc., by what I call their 'informative designators', which (roughly) are rigid designators such that if someone knows what the word means, they will know the necessary and sufficient conditions for its application. Thus 'red' is an informative (rigid) designator, but 'water' as used in the eighteenth century is an uninformative designator, because speakers then did not know the essence of water. We should count two substances (or whatever) the same substance (or whatever) iff their informative designators are logically equivalent. Then a proposition is metaphysically necessary/

possible/impossible iff it is logically necessary/possible/impossible when we substitute coreferring informative designators for any uninformative designators. In chapter 2 I consider what makes a belief that a certain event occurred or that a certain scientific theory is true, or that some assertion is possibly true, a justified belief in the sense of a belief which is probably true on the evidence available to the believer. I defend the principle of credulity, that every basic belief is probably true in the absence of contrary evidence (that is, a defeater); and I apply it to the justification of our beliefs about logical modalities, our beliefs resulting from experience, memory, and testimony, and our beliefs about the probable truth of scientific theories. I claim that it is a defeater to any belief resulting from experience, memory, or testimony that the event purportedly experienced, remembered or testified to did not cause the resulting belief.

I then apply these results in chapters 3 to 7, to examine the relation of our life of thought and feeling to what happens in our brains and so in our bodies. I argue in chapter 3 that there are two kinds of event in the world: physical events (including brain events) and mental events. Mental events are events to which the subject (the person whose events they are) has privileged access, that is a way of knowing about them not available to others. Among mental events are pure mental events, ones which do not include any physical event. Among these are beliefs, thoughts, intentions, desires and sensations, events of which the subject is often conscious and which are then conscious events. Given that events are individuated by informative designators, it then follows that pure mental events are not the same as physical events and do not (metaphysically) supervene on them. I go on in Chapter 4 to argue that not merely do brain events often cause mental events, but mental events (and in particular intentions) often cause brain events, and thereby bodily movements. Many neuroscientists have interpreted the results of recent neuroscientific experiments as showing that our pure mental events (and in particular our intentions) never cause brain events. I argue that these results do not show that, and that no experimental evidence of any kind could possibly show that, because in order to show that we would need evidence about our pure mental events which – given the principles about memory and testimony cited above – could only be obtained on the assumption that those events do cause brain events.

In chapter 5 I argue that this result that our intentions often cause our brain events needs to be expressed more carefully as the result that

persons often cause brain events when they intentionally cause bodily movements. The view deriving from Hume holds that the causes of events are other events, logically distinct from them; for example that when the ignition of dynamite causes an explosion, the ignition is a separate event from the explosion, and the first event causing the second one consists in there being a law of nature (a consequence of fundamental laws) which determines that an event of the first kind is followed by an event of the second kind. I now argue that whatever might be the case with non-intentional causation (e.g. the ignition of gunpowder causing an explosion, or a brain event causing pain), in intentional causation the cause is the person whose intention it is, a substance and not an event. A person having an intention (in acting) is simply that person intentionally exercising causal power. In chapter 6 I move on to the issue of the nature of the substance, the human person to whom pure mental events (including intentions) belong. I argue that each human person can pick out themselves by an informative designator (e.g., 'I'), whereas others can only pick out a person by an uninformative designator – since each of us knows the necessary and sufficient conditions for being who we are, and others do not. Hence since it is logically possible that I should exist without a body, it is also metaphysically possible. Each of us is a pure mental substance, having a soul as their one essential part and a body as a non-essential part; physical properties belong to us in virtue of belonging to our bodies, and pure mental properties belong to us in virtue of belonging to our souls.

Given that human persons cause brain events, the next issue is whether humans are always fully caused to cause the brain events they often do cause (and so the resulting bodily movements) by earlier brain events or mental events. I argue in Chapter 7 that it is most improbable that it could be shown that we are always so caused – since any purported laws of mind/brain interaction would be so complicated that it would be almost impossible to get enough evidence to establish or refute them. Hence when it seems to us that we are causing our bodily movements (and so the brain events which cause them) without being fully caused to do so, we should – by the principle of credulity – believe that that is how it is, and so that we have (in this crucial sense) free will. I then proceed to argue in Chapter 8 that, given that that is our situation, we are morally responsible for our actions – guilty and deserving blame for doing what we believe wrong, meritorious and deserving praise for doing what we believe to be good actions beyond obligation.

SWINBURNE ON SUBSTANCE DUALISM

LYNNE RUDDER BAKER

University of Massachusetts Amherst

Richard Swinburne's *Mind, Brain and Free Will* is a tour de force. Beginning with basic ontology, Swinburne formulates careful definitions that support his mature philosophy. He is well-known for his views on free will and substance dualism. In *Mind, Brain and Free Will*, he revisits these issues among others, because he now has 'deeper and stronger' arguments for his position. Here I want to discuss substance dualism.

The first sentence of the book says that its focus is 'the nature of human beings – whether we are merely complicated machines or souls interacting with bodies'[1] (Swinburne 2013: 1). Since I do not believe that this dichotomy is exhaustive, I shall first examine Swinburne's argument for the claim that human beings are 'souls interacting with bodies', which he interprets as substance dualism. Then, I shall offer a sketch of a different view – my own – which is materialistic in the sense that it makes reference neither to any immaterial concrete objects nor to any immaterial properties (whatever that might mean). I do not say that my materialistic view is a complete metaphysics, inasmuch as that it pertains only to the natural world. I leave it open whether there is a supernatural world.

Since I am going to offer an alternative to Swinburne's substance dualism, let me begin by enumerating ways in which I am in full agreement with him. I fully agree with Swinburne that it is *not* the case that 'mental events are merely brain events', and it is *not* the case that 'I am the same thing as my body' (Baker 1995; Baker 2000). With Swinburne,

[1] Since Swinburne begins with a dichotomy between human beings' being machines or souls, and goes on to argue that we are mental substances, I assume that he takes mental substances to be souls (i.e., immaterial substances).

I reject a Humean theory of causation in favour of a causal-powers view (Baker 2007a). And I thoroughly agree with Swinburne's rejection of the principle of the causal closure of the physical (Swinburne 2013: 104-123; cf. Baker 2009). Moreover, I also have a 'simple view' of personal identity (Baker 2012).[2] Now to turn to Swinburne.

In Chapter One, Swinburne takes considerable pains to set out his ontology. The history of the world is 'all the events that have occurred, are occurring, or will occur' (Swinburne 2013: 9). Events have as constituents substances, properties and times. So, the categories needed to tell the whole history of the world – the data of mental and physical life – are substance, property and time. Since Swinburne says that 'we can cut up the world into substances in different arbitrary ways and still tell the same world story' (Swinburne 2013: 38), we need a way to determine when 'two referring descriptions pick out the same property, substance or whatever'. Two referring descriptions pick out the same property (or whatever) if and only if that property 'can be designated by the same informative designator' (Swinburne 2013: 10). Swinburne's dualism rests on his claim that mental properties and physical properties cannot be picked out by the same informative designator. (Swinburne 2013: 69)

I. INFORMATIVE DESIGNATORS

Since Swinburne's notion of informative designators carries a lot of weight, we face the question of what an informative designator is. Swinburne says:

> For a rigid designator of a thing to be an informative designator it must be the case that *anyone who knows what the word means* (that is has the linguistic knowledge of how to use it) knows a certain set of conditions necessary and sufficient (in any possible world) for a thing to be that thing (whether or not he can state those conditions in words).[3] (Swinburne 2013: 12; my emphasis)

[2] However, since I believe in ontological vagueness, I disagree with Swinburne's account of the history of the world that holds that 'each substance com[es] into existence at a certain instant', which in 'the precise mathematical sense of the word [is] not temporally extended' (Swinburne 2013: 8). This is all the more surprising since he also says, 'All events take time.' (Swinburne 2013: 148)

[3] I assume that the informative-/uninformative-designator distinction applies only to rigid designators.

He elaborates: 'Two informative designators are logically equivalent if and only if they are associated with logically equivalent sets of necessary and sufficient conditions.' (Swinburne 2013: 12) And mental properties are not identical to physical properties because their informative designators are not logically equivalent. (Swinburne 2013: 68)

Swinburne gives some examples: 'Red' is an informative designator, he says, because anyone who knows what the word means can apply it correctly if she is favourably positioned with faculties in working order and not subject to illusion. By contrast – another example of Swinburne's – 'water' (as used in the 18th century) was an uninformative designator of the property of being H_2O because 'however favourably positioned you are and however well your faculties are working you may not be able to identify correctly some liquid not in our rivers and seas as water' (Swinburne 2013: 12). However, as used today, 'water' is an informative designator. Swinburne comments: 'Whether or not a word is an informative designator is a matter of the rules for its current use in the language.' (Swinburne 2013: 14)

The distinction between informative and uninformative designators is important for Swinburne, not only because it is used to individuate properties, but also it figures in a restriction on telling the whole story of the world: We must specify the events that make up the history of the world in terms of properties and substances picked out by informative designators. Because 'such designators will pick out the same properties and so on [if and only if] they are logically equivalent', this restriction enables us 'to tell the whole history of the world ... by listing a subset of events which entails all events' (Swinburne 2013: 23).

Let us now turn to some potential problems with Swinburne's characterization of an informative designator in terms of a speaker's knowledge of a set of necessary and sufficient conditions for its application.

First, Swinburne says, 'My knowledge of how to use "I" ... means that I know the nature of what I am talking about when I use the word.' (Swinburne 2013: 158) Later, Swinburne argues that it is the nature of a human person to be a mental (immaterial) substance. And what's true for me about my use of 'I' is true for everyone who uses 'I'. (Swinburne 2013: 158) Surely, if one's knowledge of how to use 'I' meant that one knew 'the nature of what [one is] talking about', then on Swinburne's view, everyone who used 'I', should know that she was a mental substance. But

that is clearly wrong. I (LB) can use 'I' competently, but I do not believe that I am a mental (i.e., immaterial – see note 1) subject.

Second, I have the linguistic knowledge of how to use 'arthritis', but I do not know 'a certain set of necessary and sufficient conditions (in any possible world) for a thing to be' arthritis – even though I know that arthritis is some painful condition of the joints. I don't even know any necessary and sufficient conditions for a thing to be red (Swinburne's example of an informative designator). Is a reddish-orange poppy red? Is a reddish-yellow bruise red? Of the words that I use competently, there are not many (if any) for which I know any necessary and sufficient conditions. (cf. Merricks 1998)

Third, Swinburne's characterization of informative designators not only leaves indeterminacies that are relative to 'the rules for its current use in language', as in the case of 'red'; but his characterization also relativizes an informative designator to a *speaker's* current knowledge. A young child can competently use 'water' (in its current usage) without knowing that water is H_2O; so in the child's mouth, 'water' is an uninformative designator; but in the mouth of her mother, who knows that water is H_2O, 'water' is an informative designator. It follows that, on Swinburne's characterization, a single term can be used competently by two people at the same time in the same linguistic community, and in one case it is an uninformative designator and in the other case it is an informative designator. The relativity of a word's being an informative designator, not only to time but also to speaker, seems to vitiate the usefulness of informative designators for telling the whole story of the world.

Swinburne does say that an informative designator is such that anyone who 'has the linguistic knowledge of how to use it' knows a certain set of conditions necessary and sufficient for its application, 'whether or not he can state those conditions in words' (Swinburne 2013: 12). In all of my examples – e.g., 'red', 'I', 'arthritis' – the speaker has the linguistic knowledge of how to use a term without knowing, even in some implicit way, any necessary and sufficient conditions for the term's application.

In sum, if neither 'red', nor 'I', nor 'arthritis' is an informative designator, and if 'water"s being an informative designator is relative, not only to time, but also to speaker, it is difficult to see that there are any informative designators – certainly not enough to tell the whole story of the world – on Swinburne's characterization. Hence, Swinburne's notion of an informative designator does not seem up to the task of picking out all the properties needed to tell the history of the world.

II. IDENTITY CRITERIA FOR COMPOSITE SUBSTANCES

Swinburne has a principle of identity for composites. (Composites are substances with parts – like human organisms and artefacts.) One way to express his principle is that 'there is no more to any substance than its parts (e.g., fundamental particles) and the way those parts are arranged' (Swinburne 2013: 35). Swinburne says that our normal criteria for the identity of a ship over time in terms of parts are too vague to resolve the 'ship of Theseus' puzzle. We can make them more precise, Swinburne says, in either of two ways: the continuity of planks arranged in a certain way determines the identity of the ship; or the identity of the planks determines the identity of the ship. 'We can tell the story either way without anything being omitted.' (Swinburne 2013: 31)

Can this claim about the ship of Theseus really be true, on Swinburne's view? It may make a difference in the history of the world which ship is the original ship after the change of planks – the ship with reassembled old planks or the ship with the new planks. To see this, suppose that, although the original ship was owned by Theseus, the replacement planks were owned by Minos. If we said that the continuously-existing ship was the ship of Theseus, then Theseus would own a ship all of whose planks were owned by Minos. If the original ship of Theseus had been insured, then there would be a difference in which resulting ship ended up insured. If an insurance claim were filed, which ship was identical to the ship of Theseus would make a difference in the history of the world. So, we couldn't tell the story 'either way without anything being omitted'.

Swinburne holds that we can tell the history of the world with stories in terms of fundamental particles, and also with stories in terms of 'organisms and artefacts' (and, I assume, mental substances). (Swinburne 2013: 32) However, if substances are no more than their parts in certain arrangements, it seems to follow that artefacts cannot be substances. The exclusion of artefacts as substances so construed can be illustrated by a fanciful example of two (non-identical) kinds of artefacts made of duplicate qualitatively identical parts arranged in the same way: Suppose that someone invented a device that mixed water and air and was used in a process of making soft drinks; call the device a 'drinkalator'. Suppose that it turned out that exactly the same physically possible structures that could be drinkalators could also be carburettors. But drinkalators are not identical to carburettors, which mix gasoline and air in automobiles; both types of artefacts are distinguished by the intentions and practices of the

designers, manufacturers, and users. The identity of an artefact is not determined by its parts and their arrangement: 'X is a carburettor' and 'X is a drinkalator' have different truth conditions. (Baker 1995: 195–99) No carburettor is identical to a drinkalator. So, either carburettors and drinkalators – and artefacts generally – fail to be substances or Swinburne's principle of the identity of composites is incorrect.

Other terminological matters: Swinburne defines a mental property 'as one to whose instantiation in it a substance necessarily has privileged access on all occasions of its instantiation' (Swinburne 2013: 67). And 'A mental substance is one for which the possession of some mental property is essential' (Swinburne 2013: 43). A physical property is defined 'as one to whose instantiation in it a substance necessarily has no privileged access on any occasion of its instantiation' (Swinburne 2013: 68). A pure mental property is defined as one 'whose instantiation in a substance does not entail the instantiation of any metaphysically contingent physical property in that substance' (Swinburne 2013: 68).

I shall construe the definition of a mental property to be logically equivalent to this (which seems to me slightly easier to understand): A mental property is one to whose instantiations a substance that instantiates it has privileged access – i.e., substances have privileged access to the mental properties that they instantiate.

III. SWINBURNE'S ARGUMENTS FOR SUBSTANCE DUALISM

I am not confident that I understand Swinburne's argument for mental substances from synchronic unity, but if I do, it goes like this:[4]

(1) A complete history of the world must include 'among substances with physical properties and so physical parts, substances which are such that events in those parts are the immediate causes or effects of and only of conscious events which are coexperienced with other conscious events belonging to the same substance' (Swinburne 2013: 143).

(2) The property of coexperiencing certain properties at a time 'will delimit the physical boundaries of the substance, and so help to determine which physical properties it possesses' (Swinburne 2013: 143).

[4] Swinburne 2013: 43. This is Swinburne's argument from the synchronic unity of a person. If the argument is sound, it establishes substance dualism.

(3) The property of coexperiencing certain properties at a time is a mental property. [Defn. of mental property plus the fact that one has privileged access to instantiations of coexperiencing.]
(4) If (2) and (3), then 'a substance which has conscious coexperienced properties is ... determined in part by a mental property' (Swinburne 2013: 143).
(5) If a substance which has conscious coexperienced properties is ... determined in part by a mental property, then that substance is a mental substance. [alternative definition of mental property]
(6) Humans 'are mental substances, since their spatial boundaries are determined by a mental property' (Swinburne 2013: 144).

I have a worry about (2), and consequently about (5) and (6). First (2): I wonder if 'delimit' is used equivocally. (2) may be true if 'delimit' is a causal idea, as in interactive dualism; but in order to support (5), Swinburne needs a stronger ontological reading of (4). Determination is an ontological idea, not merely a causal one. If coexperiencing certain properties at a time will 'help to determine which physical properties it possesses', then the connection between the coexperiencing and the physical properties that the coexperiencing helps to determine must be stronger than merely causal.

To put it another way, (2) is true only if 'delimit' is understood causally; but if 'delimit' is understood causally, then (2) does not support the conclusion (6). So, either the argument has a false premise or is invalid. Either way, it is unsound.

Does Swinburne's argument from diachronic unity of a person fare any better? What follows is my best idea of how the argument goes (again, I am uncertain):

(1) My use of the word 'I' is as an informative designator. (Swinburne 2013: 158)
(2) If (1), then 'I know the nature of what I am talking about when I use the word' 'I'. (Swinburne 2013: 158)
(3) What's true for me about my use of 'I' is true for everyone who uses 'I'. (Swinburne 2013: 158)
(4) If (2) and (3), then '[e]ach of us ... can continue to exist without any continuity of brain, memory or character'. (i.e., The 'simple theory of personal identity is correct'.) (Swinburne 2013: 158)
(5) If each of us ... can continue to exist without any continuity of brain, memory or character, then each of us is a 'particular subject

of experiences (that is, of conscious events), and so a mental substance who exists over time' (Swinburne 2013: 159).

(6) Each of us is a 'particular subject of experiences (that is, of conscious events)'.

(7) If (6), then each of us is a 'particular mental substance who exists over time.' [Swinburne 2013, 159–60]

(8) Each of us is a 'particular mental substance who exists over time' (Swinburne 2013: 160).

The main problem with this argument, I think, is that the truth of (7) requires that subjects of experience are mental (immaterial) substances – the conclusion of the argument from synchronic identity, which we found to be unsound. (I criticized premise (1) in the section on Informative Designators.)

To sum up, there seem to be difficulties with the application of informative designators to the individuation of properties, with the application of Swinburne's criteria of identity for substances to the ship-of-Theseus case, and with his principle of the identity of composites. Moreover, neither the argument from synchronic unity nor the argument from diachronic unity for substance dualism seems sound. In light of these at least provisional problems, it seems that a view of persons that does not countenance mental substances, but satisfies first-personal desiderata for persons should be a contender.

IV. A CONSTITUTION VIEW OF PERSONS

Since I intend this discussion to focus on Swinburne, I shall sketch my alternative only briefly.

On my view a person has a first-person perspective essentially, and a human person is embodied essentially, although she does not necessarily have the body she has now. The first-person perspective is a dispositional property that has two stages: a rudimentary stage that human infants share with higher nonhuman animals, and a robust stage that human persons develop as they learn a language. (For details see Baker 2013.)

The relation between a person and her body is constitution. Every concrete object that exists is of some primary kind or other, and has its primary-kind property essentially. *Person* is a primary kind; *teacher* is not. Constitution is a contingent relation of unity between things of

different primary kinds. Human persons begin existence constituted by human organisms that support first-person perspectives. (Later, with enough bionic replacements, a human person may come to have a nonorganic body, but the person is always embodied as long as she exists.) The unity of persons, and of much else, is provided by constitution and the first-person perspective. (For details, see Baker 2007b; Baker 2007a; Baker 2013)

Persons are of different primary kinds from bodies, human or otherwise. (Persons have first-person persistence conditions; bodies do not. Persons have innumerable causal powers that bodies lack.) Nevertheless, my view is not a property dualism. On my view, there are not two kinds of properties; there are myriads of kinds of properties, none of which is instantiable only by immaterial substances. (I do not believe that there are any finite immaterial substances.)

My view is nonreductive throughout. I am as nonreductive about artefacts as I am about natural objects: 'Every thing is what it is and not another thing', and that goes for the familiar things that we interact with.[5] Being a dialysis machine is as irreducible (and hence as much are part of basic ontology) as being a person or being a human organism.

Hence, the relevant charge against me is not property dualism; a better criticism is that my ontology includes too many concrete objects and properties, that my ontology is bloated. I accept the charge as the price for being nonreductive. So, I am no dualist; I don't stop at two kinds of substances or properties.

One advantage of the Constitution View over Substance Dualism is that there is not (or rather I cannot think of) any naturalistic way that an immaterial mind could have come into existence, but I can think of a naturalistic way that the first-person perspective could have come into existence: the rudimentary stage of the first-person perspective – consciousness and intentionality – seems to have evolved gradually over many species (from the first stirrings of intentionality in organisms with limited flexibility in responding to their environments to the more developed intentionality of mammals);[6] and the robust stage comes into existence as human beings developed syntactically complex languages with resources for first-person reference and attribution of first-person

[5] The quotation is attributed to Joseph Butler.

[6] I think that there is no doubt that dogs behave intentionally and are conscious, Descartes notwithstanding.

reference. There are thus nonmagical explanations for both stages of the first-person perspective.

Finally, I reject any claim (like Swinburne's) that objects should be understood in terms of their parts. The identity and arrangement of parts do not always (or even usually) determine the identity of wholes. Artefacts and artworks are ontologically intention-dependent; in the absence of intentions (of designers, manufacturers, users, artists), there could be no artefacts or artworks. Even if indiscernible duplicates of, say, dialysis machines spontaneously coalesced in outer space, such objects would not *be* dialysis machines. Dialysis machines are the things they are because they have the intended function to replace lost kidney-function artificially; machines with such intended functions could not exist in the absence of laboratories and medical practices. When we focus on parts and their arrangement, what is essential becomes invisible.

Despite my differences with Swinburne, I believe that he and I share certain *desiderata* in our theories about human persons:

(a) A human person is not identical to her body at any time.
(b) A human person can survive a complete change of body.
(c) Not all truths about persons are truths about bodies.
(d) The property of being a person entails the instantiation of mental properties – at least a rudimentary first-person perspective.
(e) The persistence of persons is primitive (as the Simple Theory of Personal Identity over time implies).

V. CONCLUSION

Mind, Brain and Free Will is a closely argued comprehensive work on major themes of Swinburne's. I have discussed only a few aspects of Swinburne's views bearing on substance dualism. In the belief that we can (and should) reject the dichotomy that Swinburne began his book with – that human beings are either 'merely complicated machines or souls interacting with bodies' – I offered an alternative that I believe achieves much of what substance dualists want.

REFERENCES

Baker, Lynne Rudder. 1995. *Explaining Attitudes: A Practical Approach to the Mind* (Cambridge: Cambridge University Press)

– 2000. Persons and Bodies: A Constitution View (Cambridge: Cambridge University Press)

– 2007a. 'Naturalism and the First-Person Perspective', in How Successful Is Naturalism?, ed. Georg Gasser, (Frankfurt: Ontos-Verlag) pp. 203–227

– 2007b. 'Persons and the Metaphysics of Resurrection', Religious Studies, 43 (3): 333–348

– 2009. 'Nonreductive Materialism', in The Oxford Handbook of Philosophy of Mind, ed. Brian P. McLaughlin, Ansgar Beckermann, and Sven Walter (Oxford: Clarendon Press), pp. 109-127

– 2012. 'Personal Identity: A Not-So-Simple Simple View', in Personal Identity: Complex or Simple?, ed. Matthias Stephan and Georg Gasser (Cambridge: Cambridge University Press), pp. 179-191

– 2013. Naturalism and the First-Person Perspective (New York: Oxford University Press)

Merricks, Trenton. 1998. 'There Are No Criteria of Identity over Time', *Nous*, 32 (1): 106–124

Perry, John. 2002. 'The Sense of Identity', in *Identity, Personal Identity, and the Self* (Indianpolis, IN: Hackett Publishing Company, Inc.), pp. 214–243

Swinburne, Richard. 2013. *Mind, Brain and Free Will* (Oxford: Oxford University Press)

SWINBURNE ON SUBSTANCES, PROPERTIES, AND STRUCTURES

WILLIAM JAWORSKI

Fordham University

Mind, Brain, and Free Will, Richard Swinburne's stimulating new book, covers a great deal of territory. I'll focus on some of the positions Swinburne defends in the philosophy of mind. Many philosophers are likely to have reservations about the arguments he uses to defend them, and others will think his basic position is unmotivated. My goal in this brief discussion is to articulate some of the reasons why.

I. SWINBURNE'S ARGUMENT FOR PROPERTY DUALISM

Swinburne defends substance dualism, the claim that we are pure mental substances, ones that have only pure mental properties essentially. Pure mental properties are properties whose instantiation does not entail the instantiation of any physical properties. The property of seeing a desk is an impure mental property on Swinburne's view since necessarily someone can see a desk only if he or she is causally affected by one, and being causally affected by a desk is a physical property. Seeming to see a desk, however, entails no such condition; I can seem to see a desk even if there is no desk causally affecting me. A property *P* is mental, according to Swinburne, exactly if a substance *S* which instantiates *P* necessarily has privileged access to *P*'s instantiation. If *P* is mental, then *S* can in principle know that he or she instantiates *P* in the same ways other substances know it, but there will be an additional way that *S* knows that *P* is instantiated, namely by experiencing it. Properties that are not mental are either physical or neutral. Physical properties are ones to which necessarily a substance does not have privileged access, and neutral properties are ones to which some substances have privileged

EUROPEAN JOURNAL FOR PHILOSOPHY OF RELIGION 6/2 (SUMMER 2014), PP. 17-28

access but others don't (disjunctive properties such as the property of *being in pain or weighing 50kg* are examples; both I and the desk instantiate this property, but I have privileged access to its instantiation and the desk does not).

Swinburne argues that mental and physical properties are distinct. This follows a priori, he says, from his definitions of 'mental' and 'physical' in conjunction with his account of properties. Properties, according to Swinburne, are universals, and they are abundant: any predicate whatsoever, it seems, picks out a property, and two properties are identical exactly if their informative designators are logically equivalent (this is no slip of the tongue: Swinburne is quite serious about stating conditions under which two substances, or two properties, or two events are the same, and by this he appears to mean numerically the same; although he does not endeavour to explain how two things can be one). An informative designator of some substance, property, or event, *X*, is a rigid designator of *X* that expresses *X*'s nature or essence, the conditions metaphysically necessary and sufficient for *X* being what it is. If 'red' is an informative designator, then the criteria for correctly applying it include the conditions that are metaphysically necessary and sufficient for being red. If those criteria are not logically equivalent to the criteria for correctly applying, say, 'reflects such-and-such wavelengths of light', then being red and reflecting such-and-such wavelengths of light must be different properties. This is in fact the case, says Swinburne, for knowing that something reflects such-and-such wavelengths of light does not entail that it looks a certain way to most people. The latter is an additional fact about an object that goes beyond its reflecting such-and-such wavelengths of light. It is thus true a priori that red is not identical to reflecting such-and-such wavelengths of light. It is also true a priori that redness does not supervene on reflecting such-and-such wavelengths of light, for it is logically possible that objects which reflect such-and-such wavelengths of light might look differently to us. What is true of red, according to Swinburne, is also true of mental properties; they are neither identical to nor supervenient upon physical properties. The criteria for correctly applying mental predicates and terms are not logically equivalent to the criteria for correctly applying physical ones; knowing that the former apply does not entail knowing that the latter apply and vice versa. Consequently, mental predicates and terms must designate properties distinct from those designated by physical predicates and terms.

The crucial premise in Swinburne's argument is that terms like 'red' and 'pain' are informative designators, that in applying them we grasp the conditions that make redness and pain what they essentially are. The obvious challenge to this premise comes from people like Kripke and Putnam who have built philosophies of language around examples in which competent speakers apply terms correctly while failing to grasp the essence of what the terms apply to. The term 'water' is an example: many people use it correctly to refer to water without knowing that water consists of hydrogen and oxygen atoms. Many philosophers are inclined to view terms like 'red' and 'pain' by analogy with 'water'. We can be competent in applying them and making simple inferences to and from their application (if something is bright red, then it is red; if something is red, then it is coloured, and so on) while yet remaining ignorant of what makes them what they are. If that is the case, then Swinburne's property dualism cannot be established a priori, as he claims.

Swinburne acknowledges that 'water' is not an informative designator in his sense. Prior to discovering that water was H_2O, he says, people did not fully know what they meant by 'water'. But, he insists, terms like 'red' and 'pain' are different. Language users who correctly apply these terms know the essence of what they refer to. The reason, says Swinburne, is that competent speakers who are favourably positioned with their faculties working properly and not subject to illusion know when and where terms like 'red' and 'pain' apply, they are able to make simple inferences to and from their application, and this kind of competence implies knowing what redness and pain are. It certainly seems plausible that correctly using a term like 'red' implies knowing in some sense what red is. If I am a competent speaker, favourably positioned with faculties in working order and not subject to illusion, then surely I can identify which things in the environment are red. It doesn't follow from this, however, that I know the conditions that are metaphysically necessary and sufficient for being red. By analogy, if you tell me that only people with a yellow ticket may enter the reception I can pick out the people who may enter the reception without knowing what earned them their yellow tickets. Being a competent doorman requires only being able to identify which people have a yellow ticket; it does not require me to know how or why they got it. Kripke and Putnam give us reason to think something analogous is often true of competent speakers.

Swinburne counters this with the suggestion that I cannot be wrong about something looking red to me. But exponents of the

Kripke-Putnam view can concede this point. Even if I cannot be wrong about something looking red to me, I can still be wrong about something *being* red. One way of developing this idea is to say that terms like 'red' get their meanings from certain prototypical applications. We apply 'red' to objects whose surfaces reflect such-and-such wavelengths of light to the eyes of such-and-such observers under such-and-such conditions (Swinburne countenances some of these conditions with the expression 'favourably positioned with faculties in working order and not subject to illusion'). Things get called 'red' to the extent that they resemble the prototypes. Sometimes, however, we find ourselves in circumstances that do not satisfy the conditions that define the prototypes; objects look red to us that are not 'really' red; that is, that we would not call red if we saw them in prototypical circumstances (I see your tie in unusual lighting, and believe it is red, but recognize my mistake when we enter white light). If 'red' means something like this, then it is possible that redness might be identical to a surface property whose essence we do not know even though we can competently use 'red' to refer to objects having it. Likewise, redness could be a higher-order property such as the property of having a surface property that reflects such-and-such wavelengths of light in prototypical circumstances. In either case, Swinburne's property dualism does not follow.

I confess that I found it difficult to discern an argument in *Mind, Brain, and Free Will* that would rule out views of this sort. Two considerations that Swinburne advances seem to fall short. One is the point mentioned earlier that how something looks to people appears to be an extra fact about it beyond its having a certain reflectance spectrum. Exponents of the Kripke-Putnam view can readily concede this point. When we are talking about how things look to us, they can say, we are talking about a relation between a surface and an observer as opposed to a property of the surface taken by itself. Since relations are not intrinsic properties, facts about relations are not facts about intrinsic properties. It does not follow from this, however, that redness is not an intrinsic physical property, nor that it is not a physical relation.

Second, Swinburne appeals to the kind of argument J. J. C. Smart attributed to Max Black.[1] If 'red' and 'reflecting such-and-such wavelengths of light' are not logically equivalent, as exponents of the Kripke-Putnam

[1] J. J. C. Smart, . 'Sensations and Brain Processes', in *The Philosophy of Mind*, V.C. Chappell, ed. (Englewood Cliffs: Prentice-Hall, Inc., 1962), pp. 160-172.

account must concede, then the criteria for correctly applying the one term cannot be identical to the criteria for correctly applying the other. The differences between these criteria must ultimately boil down to differences among properties. There must be certain properties whose recognition by a competent speaker underwrites the correct use of 'red', and certain properties whose recognition by a competent speaker underwrites the correct use of 'reflecting such-and-such wavelengths of light'. These properties, moreover, must be different, for it must be possible for competent speakers to recognize the instantiation of the properties that sanction the use of the one term while at the same time not recognizing the instantiation of the properties that sanction the use of the other. Consequently, it looks as though exponents of the Kripke-Putnam account must endorse a dualism of properties.

The problem with this argument is that by itself it does not support the dualism of physical and nonphysical properties that Swinburne looks to defend; it supports only the thesis that the properties which underwrite the correct use of 'red' must be different from those which underwrite the correct use of 'reflecting such-and-such wavelengths of light', and it does not follow from this alone that the former properties must be nonphysical and the latter physical. To appreciate this let us imagine for the sake of argument that physicalism is true, and that P_1, P_2, ... , P_n are all the properties that exist. Since physicalism is true by assumption, P_1, P_2, ... , P_n are all physical properties. Suppose now that competent speakers apply the predicate 'is red' to something if and only if it instantiates P_1, and that they apply the predicate 'reflects such-and-such wavelengths of light' to something if and only if it instantiates P_2. In that case, the properties that competent speakers must recognize to correctly apply the one predicate differ from the properties they must recognize to correctly apply the other, yet it does not follow that either predicate expresses a nonphysical property, for by assumption all properties are physical. The dualism that the argument supports is merely a dualism of conditions for correctly applying predicates or terms – a difference in what some philosophers call 'modes of presentation'.[2] This does not by itself support a dualism of physical and nonphysical properties. To derive that conclusion a further premise is needed to the effect that the

[2] John Perry, *Knowledge, Possibility, and Consciousness* (Cambridge, Mass.: MIT Press, 2001); Ned Block, *Consciousness, Function, and Representation: Collected Papers, Volume 1* (Cambridge, Mass.: A Bradford Book, 2007), Chapter 21, pp. 435-490.

properties which fix the referents of terms like 'red' must be nonphysical. But it is difficult to see how exponents of the argument can endorse such a premise without either begging the question against their opponents or making tendentious assumptions about properties.[3]

Swinburne replies that critics who endorse the foregoing line of argument end up multiplying entities beyond necessity since modes of presentation are extra theoretical posits. But there are two things to say in response. First, it is not clear that modes of presentation are in fact extra theoretical posits. Modes of presentations are typically posited to explain why identity statements such as 'Cicero is Tully' are informative while identity statements such as 'Cicero is Cicero' are not. If properties F_1, F_2, ... , F_n are all of Cicero's properties, and competent speakers fix the referent of 'Cicero' by appeal to F_1 and the referent of 'Tully' by appeal to F_2, then the informativeness of 'Cicero is Tully' has a straightforward explanation: it is informative because the referents of 'Cicero' and 'Tully' are fixed by appeal to different properties. Likewise, if F_1, F_2, ... , F_n are all the properties had by the property of redness (Swinburne admits that there are properties of properties), then one can explain the informativeness of the putative identity statement 'Redness is reflecting such-and-such wavelengths of light' in an analogous way: competent speakers fix the referent of 'redness' by appeal to F_1, and they fix the referent of 'reflecting such-and-such wavelengths of light' by appeal to F_2. Some philosophers choose to call properties like F_1 and F_2 'modes of presentation'. Using this terminology does not add anything to the theoretical apparatus they must already endorse to make sense of informative identity claims, so it is not clear that modes of presentation amount to 'extra' theoretical posits, as Swinburne claims. Second, there are reasons to think that philosophers like Swinburne must themselves be committed to positing modes of presentation. Ostensibly any philosophers who want to explain informative identity statements must posit such modes or something very much like them, and presumably that includes philosophers like Prof. Swinburne. In addition, without modes of presentation Black's argument cannot get off the ground, for it requires that different properties fix the referents of expressions like 'red', on the one hand, and expressions like 'reflects such-and-such wavelengths of light', on the other. This gives us reason to think that Swinburne himself must tacitly endorse modes of presentation even if he chooses not to call them that.

[3] Block (ibid.) argues for this in detail.

Based on the foregoing considerations it remains unclear how Swinburne rules out the Kripke-Putnam view. Consequently, it remains unclear why we should accept that terms like 'red' and 'pain' are informative designators in Swinburne's sense, and as a result it remains unclear why we should accept the property dualism Swinburne endorses.

II. SWINBURNE'S ARGUMENT FOR SUBSTANCE DUALISM

Similar worries attend Swinburne's argument for substance dualism. To appreciate this it's helpful to consider Swinburne's philosophical forebear, Descartes, and the exchange he had with his contemporary Antoine Arnauld, author of the fourth set of objections to the *Meditations*. Descartes argued that we could exist without bodies because thought was our only essential property. To support this premise Descartes compiled a list of the properties that people initially took themselves to have and argued that we could clearly and distinctly conceive of ourselves existing without each of the properties on the list – each, that is, except thought. We cannot form a (first-personal) conception of ourselves without thought, so Descartes concluded that we could not exist without it, and since this was the only property of which this was true, it must be our only essential property. Arnauld's worry about Descartes' argument was that the initial conception people had of themselves might be in some way impoverished, that people might have properties, perhaps even essential ones, of which they were entirely ignorant, and which therefore did not appear on Descartes' list. Since these properties would not have been subjected to Descartes' conceivability test, Descartes' conclusion would not follow; we could have essential properties of which we are entirely unaware. Descartes conceded to Arnauld that the conception we started with would have to comprise all our essential properties (it would have to be 'complete' as he put it), but he never explained why we should suppose that the conception we have of ourselves is complete in fact.

What was true of Descartes vis-à-vis Arnauld seems true of Swinburne vis-à-vis Kripke and Putnam. Kripke and Putnam made it evident that Arnauld's worry was not an abstract possibility; rather, our best concrete efforts at understanding the world have revealed that things often have properties, including essential ones, of which we can remain entirely unaware in our pedestrian dealings, and which for that reason do not factor into the meanings of the terms we use to refer to them. Just as it

is unclear how Swinburne rules out a Kripke-Putnam account when it comes to properties like redness, it is unclear how he rules it out when it comes to substances like you and I. He argues that I cannot be wrong about what 'I' refers to in the way I can be wrong about what 'water' refers to:

> 'I' or 'Richard Swinburne' as used by me ... seem to be informative designators. If I know how to use these words, then ... I can't be mistaken about when to apply them ... I cannot know how to use the word 'I' ... and still wonder whether it is I or someone else who is having that event ... My knowledge of how to use 'I', like my knowledge of how to use 'green' and 'square', means that I know the nature of what I am talking about when I use the word. (p. 158)

But it is unclear how Swinburne's conclusion follows from this. Even if I am necessarily right about what 'I' refers to, I can still be wrong about what essential properties I have. (By analogy, I can correctly admit all and only the yellow ticket holders while yet incorrectly conjecturing how they got the tickets.) It remains unclear, therefore, why we should believe that the competent use of 'I' implies knowing what I essentially am, and as a result it remains unclear how we ought to arrive at the conclusion that we are pure mental substances, ones that have only pure mental properties essentially.

The work of Arnauld, Kripke, and Putnam points to a general worry that has confronted Cartesian projects like Swinburne's since the seventeenth century. The arguments Cartesians advance typically assume that our concepts and the predicates and terms that express them leave out nothing essential to the things we think and talk about, that they are 'complete' in Descartes' sense, and as a result we can know how matters stand with regard to ourselves, our bodies, our powers, properties, and so on, merely by consulting our language and concepts. It is this assumption that Swinburne encapsulates in his claim that 'red', 'pain', 'I', and other terms are informative designators, and that I've suggested he has failed adequately to defend.

The foregoing remarks criticize the idea that mental properties are our only essential properties, but are mental properties essential to us at all? I confess that I found Swinburne's argument for this claim rather difficult to follow. Its main premises appear to be these:

(1) If humans coexperience conscious mental events of different kinds (if, for instance, they simultaneously see the trees outside

and smell the coffee brewing), then mental properties determine the physical boundaries of a substance.

(2) If mental properties determine the physical boundaries of a substance, then they in part determine what physical properties that substance has.

(3) If mental properties in part determine what physical properties a substance has, then that substance is a mental substance; that is, it has mental properties essentially.

The reasoning behind (1) appears to be this: Suppose that I coexperience mental events m_1 and m_2. We know empirically that mental events are caused by brain events, so let us suppose that m_1 and m_2 are caused by brain events b_1 and b_2, respectively. Since m_1 and m_2 are coexperienced by me, the parts of the brain involved in b_1 and b_2 must both be parts of me (if they were not, then the mental events they caused would not be coexperienced by one and the same substance as they are by me). The mental events I coexperience thus determine at least in part which physical things are parts of me. If that is true, then there is good reason to endorse premise (2), for if mental properties play a role in determining what parts I have, and those parts have physical properties which are attributable to me, then mental properties play a role in determining what physical properties I have as well. But now we come to a sticking point. If my mental properties partly determine what physical parts I have and hence what physical properties I have, it is still not evident how this implies that I have mental properties essentially, as premise (3) claims. To arrive at this conclusion Swinburne must assume that I have no physical properties other than those which are determined by my mental properties. It is plausible to suppose that I have no physical properties other than those which are determined by my physical parts, but why should we suppose that the only parts I have are ones which are determined by my mental properties? Couldn't I have parts which are mentally irrelevant, which are not involved in events that cause or are caused by mental events – hair, fingernails, white blood cells, even parts of the nervous system such as glial cells? And couldn't the continued existence of some of these mentally irrelevant parts be sufficient for my continued existence (as some animalists claim about the brainstem)? I couldn't discern an argument in *Mind, Brain, and Free Will* that would rule this out, so it remains unclear why we should accept premise (3), and hence why we should accept that we have mental properties essentially.

III. PHYSICAL PHENOMENA AND STRUCTURE

A final point concerns Swinburne's notion of bodies and physical phenomena generally. A substance is mental according to Swinburne if it has mental properties essentially, and it is physical otherwise. This might strike some readers as a rather odd way of defining 'physical', for suppose that I am a ghostly being made of ectoplasm with properties that physics cannot even in principle describe, and that I have mental properties but only contingently (when, say, the ectoplasm achieves a certain state). By Swinburne's definition I count as a physical substance (and not a mental one) despite having no properties at all that can be described by physics. The reason for this odd result is that for Swinburne being physical has nothing to do with physics. His definition is like the Amish definition of 'English'. For the Amish, being English has nothing to do with being from England; it is simply not being Amish. Likewise, for Swinburne being physical is simply not being essentially mental. This is awkward for philosophers who are careful to distinguish biological, chemical, and other special scientific properties from physical ones (just as I imagine the Amish use of 'English' must be awkward for people who are careful to distinguish the Welsh, the Scottish, and others from the English).

This definitional point is important because I think it reveals a general tendency to overlook important distinctions within the natural world. For Swinburne, as for many other Cartesians, there is nothing about human bodies that would set them apart as special denizens of the physical universe. Contrast this with a view (one might call it a contemporary hylomorphic view) which claims that structure is a basic ontological and explanatory principle, one that concerns both what things essentially are and what they can do. Put a human in a leak-proof bag and then squash it with several tons of force. The contents of the bag no longer include a human being, nor can those contents think, feel, and act as they once could. What explains the difference pre- and post-squashing? Since the physical materials remain the same (none leaked out) we want to say that what changed was simply the way those materials were organized or structured, that this organization or structure was responsible not only for there being a human before the squashing, but also for that human having the distinctive capacities it had. A view along these lines has some empirical backing, as William Bechtel, a philosopher of neuroscience, observes:

> [T]he organization of ... components typically integrates them into an entity that has an identity of its own ... Organization itself is not something inherent in the parts ... Accordingly, investigators who already understand in detail how the parts behave are often surprised by what happens when they are organized in particular ways ... In virtue of being organized systems, mechanisms do things beyond what their components do ... Not only can one study the performance of a mechanism without knowing its component parts and their operations, but what the mechanism as a whole does is typically quite different than the operations performed by its parts ... As a result, organized mechanisms become the focus of relatively autonomous disciplines ... This autonomy maintains that psychology and other special sciences study phenomena that are outside the scope of more basic sciences.[4]

According to Bechtel, a complex whole – what he calls a 'mechanism' – such as an organism, has an organization that confers on it capacities that are not had by its parts taken in isolation, and that cannot be reductively explained in terms of lower-level sciences. His work and that of others suggest that our best empirical descriptions and explanations of living behaviour posit organization (or structure, arrangement, order, configuration) as a basic ontological and explanatory principle.

Swinburne concedes that wholes are composed of parts arranged in certain ways, and that parts sometimes behave differently when incorporated into larger wholes. 'Nevertheless,' he says, 'the causal properties of larger substances such as organisms are just the causal properties of their parts ... ' (p. 32) For Swinburne, arrangement makes no causal difference to things. It might be an ontological principle insofar as it is in part what makes a whole what it is, but it is not an explanatory principle on Swinburne's view since it does not confer any powers on a whole beyond those conferred by its parts. Swinburne is not alone in thinking this. His position is the norm among Cartesians and also among physicalists. Yet if the hylomorphic alternative just described is viable, it could relieve some of the anxieties that motivate both substance dualism and physicalism, for it implies an antireductionism that preserves what is special about human existence without denying humans' essential

[4] William Bechtel, 'Reducing Psychology while Maintaining its Autonomy via Mechanistic Explanations', in *The Matter of the Mind*, Maurice Schouten and Looren de Jong, Huib, eds. (Malden, Mass.: Blackwell Publishing, 2007), pp. 172-198 (pp. 174, 185-186).

materiality.[5] Distinctively human traits, it says, are due to our distinctively human structures, structures that are nevertheless essentially embodied in physical materials. And since structures are basic ontological and explanatory principles on the hylomorphic view, ones that confer powers on a thing beyond those conferred by its composing materials, the view is robustly antireductive. Seen by comparison with a view like this, the project of trying to preserve human distinctiveness by denying essential human materiality begins to lose some of its appeal.

There is clearly a great deal more to be said on this point and on *Mind, Brain, and Free Will* in general, but I hope what I've said is enough to contribute in a small way to a further discussion of Swinburne's work.

[5] I take what is special about human existence to include eschatological factors. Elsewhere, for instance, I've argued that a contemporary hylomorphic view is compatible with the doctrine of the resurrection; see my 'Hylomorphism and Resurrection', *European Journal for Philosophy of Religion*, 5 (2013), 197-224.

SWINBURNE'S NEW SOUL: A RESPONSE TO *MIND, BRAIN AND FREE WILL*

JAMES K. DEW, JR.

Southeastern Baptist Theological Seminary

Richard Swinburne's recent book *Mind, Brain, & Free Will* is a welcomed addition to the discussion of philosophical anthropology by those of us who are not convinced by physical or material accounts of human persons. And though the book makes an important argument for substance dualism, it has much to say about other important aspects of philosophy as well. As Swinburne notes, *Mind, Brain, & Free Will* advances, develops, and occasionally re-directs many of the arguments he first gave in *The Evolution of the Soul*.

In *Mind, Brain, & Free Will*, Swinburne's primary objective is to argue for a substance dualist perspective of human persons. On this view, persons are pure mental substances (or souls), which are distinct from any physical parts or properties and also distinct from any specific mental properties. In his words: 'Each person has a "thisness", a uniqueness, which makes them the person they are quite apart from the particular mental properties they have and any physical properties (and any thisness) possessed by their body.'[1] In this view, the soul of a person is her essential part, while the body is hers only contingently. He adds, 'My soul therefore carries my "thisness". However – given the normal understanding of a human being on earth as constituted (in part) by a body – it follows that humans, unlike other possible pure mental substances such as ghosts or poltergeists, each have their body as a contingent part.'[2]

[1] Richard Swinburne, *Mind Brain, & Free Will* (Oxford: Oxford, 2013), p. 165.
[2] Ibid., p. 170.

As one who comes from a Thomistic perspective, there is much to Swinburne's book with which I agree. For example, I too think that we really do have immaterial souls that are different from our physical bodies and that we are capable of disembodied existence and surviving death. I differ with Professor Swinburne, however, over what constitutes a person and the contingency of our bodies. In my view, substance dualism seems to diminish the important role of the human body. But, my motives for that position are largely theological in nature, and as such I will focus on other things in this review.

For the remainder of this review, I will focus my comments on specific places in Swinburne's *Mind, Brain & Free Will* where I find myself unconvinced or having significant questions. Two such places come to mind: (1) Swinburne's gradual brain replacement thought experiment – what I take to be an important part of his overall argument in chapter six for substance dualism, and (2) the place and importance of mental properties in Swinburne's view of what it means to be a particular person.

SWINBURNE AND GRADUAL BRAIN REPLACEMENT

An important piece of Swinburne's case for substance dualism is his argument from the possibility of gradual brain replacement over time. He asks us to consider the following possibility:

> Suppose that P_1 undergoes an operation in which a small diseased part of his or her brain (a tenth of the whole brain) is replaced by a similar part from another brain (perhaps that of a clone of P_1) .... But now suppose that each year a different tenth of P_1's brain is removed and replaced by similar parts from another brain (perhaps that of a different clone of P_1 on each occasion). At the end of ten years there is a person whose brain is made of entirely different matter. It seems at least logically possible that – because the process has been gradual and each new part has become integrated into the brain before a new operation is done – the resulting person is still P_1.[3]

Swinburne then adds a further condition: 'Now suppose that during each of the ten operations in which brain parts are replaced, the patient remained conscious and has a series of overlapping conscious experiences lasting for the whole operation.' To clarify, we could summarize it this way:

[3] Ibid., p. 155.

> At t_1, Bob is conscious and his brain is composed of 1, 2, 3, 4, 5, 6, 7, 8, 9, & 10.

Then,

> At t_2, Bob is still conscious, but his brain has part 1 replaced by part 11, such that his brain is now composed of 2, 3, 4, 5, 6, 7, 8, 9, 10, & 11.

Supposing this process repeated itself once every year for ten years as Swinburne suggests, we would eventually end up with:

> At $t_{11,}$ Bob is still conscious but his brain has part 10 replaced by part 20, such that his brain is now composed of parts 11, 12, 13, 14, 15, 16, 17, 18, 19, & 20.

It is important to note that Swinburne's thought experiment postulates a gradual replacement of a brain over the course of ten years. During any given surgical procedure, no more than a tenth of the brain is removed and replaced. If this is really possible, then Swinburne thinks that each of us 'can continue to exist without any continuity of brain, memory, or character. It follows that the simple theory of personal identity is true'.[4]

This is an intriguing thought experiment that – according to Swinburne – leads us to substance dualism. Yet, physicalists are unlikely to be persuaded by this argument. I for one – though not a physicalist myself – suspect that this argument is built upon an unstated assumption about consciousness itself. That is, it seems like this would only favour substance dualism if it were impossible – in the first place – for purely physical organisms to have conscious experiences. If they cannot have consciousness, then an organism going through the surgical process Swinburne describes is not conscious and does not have overlapping conscious experiences. As such, there is no person and no particular human identity. But if purely physical organisms can have conscious experience – which is something he does not seem to say much about – then they may also be able to endure the kind of gradual brain replacement that Swinburne describes while also remaining conscious and having 'overlapping conscious experiences lasting for the whole operation'. I suspect that a physicalist could account for these things for two reasons: First, this scenario is significantly analogous to the metabolic process our bodies use to gradually replace parts over time;

[4] Ibid., p. 158.

Second, some physicalist philosophers have already developed theories that seem to allow for this very thing.

Consider the similarity between Swinburne's surgical process and the metabolic process our body employs to replace old parts with new parts over time. Physical organisms replace parts – through metabolic processes – all the time without ceasing to be the beings that they are. And yet, this happens even though the vast amount of the parts that compose the body at an earlier time are now gone and have been replaced by new parts. And yet, this happens while we are conscious and have overlapping conscious experiences. In short, we have our parts gradually replaced over time without losing consciousness or without ceasing to be the people that we are. Now if, through metabolic processes, physical organisms are able gradually to replace old material with new material over the course of time without ceasing to be the organisms that they are, then why wouldn't this be possible in the case that Swinburne has sketched out in *Mind, Brain, & Free Will*? What if physical organisms have the ability to gradually 'pass on' their identity to the new parts as they come into the organism? What this would require us to say is that a particular stream of consciousness – or first person perspective – is maintained by an ever-changing physical organism that gradually replaces old parts with new parts. This clearly happens in the case of metabolic processes, so why couldn't this also happen in the case a gradual brain replacement through a surgical process? One might say that there is a marked difference between the metabolic and surgical processes since one happens 'naturally' while the other is 'artificial'. But I suspect that this is not a sufficient difference since this is simply a difference in how the material arrives in the organism. Once it is there, the organism incorporates the material in the same way for both cases. All that is important here is that the physical organism be able to gradually change parts without the loss of consciousness or the life of the person in question.

Consider, also, some physicalists and materialists whose accounts of human persons might allow for Swinburne's scenario. Again, in his account a given person (1) has the parts of her brain gradually replaced over ten years until the whole brain has been replaced, and (2) the person maintains consciousness throughout the surgeries and has overlapping conscious experiences. Can physicalists account for these things? I suspect that Lynne Baker's Constitution View just might. In

her approach, the capacity for first-person perspective is the essential criterion for what it means to be a person. To be clear, having first-person perspective is more than just having some mental property or memory. It is the ability that one has to think of herself as herself. Or as she explains, 'A being with first-person perspective not only can have thoughts about herself, but she can also conceive of herself as the subject of such thought.'[5] In other words, first-person perspective requires more than me being able to have desires, intentions and plans, it requires me to realize that *it is me who has such* desires, intentions and plans. In Baker's Constitution View, a person can persist over time even if she does not have an immaterial soul, as long as she maintains first-person perspective. In other words, in light of changes in the human organism that constitutes a specific human person, Baker suggests that all that is necessary for the persistence of the person is continuity of first-person perspective and higher brain functions. She says, 'Suppose that a person slowly had her organs replaced by nonorganic parts, to the point where there was no longer metabolism, circulation, digestion, and so on, but the higher brain function remained and the person's sense of herself was uninterrupted. In this case, the person would persist but the organism would not.'[6] And so, in Baker's view, first-person perspective and higher brain function provide sufficient conditions for a person to persist over time. If her view is correct, then human persons – material persons – could maintain their consciousness throughout the gradual brain replacement scenario that Swinburne describes. As such, Baker's Constitution View might provide everything necessary to account for the persistence – and therefore survival – of a person who undergoes Swinburne's gradual brain replacement.

Peter van Inwagen's materialistic approach might also give materialists everything they need to reject Swinburne's conclusions. Like Baker, his approach allows for the parts of a material organism to change over time and for the organism – in this case a person – to persist throughout the change as long as at each successive stage of the process, the given parts are caught up into the same life as the original organism. He calls this the *Life* principle:

---

[5] Lynne Rudder Baker, 'On Being One's Own Person', in *Reasons of One's Own*, eds. Maureen Sie, Bert van Der Brink, and Marc Slors (Hampshire, UK: Ashgate, 2004), pp. 129-141 (p. 131).

[6] Lynne Rudder Baker, *Persons and Bodies: A Constitution View* (Cambridge: Cambridge University Press, 2000), p. 19.

> If an organism exists at a certain moment, then it exists whenever and wherever – and only when and only where – the event that is its life at that moment is occurring; more exactly, if the activity of the *x*s at t_1 constitutes a life, and the activity of the *y*s at t_2 constitutes a life, and the organism that the *x*s compose at t_1 is the organism that the *y*s compose at t_2 if and only if the life constituted by the activity of the *x*s at t_1 is the life constituted by the activity of the *y*s at t_2.[7]

Then, to clarify what *Life* would mean for the continuation of a particular material being, van Inwagen says:

> Suppose that the activity of the *x*s constitutes a life at *t*; suppose that a few of the *x*s cease to be caught up in that life and that the remnant continue to be caught up in a life; suppose that those of the *x*s that have ceased to be caught up in that life are 'replaced' – that certain objects, the *y*s, come to be caught up in the life the remnant of the *x*s are caught up in, in such a way that the *y*s and the remnant of the *x*s constitute that life. Suppose that this sort of replacement happens a sufficient number of times that eventually none of the *x*s is caught up in the life that has evolved, by continuous (and 'insensible', as Locke calls it) replacement of the *x*s, from the life that was once constituted by the activity of the *x*s. Is this life the life that was constituted by the *x*s?
>
> In many cases, cases of the more usual sort, the answer is undoubtedly yes.[8]

Once again, this materialistic account of human persons allows for the gradual replacement of parts over time and for a person to persist through all of the changes that take place within her body. As such, I suspect that the thought experiment offered by Swinburne does not establish substance dualism and that materialists would reject his overall argument. And, despite my sympathies with his perspective, I am inclined to lean on other kinds of arguments that support the existence of the soul.

SWINBURNE AND MENTAL PROPERTIES

I also raise a second concern with Swinburne's argument for substance dualism that, while not a defeater for it, is at least epistemologically worrisome. As he makes clear throughout chapter 6, there are no physical

[7] Peter van Inwagen, *Material Beings* (Ithaca, NY: Cornell, 1990), p. 145.
[8] Ibid., p. 149.

or mental properties that are essential to a particular person's identity. That is, if Daphne at t_2 is the same person named Daphne that was once at t_1, then Daphne persists because she has the same 'thisness' as the earlier person named Daphne, and that 'thisness' is not tied to any particular mental property (psychological continuity) or physical property. He says, 'I begin my defence of this position by arguing that it is logically possible that some person P_2 at t_2 can be the same person as a person P_1 at t_1, even if he or she does not apparently remember anything done or experienced by P_1 at t_1 or earlier and has an entirely different character from P_1, and also has a largely different body (including brain) from P_1.'[9] He later adds:

> Hence, given that an earlier person who had all the same physical parts as me, and all the same physical and mental properties as me, could, it is metaphysically possible, not be me, and could, it is metaphysically possible, be me, it follows that the difference must consist in the presence or absence of some non-physical part. I must now have a non-physical part (i.e. a part which is a pure mental substance) which makes me me, which the earlier person (even if they were in all other respects the same) would not have had if they had not been me. We may call this non-physical part of me my 'soul'.[10]

So in short, Daphne's persistence over time does not require the preservation of any of her mental properties (memories, characteristics, dispositions, etc.) or physical properties (particular body, particular body parts, particular brain, size, shape, etc.) across time. She could lose all her mental properties and physical properties and still be the same person she was at an earlier time. What is more, Daphne's original physical and mental properties could appear later in some other person other than Daphne. As such, we could have the following situation. At t_1, Daphne – who has the 'thisness' of Daphne – has the mental properties of 1, 2, and 3 and the physical properties X, Y, and Z. But, at t_2, Daphne – who still has the 'thisness' of Daphne – loses all the mental properties (1, 2, and 3) and physical properties (X, Y, and Z) that she once possessed, only to have them replaced by mental properties 4, 5, and 6 and physical properties A, B, and C. Running alongside Daphne's life, at t_1, Velma – who has the 'thisness' of Velma – has mental properties 4, 5, and 6 and

[9] Swinburne, *Mind, Brain & Free Will*, p. 151.
[10] Ibid., p. 170.

physical properties A, B, and C. Then at t_2, Velma – who still has the 'thisness' of Velma – loses all the mental properties (4, 5, and 6) and physical properties (A, B, and C) that she once possessed only to have them replaced by mental properties 1, 2, and 3 and physical properties X, Y, and Z. So, just to clarify, we would have the following:

At t_1:

> Daphne has mental properties 1, 2, and 3 and physical properties X, Y, and Z.
>
> Velma has mental properties 4, 5, and 6 and physical properties A, B, and C.

At t_2:

> Daphne has mental properties 4, 5, and 6 and physical properties A, B, and C.
>
> Velma has mental properties 1, 2, and 3 and physical properties X, Y and Z.

Yet,

> Daphne at t_1 = Daphne at t_2.

And,

> Velma at t_1 = to Velma at t_2.

This is far more drastic than what would happen in Locke's Prince and the Cobbler thought experiment. For on this account, persons may switch bodies (and thus physical properties), but they maintain all their mental properties (memories, dispositions, desires, etc.). And so while the Prince may be terribly confused about how he got into the Cobbler's body (and vice versa), he is not confused about who he actually is. He remembers his life and possesses all the former mental properties that he once had. By contrast, Swinburne's approach makes it possible to lose both the physical and the mental properties. I can imagine that both Daphne and Velma at t_2 are horribly confused about who they really are. At t_2, Daphne is sure to think that she is Velma since she now has all of Velma's old mental and physical properties, and Velma is sure to think that she is Daphne for the same reason. Both have become lost but are not even aware that it has happened. They are deceived into thinking that they are someone that they are not.

My concern here is not with Swinburne's contention that we have a 'thisness' unique to each of us. My own Thomistic leanings incline me to think that we do have an immaterial soul, even if Swinburne and I might differ over the nature of that soul. My concern is with, on his model, our complete inability to identify ourselves across time given the contingency of our mental and physical properties. If Daphne has a complete loss and replacement of her mental and physical properties between t_1 and t_2, such that she now has all of Velma's old mental and physical properties, then even Daphne will not be able to properly identify herself in the future. She will think she is Velma when in fact she is not. The same is true for Velma who would now have all Daphne's old physical and mental properties and be equally confused and deceived. As I mentioned earlier, this is certainly not a logical defeater for Swinburne's position, but it does seem to be a troubling and unfortunate consequence of it.

CONCLUSION

On the whole, Swinburne's book is helpful and constructive. He has much to say about a number of important philosophical issues that need to be taken seriously. His critique of Thomas' view of the soul has caused me to question my own approach. For this I am grateful and I look forward to working through this difficult issue. And his thought experiment for gradual brain replacement is intriguing and enlightening. In the end, however, I find myself unconvinced by some key parts of his argument and concerned about the way his approach dismisses mental and physical properties of human beings.

SWINBURNE ON THE CONDITIONS FOR FREE WILL AND MORAL RESPONSIBILITY

DAVID P. HUNT

Whittier College

Richard Swinburne's rich and stimulating *Mind, Brain, and Free Will* synthesizes data and theories from diverse areas of philosophy into a single unified account of what agents are and how agency is exercised. The resulting account is currently unfashionable, but that says more about fashions than it does about the viability of Swinburne's account. Since I'm largely sympathetic to the picture Swinburne paints, I will offer friendly criticisms of some details of the landscape, focusing on those that most closely concern free will.

The first part of the book argues for positions in ontology, epistemology, and philosophy of mind that lay a foundation for Swinburne's account of agency. Many of these arguments are well worth engaging, but I will simply summarize here the conclusions that are most relevant to this review. (1) Substance causation, not event causation, is fundamental. 'The regularities constituting or underlying laws of nature are regularities, not of ... events, but ... in the causal powers and liabilities ... of actual substances.' (p. 7) Swinburne calls this the 'substances-powers-and-liabilities account', or 'SPL'. (2) We are essentially simple mental substances. Human beings 'consist of two parts – [the] soul (the essential part) and [the] body (a nonessential part), each of them separate substances' (p. 170). Swinburne argues here for 'the simple view' on which 'personal identity is a separate feature of the world' rather than 'analyzable in terms of degrees of continuity' (p. 150). '[E]ach person has a "thisness" which makes him or her that person, a "thisness" other than any thisness possessed by the matter of their brains', so that 'being that person is compatible with having any particular mental properties or any physical properties (and so body) at all' (p. 151). (3) It follows

from (1) and (2) that we exercise a species of causation reflective of the kind of substance we are. 'Such a substance could exercise its power to cause some effect because it intends – that is, tries – to cause that effect, and not because it has a propensity to cause the effect.' (p. 133) This is 'intentional', as opposed to 'inanimate', causation; the more common term is 'agent causation'.

This is of course a highly controversial mix, but the individual parts gain plausibility through their connection to the whole. Agent causation, for example, occupies a much better position against its critics if SPL is correct. I might add that while Swinburne does not situate this book within the Christian theism for which he's argued so powerfully elsewhere,[1] a fair amount of mutual support could be had if he did. Taking agent causation again as an example, the original causal powers belonging to God are presumably agent- rather than event-causal; a theist, then, can either *add* event causation to this original agent causation, or employ agent causation as a model for all causal explanation.[2] It's clear which of these is simpler.

Human agents in fact stand in intimate causal relation to bodies, most intimately to brains. This raises the question how the intentional causality exercised by minds is related to the inanimate causality exercised by the material substances with which they are conjoined, in particular, by those constituting their brains. It is only when the agent can make a difference to the outcome – when an exercise of agent causation brings about a different result (a different intention) than would have been produced by material causes alone – that the agent has free will and qualifies as morally responsible for his or her action.

What are the conditions under which this may happen? It turns out that they are fairly limited. Swinburne has some interesting things to say about how mind and brain might interact on these occasions, but I will take up the question on the psychological level. We are motivated to act by our desires and our value beliefs, 'beliefs about the objective intrinsic goodness or badness of doing actions of different kinds' (p. 175). ('Moral beliefs' are a species of value beliefs that concern 'a special kind of overriding goodness' and 'overlap substantially with the beliefs of most

[1] The word 'God' does not even appear in the book's index.

[2] For Swinburne, agent causation provides this model *because agents are substances*. A different but equally simple approach is offered by occasionalism, on which the only causation *is* agent causation.

other humans' – p. 176.) The problem is that '[b]eliefs are by their very nature involuntary' (p. 77), as are desires (p. 85). But the most common configurations of beliefs and desires make it 'inevitable' that we intend as we do. These are cases in which our intentions

(1) 'are caused by a strongest desire and we have no contrary moral belief',
(2) are caused 'by a strongest moral belief when we have no contrary desire', or
(3) 'simply execute (in the way which we believe to be the quickest way) some ultimate intention.'

In contrast, there are just two circumstances in which the outcome is *not* inevitable:

(A) where 'we have equally strong competing desires and moral beliefs', or
(B) 'where the desires and moral beliefs are in opposition to each other' (p. 201).

In A, where my beliefs and desires are tied in strength, 'I will have to make an arbitrary decision', though this 'will be fully rational, for whatever I do, I have a reason to do it, and no better reason for not doing it' (p. 184). It's only in B that the conditions exist for an exercise of full-on free choice. 'Here I have to decide whether to yield to desire and do the less good action, or to force myself – contrary to my strongest desire – to do the best action. ... This situation I will call the situation of *difficult moral decision*.' (p. 184)

It's worth comparing Swinburne's position with the very similar one endorsed by Peter van Inwagen.[3] Van Inwagen identifies three sets of circumstances in which we can't choose otherwise:

(i) when inclination is unopposed by duty;
(ii) when duty is unopposed by inclination; and
(iii) when it's obvious what to do.

These clearly parallel 1-3 above. Van Inwagen also identifies three circumstances in which we *can* choose otherwise, all involving conflicting alternatives where it isn't obvious what to do:

[3] Peter Van Inwagen, 'When Is the Will Free?' in *Agents, Causes, and Events: Essays on Indeterminism and Free Will*, ed. Timothy O'Connor (Oxford: Oxford University Press, 1995), pp. 219-238.

(a) Buridan's ass cases;
(b) when duty conflicts with inclination; and
(c) when conflicting values are incommensurable.

His (a) and (b) clearly parallel Swinburne's A and B, but Swinburne does not make explicit room for cases like (c). These might be shoehorned into A, on the grounds that incommensurable competing moral beliefs are, if not 'equally strong', at least *not measurably unequal in strength*. But because these cases often elicit a profound wrestling with one's deepest values, they don't really belong with cases that can be settled with a coin toss.

Van Inwagen's argument has elicited responses from a number of critics, including John Fischer & Mark Ravizza.[4] Critiques of Van Inwagen's restrictions on the exercise of free will are equally critiques of Swinburne's. I will make two comments on Swinburne's position, focusing just on the role of beliefs.

First, I'm not persuaded that beliefs are as immune to the will as Swinburne supposes. He argues that the involuntariness of belief (unlike that of thought and desire) 'is a logical matter, not a contingent feature of our psychology' (p. 77). The key step in the argument is this: 'But then [if we thought it was up to us whether or not to believe that *p*] we would know that we had no reason to believe that our belief that *p* was in any way sensitive to whether or not *p* was true; and in that case we couldn't really believe it.' (p. 77) But the fact that people actually do this is the best evidence that it is possible; it's then the job of a good theory to account for this possibility. It's hard to account for it on the assumption 'that a belief that some proposition has an (epistemic) probability on the believer's evidence greater than ½ ... is logically equivalent to a belief in that proposition ... ' (p. 76). There are people – perhaps even many people – who believe not just 'in' God (which may be a different use of 'belief'), but believe to be true the proposition *that God exists* and who would demur if asked whether they believe that the probability of this proposition is greater than ½ on their total evidence. William James famously argued that there are conditions under which such belief is not only possible but epistemically permissible. James' will to believe and Kierkegaard's leap of faith are especially likely to take as their objects

[4] John Martin Fischer & Mark Ravizza, 'When the Will Is Free', in O'Connor, op. cit., pp. 239-269.

value beliefs, the beliefs whose involuntariness is supposed by Swinburne to play such an important role in limiting our free will.

Second, the effect of beliefs on our intentions, like our having the beliefs in the first place, may be more subject to the will than Swinburne allows. Of this effect he writes:

> I could not believe that some action was really morally good to do ... and yet not see myself as having a reason for doing it. And I could not see myself as having a reason for doing it unless I had some inclination to do it. And the better I believe some good action to be, the greater as such is my inclination to do it. (p. 178)

But what seems to have fascinated Augustine about his theft of pears is that it didn't fit this paradigm. Perhaps Swinburne means to block this response with his very broad notion of a value belief, so that when Augustine (or Raskolnikoff) acts contrary to a moral belief that he has precisely because it is a moral belief, he does so in the service of a non-moral value belief that he has. The question is whether there is good independent evidence that such a non-moral value belief is present, or whether its presence is posited simply as a requirement of the theory.

Leaving to one side the frequency of free agency among human beings, let us turn to Swinburne's thoughts on an important challenge to free will, as he understands it. This is Harry Frankfurt's famous critique of the Principle of Alternate Possibilities, or 'PAP', which Swinburne formulates as follows:

> A does *x* freely only if he could have not done *x* (i.e. could have refrained from doing *x*). (p. 203)

(PAP, as it has been discussed in the literature, is actually a principle about moral responsibility; what Swinburne calls 'PAP' is really a freedom version of PAP.) Frankfurt's counterexample involved a case in which Jones decides to kill Smith and both of the following are true: (i) if Black fails to detect a prior sign that Jones will decide to kill Smith, Black intervenes to *cause* Jones to decide to kill Smith; (ii) Black does detect the sign and does not intervene, so Jones decides on his own. Frankfurt believed that in this scenario Jones is morally responsible for his decision while having no accessible alternative to the decision, so PAP is false.[5]

[5] Harry Frankfurt, 'Alternate Possibilities and Moral Responsibility', *Journal of Philosophy*, 66 (December 1969), 829-39.

Swinburne endorses a response to Frankfurt, called the 'dilemma defence', that is associated with David Widerker.[6] Either the prior sign causally determines Jones's decision, or it doesn't. If the former, Frankfurt's scenario begs the question against the libertarian; if the latter, it remains possible for Jones to refrain from deciding to kill Smith – not ultimately, of course (Black will ensure that Jones makes the decision he wishes him to make), but *at t*, the time at which Jones actually decides to kill Smith. All that's needed to save PAP from Frankfurt counterexamples, then, is to clarify it with the help of some temporal indexing:

A does *x* freely at t only if he could have done not-*x* at t instead.

This principle, which Swinburne calls 'PAP*', is 'surely true' (p. 204).

I think that Swinburne's confidence in PAP* is misplaced. Alternatives can always be found in Frankfurt cases; the question is whether they are sufficiently 'robust' to ground all the responsibility ascriptions we wish to make. If Jones can refrain *at t* from deciding to kill Smith, this might explain how he can be morally responsible for deciding *at t* to kill Smith, but not how he can be morally responsible for deciding to kill Smith *full stop* (as he surely is), since he has no alternative to this.[7] But the dilemma defence has also given rise to new counterexamples which violate PAP* as well as PAP. Some feature 'buffered' scenarios in which the agent must complete an intermediate step (traverse a psychological buffer, so to speak) before he is in a position to decide otherwise.[8] Widerker himself, ironically, no longer supports the dilemma defence, and has developed his own counterexample to PAP, which he calls 'Brain-Malfunction-W'.[9]

But there is a vast literature on Frankfurt counterexamples to PAP, and it's unreasonable to expect Swinburne's brief remarks to do it justice.

---

[6] David Widerker, 'Libertarianism and Frankfurt's Attack on the Principle of Alternative Possibilities', *Philosophical Review*, 104 (1995), 247–61. A similar move can also be found in Robert Kane, *The Significance of Free Will* (New York: Oxford University Press, 1996), and in Carl Ginet, 'In Defense of the Principle of Alternative Possibilities: Why I Don't Find Frankfurt's Arguments Convincing', *Philosophical Perspectives*, 10 (1996), 403–17.

[7] See David Hunt and Seth Shabo, 'Frankfurt Cases and the (In)significance of Timing', *Philosophical Studies*, 164 (March 2012), 1-24.

[8] David Hunt, 'Moral Responsibility and Buffered Alternatives', *Midwest Studies in Philosophy*, 29 (2005), 126–145.

[9] David Widerker, 'Frankfurt-Friendly Libertarianism', in Robert Kane, ed., *The Oxford Handbook of Free Will*, 2nd edition (Oxford: Oxford University Press, 2011), pp. 266-287.

The point I would like to make is not that Swinburne fails to settle the debate in PAP's favour, but that he doesn't need to do so. If the alternatives requirement for free will is abandoned, there is still the sourcehood requirement, for which (like the alternatives requirement) there are both compatibilist and incompatibilist interpretations.[10] For an agent-causal libertarian like Swinburne, it's the sourcehood requirement that is fundamental anyway, and it's not surprising that Swinburne's own definition of 'free will' – 'the agent acts intentionally without their intentions being fully determined by prior causes' (p. 202) – is a pure statement of incompatibilist sourcehood, unsullied by any reference to alternatives. If PAP is false, alternative possibilities (like the red spots signifying measles) are a symptom that ordinarily accompanies free will, though they are metaphysically distinct from it, as shown in the extraordinary cases constituting Frankfurt counterexamples. The underlying condition, of which alternatives are normally symptomatic, is simply a particular kind of causation, the 'intentional causation' exercised by agents, when this is genuinely effective (i.e., makes a difference not explained by inanimate causation alone). This might be a fairly frequent occurrence. Swinburne's limited conditions for free will rest on two premises: that most intendings are inevitable, given the agent's beliefs and desires; and that inevitable intendings – intendings for which the agent has no accessible alternatives – are unfree. I have already suggested that the first premise is too strict, but if PAP is rejected, this argument for limited free will can be resisted at the second premise as well. I think this is a result that Swinburne should welcome.

Free will is important because it's a requirement for moral responsibility, the subject of the book's last chapter. What is the scope of our moral responsibility? To answer this question, Swinburne draws on moral intuition (of course), but also on theories developed earlier in the book. An example is his judgment that the mere passage of time, no matter how much the agent has changed, does not diminish responsibility (p. 226). This is said to follow from his 'simple' account of personal identity, according to which a mental substance has the same essential properties and 'thisness' throughout life; changes in contingent properties, such as memory and character, would not then affect the individual's responsibility. This is not an issue on which ordinary judgment speaks with a single voice, and Swinburne explicitly contrasts

[10] Widerker now characterizes himself as a 'source incompatibilist'.

his position with that of Locke. In the face of someone who protests, 'But I'm not the same person I was 20 years ago!' Swinburne is in effect responding, 'There is (perhaps) a sense in which this is true, and a sense in which it is false; unfortunately, the sense in which it is false is the sense relevant to moral responsibility.' Insofar as one's moral intuitions line up with Swinburne's rather than with Locke's, this may provide some retroactive support for the account of personal identity from which this result is supposed to follow.

I propose to review four further areas in which Swinburne's conclusions about moral responsibility are at least controversial. (The areas are interconnected, so distinguishing among them is somewhat artificial.) The first of these concerns the conditions for praiseworthiness (there are of course companion conditions for blameworthiness). Swinburne's initial claim is that, *normally*, an agent is to be praised only for actions believed by the agent to be morally good but not morally obligatory. The point of the qualifier is soon evident, because praise is also clearly relative to another standard reflecting our expectations of other people and how difficult it was for them to perform the action under assessment (p. 212). It turns out, then, that an agent may be praised for an action that is morally obligatory if it is sufficiently difficult for the agent to choose (given opposing desires), and not praised for an action that is morally good but not obligatory if it doesn't require any extra-normal effort.

I have a couple of worries about this account. One is that the account is complex, combining the initial claim (that an agent is to be praised only for actions believed by the agent to be morally good but not morally obligatory) with a set of exceptions, where the exceptions could just as well have constituted the norm and the norm the exceptions. An alternative account is simply that a person is praiseworthy for doing what (they believe) is morally good – whether or not they believe it to be morally obligatory – when they could have acted otherwise but didn't. In this context, a third party who exclaimed 'Good for you!' would not be saying anything false; but since praiseworthiness comes in degrees, it is often not worth pointing out, and the 'conversational implicature' of doing so would be misleading. This account is simpler and more unified, and for that reason seems to me to be better.

The other worry is that degree of praiseworthiness does not always track degree of difficulty, and sometimes it even seems inversely proportionate to it. I'm inclined to think that a woman who rushes into a burning building to save her child, without stopping to think about it, is

more praiseworthy than a woman who does so only after wrestling with a 'difficult moral decision'. Swinburne later discusses the case of a 'hero who is caused inevitably ... to do a supererogatory action', and argues that we should not praise (or praise so highly) the hero's action; rather we should admire the character from which the action flows, and praise any earlier actions that led to the development of that character (p. 220). This seems to me overly restrictive of what the hero can be praised for, but it also makes it hard to understand how God can be praiseworthy for anything he does. Perhaps divine praiseworthiness rests on a wholly different analysis than human praiseworthiness, or 'praising' God is just a loose manner of speaking (whose real content is admiration for God's character). But a theory on which God is genuinely praiseworthy, in a sense that is continuous with the sense in which humans are praiseworthy, is surely preferable.

The other controversial issues I would like to mention concern what we are culpable *for*. Swinburne's earlier argument that free will is restricted to a fairly rare set of circumstances leads to similar restrictions on moral responsibility. Here are three such restrictions.

We are culpable only for what we do freely at the time we do it. Most of those who restrict free will do not similarly restrict moral responsibility. They are able to do this because they endorse so-called 'tracing' principles, under which a person can be held responsible for an action they didn't freely choose if the action follows from an earlier action which they did choose. (There are different accounts of how the one action must 'follow from' the earlier action; an epistemic condition will surely be part of the mix.) But Swinburne rejects tracing in favour of the strict view that a person whose own choices led to their being unable to fulfil an obligation 'may be culpable – not for the failure to fulfil the obligation at the later time, but for allowing themselves to get into a situation where they believed that it would be improbable that they would be able to fulfil the obligation' (p. 213). But most people's judgment is that when a drunk driver kills a pedestrian, the driver is culpable not only for earlier actions that could have been avoided (such as drinking to excess), but also for actions that the driver was then too impaired to avoid (such as killing the pedestrian). This requires tracing.[11]

[11] Manuel Vargas, for example, writes that 'one of the nice features about tracing is that it is one of a few things to which nearly all parties in the debate about free will appeal to with equal enthusiasm'. 'The Trouble with Tracing', *Midwest Studies in Philosophy*, 29.1 (2005), 270.

We are culpable only for trying, not for the success of our efforts. An excellent justification for this position is that, once we make our contribution to events by trying, the matter is then out of our hands: it's up to the world, not to us, whether our efforts succeed. How can we be blamed for *that*? But when Swinburne concludes that '[s]omeone is just as culpable for trying to blow up a civilian aircraft although prevented from doing so by the police discovering the bomb, as they are for succeeding in blowing up the aircraft' (p. 211), I find my intuitions putting up some resistance. The successful bomber certainly seems to have more on his conscience, more *for* which he is culpable. The same is true of the drunk who kills someone in comparison with the drunk who manages to get home without hitting anyone, each of whom (as we've just seen) is no more nor less culpable than the other by Swinburne's lights. In the case of the successful bomber as in the case of the impaired driver, a tracing principle might produce results more in keeping with ordinary judgment. The two cases also raise the question of 'moral luck', and what a good theory should do with this problem. The failed bomber is the beneficiary of moral luck: through no credit to himself, he has been spared responsibility for multiple deaths. The question is whether such luck is a real feature of the moral landscape, a tragic concomitant of the human condition, or an illusion to be dispelled by the right moral theory. Swinburne's restriction on culpability to cases of trying clearly belongs in the latter camp; whether that's a virtue or a vice is a question I'll leave open.[12]

We are culpable only for acting contrary to our value beliefs. As we saw earlier, Swinburne regards our beliefs at any particular time as givens; for this reason, when we act in accordance with our value beliefs (beliefs whose very nature is to motivate action), we cannot be blameworthy for so acting. We can be culpable then only when acting contrary to our value beliefs. Swinburne offers the following example of culpability being limited by the agent's value beliefs: 'Some people believe that stealing from the rich is not wrong; and so if I have this belief and also the belief that you are rich, I would not be culpable for stealing from you.' (p. 211) Whether or not this is the right result in this particular case, the principle seems too strong. It might, for example, justify the conclusion that Hitler's culpability was much more restricted than anyone would have thought.

[12] The term 'moral luck' was introduced by Bernard Williams in his 'Moral Luck', *Proceedings of the Aristotelian Society*, supp. Vol. 50 (1976), 115-35.

If we learned that Hitler's choices conformed very closely to his value beliefs, this would not (and should not) lead us to a significantly different assessment of his culpability. Perhaps he was obligated to form better beliefs; perhaps he was obligated to act contrary to his value beliefs. But in one way or another, Hitler's culpability needs to be tied less tightly to his value beliefs. Another way to arrive at the same moral is to imagine a parent who takes to heart Kant's dictum that the only thing that is good without qualification is a good will; persuaded that culpability is relative to one's value beliefs, the parent sets about instilling in the child those value beliefs to which it is easiest, and thus most likely, that the child's choices will conform. Culpability-avoidance is not the only quality that parents should try to foster in their children, but something has surely gone wrong if this quality is easier to achieve the more lax the child's value beliefs.

In conclusion, it should be evident that I find human choice and the moral responsibility that comes with it more mysterious than Swinburne makes it out to be. Perhaps Swinburne does as well. Peter van Inwagen, at a key point in his *Essay on Free Will*, summed up the situation as follows: 'I must choose between the puzzling and the inconceivable. I choose the puzzling.'[13] Free will *is* puzzling; to cite the title of a later article by Van Inwagen, it 'remains a mystery'.[14] But philosophers aren't content with this situation; it's our job to work on the puzzle and make the mystery somewhat less opaque. This requires, by the very nature of the enterprise, abstracting from reality what is amenable to philosophical analysis. This Swinburne does with great skill, and the resulting book is an impressive example of how philosophical order can be imposed on the messy phenomenon of human free will.

[13] Peter van Inwagen, *An Essay on Free Will* (Oxford: Clarendon Press, 1983), p. 150.

[14] Peter van Inwagen, 'Free Will Remains a Mystery', in Robert Kane (ed.), *The Oxford Handbook of Free Will*, 1st edition (Oxford: Oxford University Press, 2002), pp. 158-77.

RESPONSE TO REVIEWERS

RICHARD SWINBURNE

University of Oxford

Many thanks to the Editors of *The European Journal* for organising this symposium, and to all the reviewers for the work involved in grappling with my arguments.

I. INFORMATIVE DESIGNATORS AND PROPERTY DUALISM

I claimed that the history (in an objective sense) of the world is just the succession of events (which – briefly – are the instantiations of properties in substances at times), and that we can tell it in many different ways by cutting the world up into substances and properties in different ways; but that to do so we need to pick out substances and properties by informative [rigid] designators. I defined an 'informative designator' as follows:

> For a rigid designator of a thing to be an informative designator it must be the case that anyone who knows what the word means (that is, has the linguistic knowledge of how to use it) knows a certain set of conditions necessary and sufficient (in any possible world) for a thing to be that thing (whether or not he can state those conditions in words.) Two informative designators are logically equivalent if and only if they are associated with logically equivalent sets of necessary and sufficient conditions. To 'know' these conditions for the application of a designator – as I shall understand this expression – just is to be able (when favourably positioned, with faculties in working order, and not subject to illusion) to recognize where the informative designator (or, if it is defined in words, the words by which it is defined) applies and where it does not and to be able to make simple inferences to and from its application. (*Mind, Brain, and Free Will*, p. 12)

I illustrated the difference between an informative designator such as 'red' and an uninformative one such as 'water' (as understood in the eighteenth century). We know what it is for a thing to be red because we can always recognise red things under the stated conditions; but in the eighteenth century when to be 'water' meant to have the same essence as the stuff in our rivers and seas, we could only recognize water when it was in our rivers and seas. With aid of this distinction I argue that we would leave out an important part of the history of the world unless we included in that history both pure mental events (sensations, beliefs, etc.) to which subjects have privileged access and physical events which are publicly accessible events (or events which entail the occurrence of events of both kinds), and both pure mental substances and physical substances (or substances which entail the occurrence of substances of both kinds).

I have quoted my definition of 'informative designator' at length, because it seems to me that neither Lynne Rudder Baker nor William Jaworski have understood it fully or grasped its significance. Baker's first objection (p. 8) to the utility of this concept is that since on my view that 'I' is an informative designator which designates a mental substance, everyone who uses 'I' competently should believe that they are a mental substance, and of course they don't. But Baker has not appreciated the sentence of my definition which states that to 'know' the conditions for the application of the designator '*as I shall understand this expression*' just is to be able to recognize when it applies and when it does not; knowing the necessary and sufficient conditions for a thing to be that thing is 'knowing the nature of what I am talking about'. I use the point about 'I', together with a claim about the logical possibility of certain thought experiments, to argue over several pages for the truth of the philosophical thesis that I am a mental substance. To do this I use arguments which, even though in my view they are sound, not everyone will believe to be sound, and so not everyone will believe the philosophical thesis. Yet clearly someone can know something without believing all its logical consequences. Jaworski has a similar general worry – he writes (p. 24) that 'even if I am necessarily right about what "I" refers to, I can still be wrong about what essential properties I have'. But on my definition of knowing *a* set of conditions necessary and sufficient for the application of a designator, 'being necessarily right' about what it refers to *is* knowing one set of essential properties for being me. There are many different logically equivalent sets of necessary and sufficient conditions

for anything to be what it is; I only need to know one particular set in order to use the word 'I' correctly. But knowing one set of conditions for applying 'I' is quite sufficient to rule out the possession of any physical properties as metaphysically necessary – in virtue of my arguments to show that physical and pure mental properties and so the existence of physical and pure mental substances never entail each other. Of course there may be physical substances with physical properties which are *causally* necessary and sufficient for my existence – but that is irrelevant to my argument.

Then Baker claims (p. 8) that she knows how to use the word 'arthritis' but doesn't know necessary and sufficient conditions for its application. But the meaning of a technical term is determined by a 'group of experts' (my p. 10), and on the assumption that the experts do know the necessary and sufficient conditions it will be an informative designator; otherwise it is an 'uninformative designator'. And then Baker claims (p. 8) that she does not know the necessary and sufficient conditions for something to be 'red' because of doubt about whether reddish-orange things and such like are red. But a competent language speaker will know that a reddish-orange thing is a borderline case of being red and so (given the falsity of the epistemic theory of vagueness) will know that the necessary and sufficient conditions for being red are not satisfied, and will also know that the necessary and sufficient conditions for being not-red are not satisfied. Baker is mistaken in supposing (p. 8) that on my view whether a word is an 'informative designator' normally varies with the speaker. My view is that words have a 'correct use' (my p. 17) in a language; and normally whom or what is designated by a word is the same whoever uses it. The correct use today of 'water' by everyone is to designate H_2O. However what is designated by an indexical clearly varies with the speaker – what 'you' or 'here' designate varies with who use these words, when, and where. So what 'I' designates, while being always an informative designator, varies with the speaker. And, I argued, the correct use of 'I' or of my own name by myself is that of an informative designator; but that the correct use of my name by someone else is that of an uninformative designator. This is the sole case where whether a word is an 'informative designator' varies with the speaker. 'Richard Swinburne' as used by me is an informative designator, whereas as used by others it 'is an uninformative designator' (my p. 165). This is because 'Richard Swinburne' is used to refer to the actual person who has a certain body. But others do not know the necessary and sufficient conditions for being

that person – since that person could now exist without a body or in a different body; and others would not necessarily be able to recognize that person under those conditions. They apply the word, as people applied 'water' in the eighteenth century, on the basis of conditions which are not necessary and sufficient for being that person. I also use 'Richard Swinburne' to refer to the actual person who has a certain body. But I, being aware of myself, am in a position to know who that person is, and so to know those necessary and sufficient conditions for being Richard Swinburne. I alone necessarily would be able to recognize myself now if I had a different body, or if I did not have a body at all. I conclude that none of Baker's criticisms of the utility of the concept of an informative designator have any force.

Jaworski has a similar package of objections to the way I apply the concept of informative designator. He claims (pp. 21-22), that showing that 'red' and 'reflects light of such-and-such wavelengths' are not logically equivalent and so pick out different properties, does not show that one is mental and the other is physical. He is correct in this claim. 'Red' is in my sense a physical property, because the substance in which it is instantiated – a surface – does not have privileged access to its instantiation in it; all observers (when favourably positioned, with faculties in working order, and not subject to illusion) have equal access to whether it is instantiated; and of course 'reflects light of such-and-such wavelengths' is also a physical property. I introduced this example in chapter 1 to illustrate how we need to distinguish between properties – by their informative designators not being logically equivalent; and I deliberately chose an example which was not an example where one property was mental and the other property was physical. But having introduced the concept of an informative designator, I proceeded to argue that there are innumerable properties individuated by informative designators which are such that the substance in whom they are instantiated has privileged access to their instantiation, and which are therefore not the same properties as ones which are such that the substance in which they are instantiated does not have privileged access to their instantiation. Contrary to Jaworski, I do not merely 'tacitly endorse' (p. 22) 'modes of presentation'; I explicitly claim that they are themselves properties – 'a mode of presentation ... is just as much a real characteristic of any object as any property' (my p. 26). My objection was to introducing into our ontology a category of 'mode of presentation' separate from the category of property. Certainly some properties might be distinguished from

each other by different modes of presentation (that is, properties distinct from the properties which they were picking out), but to suppose that all properties could be so picked out would lead to an infinite regress; some properties (those having informative designators) must be recognizable as what they are without intermediaries. However all substances, unlike all properties, may be picked out by different modes of presentation (i.e. different properties); and so I am not denying what Jaworski has to say (p. 22) about 'Cicero' and 'Tully' with the caveat that (except for their use by Cicero himself) they serve as uninformative designators.

Jaworski claims (p. 20) – correctly – that we may find ourselves in unusual conditions and call something 'red' which we would not call 'red' under ideal conditions and come later to recognize our mistake. Then he goes on to claim that if 'red' is like this, then it is possible that redness may be identical to a surface property whose essence we do not know even though we can competently use 'red' to refer to objects having it. But I cannot see why that is supposed to follow. To repeat – words mean what competent language users mean by them. And in my view 'red' is used to designate a certain aspect of the way things look (to most people); and the way things look is the sole determinant of whether they are red. If (implausibly) I am mistaken and 'red' is not used in that way, then we would need another word which is used in this way to designate that aspect (or words which entail that aspect) if we are to give a full description of the world. Jaworski suggests that while all speakers understand 'red' in the same sense, some competent speakers might pick out an object as 'red' by the way it looks, and other speakers might pick it out by the wavelengths of the light it reflects. 'Red' would then have two criteria for its application, each sufficient and only disjunctively necessary; red objects could be picked out by either of two separate properties. But again if that's how 'red' works then – to fully describe the world – we need an informative designator for each of the two properties. My definition that two rigidly designating predicates which are not logically equivalent designate different properties is not (p. 22) a 'tendentious assumption about properties'; it is a definition which I explicitly provide since there are alternative criteria which one could choose to use for the identity of properties. I chose this criterion because with the aid of it we can tell the whole history of the world by listing only a set of events (instantiations of properties in substances at times) which entail all events which ever occur. I am not denying, as Jaworski seems to suppose (pp. 23-24), that properties and substances

may have essential properties of which we are unaware; but I do claim that – if so – those essential properties must be entailed by the ones of which we are aware when we pick them out by informative designators. To deny that is to use different criteria for 'same property' from the one which I have good reason to propose.

II. COMPOSITE SUBSTANCES

Baker criticises my principle of the identity of composites. One way to express the principle, she claims (pp. 9-10), is 'there is no more to any substance than its parts (e.g. fundamental particles) and the way those parts are arranged'. But – crucially my principle, as I stated it (my p. 35) is 'there cannot (logically) be two things which have all the same parts having all the same properties, arranged in the same way'. 'Having all the same properties', is crucial (and indeed makes 'arranged in the same way' redundant). Among the properties of a thing are (my p. 5) its relations to other things including (see e.g. my p. 33) properties of spatio-temporal continuity with earlier things. Hence Baker's drinkalator – carburettor example (pp. 9-10) does not count against my principle. All the parts of a drinkalator have the property of being part of a device used for making soft drinks; and the parts of a carburettor do not have this property. All that is ruled out by my principle is that the world could be different if instead of one such machine there was another such machine made of the same parts with the same properties (including the same past-related properties) arranged in the same way. Then I claimed that we can tell the story of the ship of Theseus either by using the 'same planks' criterion or by using the 'gradual replacement' criterion for the identity of the ship 'without anything being omitted'. Baker objects (p. 9) that which later ship is the original ship makes a great difference to the history of the world, because it affects who owns which ship. But on both accounts it may be part of the story that the courts determined which of these ships was the ship of Theseus. But (barring the – to my mind – implausible epistemic theory of vagueness) the courts did not discover the answer to a deep metaphysical issue; they merely determined which criterion should be used for determining who owned which of the two subsequent ships. Once we know what happened to all the planks and what the courts decided, we can still tell the story in two mutually entailing ways. I doubt if the decision of a court (e.g, that 'same planks' determines same ship)

is enough to settle how the expression 'same ship' should normally be used; but even if it does settle it, we can still tell the whole story by using a different criterion (e.g., gradual replacement) from that used by the courts, adding that the courts called the ship formed of the same planks the original ship and so determined that the owner of the original ship now owned, not the original ship but the ship formed of the same planks.

III. MENTAL SUBSTANCES

A mental substance is – on my definition – one for the possession of which some mental property is essential. I argue that we humans are mental substances because we co-experience at a time and over time different mental properties. Baker (pp. 10-11) outlines correctly the structure of my argument from co-experience at one time (e.g. of sensations of different kinds), which I call the argument from the synchronic unity of the human person. She doubts its validity because she suspects that I am using 'delimit' as a causal notion in (2), and that (2) will only support the ontological claim of (4) if it uses 'delimit' in a stronger sense. 'Delimit' is indeed being used in a causal sense in (2), but my claim is that the causal facts determine the ontological facts. What makes a brain my brain (ontological fact) is that my mental events are caused by, or cause, events in that brain (causal fact). It then follows that if I have any physical properties (as Baker and almost all of us hold), then necessarily a mental property of mine determines that certain of my brain properties are (contingently, because of the causal connection) among those physical properties. I claimed at this place in the discussion as elsewhere in the book, that we can cut the world up into substances and trace their continuities in different ways without leaving anything out of the history of the world. In particular we could regard the physical part of me as only a brain – after all, my brain in a vat is still me, or even as only part of a brain, in interaction with another substance, constituted by the rest of my body; and the history of the world which assumed this would entail and be entailed by its history described with our more normal categories. My point at this stage of the discussion, was that if we treat me as having any physical properties at all, certain brain properties must be among them in virtue of their causal relations to a mental property. A mental property determines the minimal set of physical properties which, if I have any physical properties, are essential if a substance is to be me at all. And so, whether or not any physical properties are necessary for

my existence, a mental property is necessary. Jaworski (p. 25) is right to claim that I have to hold that 'I have no physical properties other than those which are determined by my mental properties'. But – and I should have made this point explicit in the text – all my other physical properties (on our normal understanding of my physical boundaries) cause or are caused indirectly by (i.e. via a causal route) my brain properties which cause or are caused 'immediately' (my p. 143) by my mental properties – given that the causation is close enough to regard them as belonging to a single organism. What makes the other physical properties mine is determined (immediately or indirectly) by my mental properties. The rest of my brain is mine because events in it cause the events which are the immediate causes of my sensations, and so on. Even my hair is mine, because it goes where my brain events cause the rest of my body, to which it is attached, to go. So if I had no mental properties, neither would I have any physical properties. If I no longer had any conscious events or continuing mental states (e.g. beliefs and desires which I can have while not being conscious, and even when in a coma), I would not exist – on our normal understanding of what it is for 'me' to exist, which I was trying to analyse. So I reject Jaworski's suggestion (p. 25) that 'the continued existence of some ... mentally irrelevant parts [might] be sufficient for my existence'. My corpse is not an existing me.

Although my argument from the diachronic unity of the human person is supported by my argument from its synchronic unity, it does not in fact need it – contrary to Baker's claim. In fact it consists of several connected arguments, none of which need the argument from synchronic unity. I opened the section on diachronic unity with the claim that (among pure mental events, to which – see my pp. 71-72 – my discussion thereafter of 'mental events' was almost entirely confined) all conscious events, such as having a pain or a thought last for a period of time. (As Baker points out, I wrote carelessly that 'all events take time' (my p. 148); I should have written 'all conscious events take time'.) I then argue that for a person to have a conscious event such as a pain lasting for a period (e.g. one second) is for that person to have a pain lasting for the first half of that period and also a pain lasting for the second half of that period. In being aware of one's one-second pain, one is aware of oneself as experiencing pain for the first half of that period and then experiencing pain for the second half of the period, and so of oneself continuing over time. Yet (given – as I argue later -that 'I' is an informative designator) that I have a pain (or any other conscious event) does not entail and is

not entailed by any physical event at all. So the substance involved in that event (myself) could – it is metaphysically possible – exist even if there were no physical events. That entails that I am a pure mental substance. What applies to me applies to other humans, and since the argument is simply an argument from a being being conscious to it having a soul as its one essential part, it applies to any conscious being. But all that that argument shows is that I have a soul as my one essential part for the 'specious present' of one conscious event. I then extend the argument to show that I have a soul as my one essential part for as long as I have a series of overlapping conscious events. I then appeal to various thought experiments to show the logical possibility and then the metaphysical possibility of any conscious being continuing to exist over intervals of not being conscious without that consciousness being dependent on any particular brain and then on any brain at all. Hence being conscious entails being such as not – metaphysically – to need any physical properties in order to exist. Since consciousness entails being a pure mental substance, not being a pure mental substance entails not being conscious. James Dew (p. 31) calls it an 'unstated assumption' that 'it is impossible for purely physical organisms to have conscious experiences'. I think that the argument for that is sufficiently explicit in the arguments from synchronic and diachronic unity summarized above to be found in my chapter 6.

Dew misunderstands my thought experiment in which a tenth of my brain is replaced each year until after ten years none of the original brain remains. He asks (p. 32) 'why couldn't it ... happen' that the same 'stream of consciousness – a first person perspective – is maintained by an ever-changing physical organism'? If he means by the 'same stream', a ten-year stream of overlapping conscious experiences, that seems unlikely in view of the human need for sleep, some of which seems to be non-conscious. But if it did happen, the resulting person would indeed be the same person as the original person. But given a few intervals of non-consciousness in such an experiment, as I wrote, it still seems logically possible that the resulting person would be the same as the original person. But, I go on to say, it seems 'also logically possible' that the resulting person would not be the original person. I then go on to argue that what is logically possible is in this case also metaphysically possible. I then conclude that because each scenario is compatible with all the data about the physical parts, and the physical and mental properties of the original and resulting persons, by the principle of the identity of composites the final resulting

person would need to have a same non-physical part (the same soul) if he is to be the same person as the original person.

Baker summarises her own view – a view of which Dew writes approvingly – that (to quote Dew, p. 33) as well as some sort of bodily continuity, 'all that is necessary for the persistence of the persons is continuity of first-person perspective' which is 'the activity one has to think of herself as herself'. The continuity of 'first-person perspective' is explicated as the continuity, not of a conscious thought, but of an 'ability to think of herself as herself'. But what is it for such an ability possessed by a person P_2 on waking up in the morning to be continuous with the ability possessed by the person P_1 who went to sleep the previous evening, and whose body was continuous with the body of P_2? If bodily continuity and/or continuity of memory and character is necessary and sufficient for the resulting ability to be continuous, we run into all the problems raised for the complex theory of personal identity by brain transplant thought experiments. For example, 'continuity of first person perspective' with the previous person could be possessed by more than one subsequent person; and so it cannot provide an answer to which (if either) subsequent person is the original person, and – unless personal identity is a matter of degree (a view which I gave arguments to reject on pp. 154-5), there must be an answer to this question. And the answer which I provide which is immune to such problems, is that 'continuity of first-person perspective' must be construed as the continuing existence of an indivisible part of the original person, her soul; and then, I now add against Baker, no one needs to 'think of herself as herself' in order to have some primitive conscious events, e.g. sensations, and so to continue to exist. Peter van Inwagen's theory, to which Dew is also sympathetic, that personal identity over time is constituted by the continuing life of a physical organism, also runs into all the problems of any complex theory of personal identity. To take a slightly different example, it runs into the problem that any answer to the inevitable question of how many bodily parts can be replaced how gradually for the person still to be the same person will seem highly arbitrary. The obvious non-arbitrary solution to this question is that the truth about when a person continues to exist is a truth over and above any truth about how many bodily parts have been replaced how gradually, but that the fewer parts are replaced and the more gradual the replacement, the more probable it is that the same person continues to exist.

Both Dew and Baker are troubled by consequences of my theory. Dew (p. 33) claims that my theory has the 'troubling and unfortunate consequence' that we may be completely unable 'to identify ourselves across time'. But I argued that under normal circumstances (i.e. when brains are not split, and memory and character are continuous) it is 'enormously probable' that I am the same person as any previous person who had the same brain. To ask for more than that is unreasonable. It is only under extremely abnormal circumstances that there will be a serious doubt about which previous person was me. And, as Dew admits, that there will be a serious doubt under such circumstances is no argument against the theory. Baker (p. 13) regards it as a disadvantage of my theory that 'there is not (or rather I cannot think of) any naturalistic way that an immaterial mind could have come into existence'. I too cannot think of such a way, but I – unlike Baker – am very happy to endorse this consequence of my theory – given what I regard as strong arguments in favour of that theory.

IV. FREE WILL AND MORAL RESPONSIBILITY

David Hunt has confined his comments to my views on what it is to have 'free will' of the kind that makes us morally responsible. He has read my text very carefully – more carefully than I read some of the texts dealing with PAP, 'The Principle of Alternate Possibilities', when – as Hunt points out – I represented (on my pp. 203-4) various authors as making a claim about free will (in my sense), whereas in fact their claim concerns moral responsibility. Fortunately, as Hunt kindly remarks (p. 45), this error (for which I apologize to my readers) makes no difference to my main arguments in chapter 8 that free will in my sense – an agent acting 'intentionally without their intentions being fully determined by prior causes' (my p. 202) – is necessary for moral responsibility.

Generally Hunt is very sympathetic to my views on these topics, and so I am left to deal with some fairly minor issues. I stand by my argument that belief (that is, belief that some proposition is true) is immune to the will. My claim is that we cannot change our beliefs at will, that is immediately by a decision. However I wrote (p. 77) that 'I can try to brainwash myself, so as to come to hold later a certain belief specified in advance; but I will only succeed if I get myself to be caused to hold the belief at the later time by some cause, e.g some brain event, which

I am not at that time intentionally causing'. Pascal had a suggestion of a procedure for doing this – 'taking holy water, having masses said, and so on. That will make you believe quite naturally'.[1] This point should be enough to deal with Hunt's claim that people actually do voluntarily change their beliefs; to do so requires a period of brainwashing and is not guaranteed success. Even if some, maybe most, people do not accept my philosophical thesis that belief that p is belief that p is more probable than not -p, that doesn't show that the thesis is false! It may show only that some people do not fully understand the logical consequences of 'S believes that p'. Hunt asks for independent evidence that when we do an action intentionally, we always believe that it is in some way a good action to do. He claims (p. 43) that Augustine's action of stealing pears when young, 'didn't fit this paradigm'. It seems to me that Augustine's description of his motivation in stealing the pears does exactly fit the paradigm, even though Augustine himself seems to doubt it. For he writes of his 'pleasure' in stealing, and that he 'loved' the act, and 'loved' the evil in him. You cannot get 'pleasure' from something, or 'love' something, without thinking it in some respect a good thing. But however that may be, the passage of mine which Hunt cites concerns moral beliefs, not any value beliefs; 'moral beliefs' on my definition are (roughly) value beliefs about the overall goodness of actions, of a kind overlapping with the views of many other actual humans; and I spell this out more carefully by illustrating what 'overlap' amounts to. Those are the kind of value beliefs which – I argue – are such that we are morally responsible for acting on them or not acting on them. My principal claim is that we could not have a moral belief without having some inclination to act upon it. Contrary to Hunt's claim (p. 48) that I hold that we are culpable 'only for acting contrary to our value beliefs', my view is that we are culpable only for acting contrary to our moral beliefs; everyone has value beliefs, but not everyone has moral beliefs. Psychopaths – such as perhaps was Hitler – are not, in my view, culpable for their actions, unless they have moral beliefs. But of course they may be culpable for acting so often contrary to their moral beliefs that they ceased to have any.

Hunt raises two further issues about the kinds of actions for which – given that we have free will (in my sense) – we are morally responsible. The first is this: I claimed that normally we are praiseworthy only for

[1] Blaise Pascal, *Pensées*, translated by A.J. Krailsheimer (London: Penguin Books, 1966), No. 418.

doing actions which we believe to be good but not morally obligatory; but that we are praiseworthy for doing actions believed obligatory only if we do so by overcoming strong contrary temptations (my p. 212-13). Analogously I added later (my p. 220), that we are not praiseworthy for doing an action believed good but not obligatory, if we were not subject to any contrary desire. Hunt rightly claims that I could have given a simpler account of this total view. The complexity arose, I suspect, because of my own doubts about the (p. 220) claim at the stage of the argument when I put forward the (p. 213) claim. However I stand by this view, which results from my reflection on the principles lying behind what I suggest are our intuitive judgements about many particular cases; and I do not find Hunt's judgement about the woman who rushes into the burning building at all persuasive. She instinctively does the right thing, and so deserves admiration for her character from which her action inevitably and unthinkingly flows, but she can't deserve any extra admiration for acting on that character when circumstances require it. So she doesn't deserve the kind of admiration which belongs to someone who fights contrary temptation in order to do the right thing. And although we should be enormously grateful to God for our creation, I do not think that we should regard God as praiseworthy in the same sense as humans are praiseworthy, except in respect of his actions when incarnate and when in my view he was tempted to do and could have done less than the best, even though he could not have failed to fulfil his obligations.[2] The second issue concerns whether we are more blameworthy for successful wrongdoing than for unsuccessful wrongdoing. I stick by my view that on a libertarian account of free will it is irrational to hold a bomber more blameworthy when his bomb works than when it doesn't. We owe reparation for the harm we cause, and the successful bomber will owe a lot more reparation than the unsuccessful one; but we must not confuse his resulting indebtedness with his being more blameworthy. And I mentioned (my pp. 211-12) additional reasons why the criminal law is right to punish successful bombers more severely than unsuccessful ones.

[2] On the point that if God became incarnate as Jesus Christ, he could not have done wrong, but he could have done less than the best, and is praiseworthy for not doing less than the best, see my *Was Jesus God?* (Oxford: Oxford University Press, 2008), pp. 44-47.

NESTED MODES, '*QUA*' AND THE INCARNATION

ALEXANDER R. PRUSS

Baylor University

Abstract. A nested mode ontology allows one to make sense of apparently contradictory Christological claims such as that Christ knows everything and there are some things Christ does not know.

I. INTRODUCTION

According to traditional Christian doctrine, Christ is God and Christ is human. God knows everything. A human being does not know everything (if only because a human being's finite mind cannot hold all the infinitely many mathematical truths). Thus Christ knows everything and does not know everything, surely a contradiction. Likewise, God always feels infinite joy, while Jesus Christ was tortured to death. Thus, Christ felt infinite joy and yet was tortured to death. A standard response to these apparent contradictions is that Christ *qua* God feels infinite joy and knows everything, but *qua* human he was tortured and there are things he does not know. However, this '*qua*' not only seems to be an obscure cheat, but it is not clear how it removes the contradiction (Morris, 1987).

The '*qua*' move is not a cheat. I will start with a nested mode or trope ontology inspired by Spinoza and Aristotle. Then I will discuss how we would express facts within that ontology in English, and show that very naturally one gets locutions that have a formal structure similar to the Christological case, including that mysterious '*qua*'. I will outline the beginning of a semantics for the Christological locutions. Finally, I will discuss the possibility of extending the story outside of the nested mode context, and how one might reconcile the account with divine simplicity.[1]

[1] For an excellent thorough recent discussion of the ways of understanding the Christological '*qua*', see Pawl (2015).

II. NESTED MODES AND ESSENCES

Modes are particular instances of properties. They are often called 'tropes' in contemporary metaphysics, but I will use the older term 'mode' so as not to suggest a trope-bundle ontology (what I say may be compatible with a trope-bundle ontology, but does not require it). The accidents of Aristotelian ontology should be seen as modes, and in those Aristotelian ontologies that suppose individual essences, essences are modes as well. And the term 'mode' is found throughout the work of Descartes, Leibniz and Spinoza. Beyond trope theory proper, the qualia of some contemporary philosophies of mind are probably best seen as modes of conscious being or modes of thought.

On a mode ontology, whenever an object satisfies a fundamental unary predicate, it has a corresponding mode. Thus, if 'is negatively charged' is a fundamental unary predicate, and Ellie the electron is spinning, then Ellie has a mode of spin, *Ellie's spin*. If Pam the positron is also spinning, even if its spin is just like Ellie's, Pam's spin is still something distinct from Ellie's.

There will also be modes corresponding to non-unary predicates. If 'loves' is fundamental, then when Romeo loves Juliet, there is at least one mode of love. There are a number of options here:

(1) There is a single relational mode, which can be denoted either 'Romeo's loving Juliet' or 'Juliet's being loved by Romeo', and both Romeo and Juliet have that mode.
(2) There are two relational modes, Romeo's loving Juliet and Juliet's being loved by Romeo, and Romeo and Juliet have them respectively.
(3) There is only a single relational mode, Romeo's loving Juliet, and Romeo has it.
(4) There is only a single relational mode, Juliet's being loved by Romeo, and Juliet has it.

Given that coming to be loved appears to be a mere Cambridge change, while coming to have a mode appears to be a non-Cambridge change, we have good reason to opt for (3). But perhaps I am wrong about that, and anyway in the case of other predicates other options may be more appropriate.

I leave open whether there are modes corresponding to non-fundamental predications (but see Section 3).

Now, a central insight in Spinoza's ontology is that modes can nested. Thus, Ellie's spin might be upward, and then assuming the direction of spin is fundamental, Ellie's spin will have its mode of upwardness. In this case, we will say that the upwardness is a *remote mode* of Ellie while it is presumably an *immediate mode* of Ellie's spin. Note that this example suggests a promising way of connecting nested mode ontology with theories of determinables. Perhaps whenever ν is a mode of μ, then μ corresponds to a determinable while ν to its determinate.

In general, we can say that ν is a remote mode of *x* (where *x* is an individual or a mode) provided that there is a mode μ such that (a) μ is distinct from ν and *x* (in Section 6 we will discuss whether something could in be a mode of itself), (b) ν is a mode of μ and (c) μ is a mode of *x*. And an immediate mode of *x* is a mode of *x* that is not a remote mode of *x*.

Spinoza, of course, goes overboard on this: he makes everything, with the exception of God, be a mode of something else, and so we ourselves end up being infinitely remote modes of God. But we need not follow him there.

Next, let us introduce an Aristotelian element. One of the modes of an individual is an *essence*. Being a mode, this is an individual essence: although you and I have humanity as our essences, yours is yours and mine is mine.

We now have a decision point that will be crucial. We could take an essence to be simply yet another immediate mode of an individual. But there is a more daring move possible, one that both has metaphysical benefits and will be crucial to our account of the Incarnation: we could suppose that only essences are immediate modes of individuals.

The main metaphysical benefit of this supposition is that it allows us to give an elegant account of what an essence is. As Fine (1994, 1995) has shown, the Aristotelian notion of an essence should not be analyzed in terms of *modally essential* properties, i.e., properties that an entity could not fail to have. Essences are more explanatorily fundamental than the accidents.[2] The supposition that essences are the immediate modes of individuals then provides us with an elegant account of essences that highlights this fundamentality. For all other modes of an individual will be modes (immediate or remote) of essences, but it is reasonable to take

[2] For other attempts to account for the fundamentality of essences, see Gorman (2005) and Pruss (forthcoming).

a mode of x, at least when the mode is distinct from x, to be dependent on x. (How our hypothesis connects with modal notions of essential properties is a question for further investigation.)

Thus, given our supposition, we can give an elegant account of the notions of individual, essence and accident in terms of modes and remoteness. An individual is something that is not a mode of something else. An essence is an immediate mode of an individual. An accident is a remote mode of an individual.

I will henceforth make the supposition that only essences are immediate modes of individuals.

If we combine this supposition with the hypothesis that the nesting of modes corresponds to the determinable-determinate relationship, then we conclude that all our other fundamental properties are determinates of the property corresponding to our essence, presumably our humanity. My believing that the Pythagorean Theorem is true is then a way of my being human, just as being green is a way of having color. This is an attractive way of thinking about ourselves.

III. TALKING ABOUT NESTED MODES

Suppose that Sally has a mode of running and her mode of running has a mode of quickness. Then we could correctly say:

(5) Sally is running quickly.

and conclude from this that:

(6) Sally is running.

and

(7) Sally is quick.

But the second conclusion can mislead. For suppose that she is also texting. We had better not conclude from (7) and

(8) Sally is texting

that

(9) Sally is texting quickly.

It is quite clear in our mode ontology why the inference here fails. Sally's quickness is a mode of her running, but perhaps not a mode of her texting. For Sally to be texting quickly, it is not enough that she have

a quickness and a texting, but she would need a quickness that is a mode of the texting[3] (assuming, contrary to fact, that texting is fundamental).

Suppose that Sally's texting is slow. Then just as we got to say that she is quick, we get to say that

(10) Sally is slow

(or maybe we prefer the wording: 'Sally is being slow') or even

(11) Sally is not quick.

There is, nonetheless, an air of contradiction between (7) and (10), and even more so between (7) and (11). To clear this up, we might expand on these to say:

(1) Sally is quick *qua* someone running.
(2) Sally is slow *qua* someone texting.
(3) Sally is non-quick *qua* someone texting.

For instance, the last of these says that she is texting, but her texting isn't quick.

Thus, expressing innocent claims about nested modes leads quite naturally to adverbial predications like (5), to apparently contradictory claims like (7), (10) and (11), and finally to '*qua*' claims.

Nonetheless, not all sentences of the above sort need correspond so neatly with claims about nested modes. In the above, we assumed that 'is running', 'is texting', 'is slow' and 'is quick' are fundamental. But of course in the case of 'is texting' this is almost surely not so, and the case of the other predicates is also not clear.

There are two possibilities here. If we have an abundant mode ontology, where every correct predication corresponds to a mode, then the above needs no adjustment. But if we have a sparse mode ontology, where only fundamental correct predications correspond to modes, we need to complicate matters. Presumably, non-fundamental claims are grounded in fundamental ones, and the above story will apply in the fundamental cases. Likewise, it is plausible that if we can resolve the apparent contradictions in the cases of fundamental Christological predications, then it should be possible to generalize to the non-fundamental case. There may be difficulties, but apart from the special case of negative predications – which we will discuss in Section 4.4 – we

[3] This is the same issue that Geach (1956) famously discusses for 'good': that a basketball player is good and also a golfer does not imply that she is a good golfer, much less a good human being.

leave those for future research. Thus, from now on, I will simply pretend that all the positive predicates we are concerned with are fundamental.

Finally, a verbal point should be noted. When Sally's running has a mode of quickness, it is correct to say both that Sally's running is quick *and* that Sally is quick. But sometimes a different predicate is applied in the case of the mode and the individual. For instance, suppose Sally's texting is rife with spelling errors. But we shouldn't say that Sally is rife with spelling errors. In fact, even in the case of running, when we say that Sally is quick and that Sally's running is quick, we are using 'is quick' in two different senses.

The verbal point is particularly apposite when we are talking of modes of the essence in our mode ontology. For instance, when Socrates knows *p*, there is a knowledge-of-*p* mode, which on this ontology is a mode of Socrates' humanity. In virtue of this knowledge mode, Socrates knows *p*. But it is incorrect to say that Socrates' humanity knows *p*. We do not, in fact, have a convenient way of expressing the predicate that applies to Socrates' humanity in virtue of its possession of the knowledge mode. Perhaps we might say that Socrates' humanity is determined to human knowledge of *p* or that Socrates' humanity confers knowledge of *p* on him. It should not, however, surprise us if we do not have readymade terms when we analyze things so finely.

IV. CHRISTOLOGICAL APPLICATION

4.1 Basic account

Suppose now that – as far as we know uniquely – Christ is an individual with *two* essences: humanity and divinity. And then each of these essences has the kinds of modes proper to it.

When Socrates feels pain, Socrates' essence – his humanity mode – has a pain mode. When Christ suffers, his essence of humanity has a pain mode. That Socrates' humanity has a pain mode is sufficient to make it true that Socrates feels pain. That Christ's humanity has a pain mode is sufficient to make it true that he feels pain. 'Christ feels pain' is just as literally true as 'Socrates feels pain', and both statements are true for the same kinds of reasons.

Likewise, just as God the Father has a mode of infinite joy, which is a mode of his divinity, so too Christ's divinity has a mode of infinite joy. And just as the fact that the Father has a divinity with a mode of infinite

joy makes it true that the Father has infinite joy, so too that Christ has a divinity (the numerically same divinity, according to orthodox Trinitarianism) with a mode of infinite joy makes it true, and true in exactly the same sense, that Christ has infinite joy.

At the crucifixion, it would thus be correct to say:

(1) Christ has infinite joy
(2) Christ is in horrendous pain.

The two claims are no more in conflict than the claims that Sally is quick and Sally is slow.

And just as in Sally's case we removed the appearance of contradiction by saying that she is quick *qua* running and slow *qua* texting, here we can say:

(1) Christ has infinite joy *qua* divine
(2) Christ is in horrendous pain *qua* human.

There is nothing particularly mysterious about the '*qua*' in (17) and (18): it simply indicates which mode – the divine essence or the human essence – the joy and pain modes are respectively modes of.

We could also express (17) and (18) adverbially, though the locutions would sound a little strange:

(1) Christ is infinitely joyfully divine.
(2) Christ is horrendously painfully human.

When only one of Christ's essences is conversationally salient, we can omit the mention of essence, and simply make claims like (15) and (16).

In the case of ordinary human beings, we could always add '*qua* human' to attributions of pains, pleasures, and the like. But there would be no point, since only one essence – an essence of humanity – is salient when we talk of an ordinary human being, as an ordinary human being (one can take this as partly stipulative of 'ordinary') has only one essence.

4.2 Metaphysical truth conditions

Facts about a nested mode ontology are naturally expressed with *qua* locutions. When dealing with positive *F*, a metaphysical truth condition (cf. Sider 2011, Section 7.4) for

(3) *x* is *F qua G*

is something like

(4) has a *G*ness mode which in turn has an *F*ness mode.

As noted at the end of Section 3, '*F*ness mode' should be understood loosely: '*F*ness mode' need not describe a mode by virtue of which *x*'s *G*ness mode is *F*, but rather perhaps a mode by virtue of which *x*'s *G*ness mode makes *x* be *F*.

But of course Scripture and liturgy do not use locutions like (21). Instead, we have simple predications, like that Christ existed before all (created) things or that Christ died on the cross. How do we give metaphysical truth conditions for such predications?

First, some sentences use a predicate that simply expresses the individual's essence. Call these 'essence predicates'.[4] For these, the truth conditions are somewhat trivial: the predicate applies to the individual if and only if the individual has a requisite essence. Thus, the truth condition for 'Christ is human' is simply that Christ has a humanity mode.

Now consider a positive non-essence predicate *F*. Depending on wording and context, there will be two possibilities for a truth condition for

(5) *X* is *F*,

where *X* is a term denoting Christ. First, we might have *unspecified truth conditions*, namely:

(6) *X* has an essence that has a mode of *F*ness

or we might have one of the two *specified truth conditions*:

(7) *X*'s divine essence has a mode of *F*ness

or

(8) *X*'s human essence has a mode of *F*ness.

Sometimes the choice of designator *X* makes clear whether (25) or (26) is appropriate by making either divinity or humanity relevant. If *X* is 'Jesus', 'the Son of Mary' or the like, then humanity is likely to be relevant, while if *X* is 'the Son of God', 'the Second Person of the Trinity' or a similar term, then divinity is likely to be relevant. On the other hand, the term 'Christ' does not clearly make one or the other nature relevant.

In the context of a conversation that defers linguistically to Christian orthodoxy, where the interlocutors use language in the way Christian

[4] They are not to be confused with 'essential predicates' in the modal sense, namely those predicates that must apply to an individual if that individual is to exist.

orthodoxy traditionally does,[5] the predicate *F* may itself be an indicator of whether we should adopt the reading in (25) or the one in (26). Divinity is relevant when we speak of Christ creating and humanity when we speak of Christ being born.

But often unspecified truth conditions seem appropriate. A competent speaker can responsibly say of a distant animal that it is running without having to know whether it is a dog or a fox, so that even if the animal is a fox, we perhaps should not say that truth condition is

(9) The animal's vulpine essence has a mode of running

but

(10) The animal has an essence that has a mode of running.

Likewise, when wording and context do not make salient a particular essence of Christ, the unspecified reading (24) may well be more appropriate. In theological discussion, liturgy and other specialized contexts, one of the two specified truth conditions for Christological predications may well be appropriate. But in more ordinary contexts, the unspecified truth conditions are likely to often be appropriate. As always with such contextual matters, precise rules are not available.

4.3 The grammar of 'qua'

Pawl (2015) classifies '*qua*'-based responses to the apparent Christological contradictions on the basis of the role played by the term '*qua*': it could signal tense, signal a tacit sortal, or modify the subject, predicate or copula. Perhaps surprisingly it is not quite clear how the above account fits into Pawl's classification. On the present account, in the context where Sally's running is salient, all of the following sentences will have the same metaphysical truth condition, namely that Sally's running mode has a quickness mode:

(1) Sally is quick.
(2) Sally is quick *qua* running.
(3) Sally's running is quick.
(4) Sally is running quickly.

At this point there appears to be a choice point for our theory. Do we take (30) to mean something like (31) or something like (32)? If we go

[5] Which may or may not mean that the interlocutors themselves subscribe to this orthodoxy. They may defer linguistically and then go on to criticize.

for (31), then it seems that our '*qua*' modifies the subject, making clear that the subject is Sally's running mode. On the other hand, if we go for (32), then it seems the predicate is being modified. Whereas it seemed that the predicate was simply 'is quick', it really is 'is running quickly.' But perhaps this is only an illusion of a choice point. After all, all the four sentences in this context have the same metaphysical truth conditions. Perhaps there is but one fact here that is expressed in four different ways: elliptically in (29), with '*qua*' in (30), and then in two grammatically distinct ways in (31) and (32).

We can, however, say something more in terms of Sider's (2011) notion of carving at the joints, if we take the nested mode ontology to be not only a correct ontology, but a correct fundamental ontology. On the nested mode ontology, it is (31) that carves closest to the joints of reality, as the ideology of (31) is precisely that of the nested mode ontology, and so there is a sense in which the present account is a subject-modification account. But if our only purpose is to defend the coherence of the doctrine of the Incarnation, then we do not need to embrace nested mode ontology as the fundamental ontology. One might, for instance, say that while it's true to say that there are nested modes, there is some more fundamental account of nested mode language. In that case, (31) may no longer carve closest to the joints.

4.4 Negative predications

In the joy and pain examples of Section 4.1, I made things easy for myself: I considered what were at least at first sight[6] positive predications. Indeed, one might be unimpressed by the account I gave, since perhaps even an ordinary human could experience great joy and great pain at the same time.

But now consider something which is much more plausibly contradictory. Let *p* be some proposition that God knows but no human being knows. Then it seems we have:

(1) Christ knows *p*
(2) Christ does not know *p*.

And here the contradiction appears much more blatant.

[6] One might think with Augustine that evil is privation, and hence pain, or at least whatever is responsible for the badness of the pain, is at least partly constituted by the absence of something.

Now, if there are negative modes, like a mode of ignorance, then there is no special difficulty here. Christ's divinity has a mode of knowledge of *p* while Christ's humanity has a mode of ignorance of *p*. But it is implausible that for every case where *x* lacks a positive mode, *x* has a distinct mode of lacking such a positive mode. That would make each of us not only have infinitely many modes, but modes beyond cardinality.[7]

But of course negative predications also occur in perfectly ordinary nested mode cases. When Sally is running quickly but texting slowly, then she is quick and not quick. She is quick because her mode of running has a mode of quickness, and she is not quick because her mode of texting lacks a mode of quickness. Again, inserting '*qua*' will remove the air of contradiction.

Thus, Christ knows *p* because his divinity has a mode of knowledge of *p*, and Christ does not know *p* because his humanity lacks any mode of knowledge of *p*. Note that the grounds for which it is correct to say that Christ knows *p* are just like grounds that make it correct to say that the Father knows *p*: that there is a mode of knowledge had by the mode of divinity. And just as Socrates may not know *p*, because his humanity lacks a mode of knowledge of *p*, so too Christ's ignorance is grounded in *his* humanity's lacking a mode of knowledge of *p*.

One might object as follows. The right metaphysical truth conditions for

(3) Socrates knows *p*

are not

(4) Socrates' human essence has a knowledge-of-*p* mode.

For it is epistemically possible for (35) to be true even if it turns out that Socrates is non-human. Rather, we should take the truth conditions for (35) to be:

(5) Socrates has an essence that has a knowledge-of-*p* mode.

But then

(6) Socrates does not know *p*

[7] I am at best acquainted with finitely many sets. If for every set *a* that I am not acquainted with I have a distinct mode of non-acquaintance with *a*, then I will have a mode of non-acquaintance for each set, except for finitely many. And the class of all sets has no cardinality, nor does it gain a cardinality when we take finitely many members away.

will have as its truth conditions:

(7) Socrates has no essence that has a knowledge-of-*p* mode.

By parallel, then, in a context where unspecified truth conditions are appropriate, we should say that (34) implies:

(8) Christ has no essence that has a knowledge-of-*p* mode.

And (40) is false on our account, since God has an essence with a mode of knowledge of *p*, namely his divine essence.

Now, first of all, it is not clear that *metaphysical* truth conditions have to track what is epistemically open to the individual (Sider 2011, Section 7.4 expressly denies it). It might be epistemically open that tables have no particles in them, but the true metaphysical truth conditions for 'There is a table' might still be that there are particles arranged tablewise. So it is still open to insist that the right metaphysical truth conditions for (35) are given by (36), and hence the metaphysical truth conditions for (38) are:

(9) Socrates' humanity has no knowledge-of-*p* mode.

That said, it is better if the truth conditions remain closer to being epistemically available to the individual, as was discussed in the case of the distant fox.

But even granting that (37) is the right truth condition for (35), the objection assumes a wide scope[8] reading of the negation in (38), namely that (38) denies (35). But one can also take the negation to have narrow scope, in which case the nested mode ontology's truth condition could be:

(10) Socrates has an essence that has no knowledge-of-*p* mode.

Similarly, when we hear that Sally is not quick, we do not assume that she is nowise quick. Rather, usually there is a contextually relevant attribute and we are being told that she is not quick at that, and even when context does not specify the attribute, we do not assume that she is nowise quick.

We do not normally need to make the narrow–wide scope distinction in the case of attributes like knowledge of *p*, but that is because normally we deal with beings like Socrates that have only one essence, and so there

[8] Or medium wide. Perhaps (39) should be read as committing one to the existence of Socrates, while the widest scope reading of the negation in (38) would not.

is only one contextually relevant essence that might or might not possess a knowledge-of-*p* mode. Given the auxiliary premise that Socrates has only one essence, one can indeed derive (39) from the denial of (36), and normally we tacitly accept such premises.

But of course once we start speaking about beings with more than one essence, we need to make a decision on how to read negative sentences. A natural move is to take negations to be narrow scope, unless they are qualified or strengthened in a way that forces a wide scope reading, as in: 'Christ nowise does wrong.'

V. WITHOUT NESTED MODES

One may even be able to have the main outlines of the above story without nested modes. Instead of taking

(11) Socrates knows *p*

as grounded in

(12) Socrates' humanity has a knowledge-of-*p* mode,

we could take it as grounded in:

(13) Socrates' humanity is *p*-knowledge-conferring,

where we do not take the application of the predicate 'is *p*-knowledge-conferring' to be grounded in a knowledge-of-*p* mode. We can understand the predication in (45) in line with our favorite non-trope theory of predication, whether Platonist, conceptualist or nominalist. We still need a commitment to individual essences, but not to any remote modes.

If we proceed in this way, however, we lose the analogy between Christological predications and sentences about Sally's being quick (*qua* running or *qua* texting). That loss makes the theory less plausible in a way that perhaps cancels out the benefits of not having to have to nested modes. Nonetheless, an ingredient from the modified theory may be useful in the next section. And even if less plausible, the theory may have a use in showing the coherence of the doctrine of the Incarnation, much in the way that van Inwagen (1995) has used relative identity theory to show the coherence of the doctrine of the Trinity without endorsing relative identities in ordinary cases.

VI. DIVINE SIMPLICITY

The above account of Christological predication requires that God have a divine essence, and that this essence in turn have modes. On its face this is incompatible with divine simplicity: it implies something like the essence-accident complexity that is denied by proponents of divine simplicity like Aquinas (1920, I.3.6).

But this is too quick. When we say that God has an essence ε and the essence has a mode μ, this does not force us to admit three things. For it might be that they are all identical, so that God = ε = μ, just as an officer, a doctor and a wife may all be one person. Granted, the mode μ is a mode of the essence ε, and the essence ε is an essence of God, but just as one can be one's own enemy, so too perhaps something can be its own mode and its own individual essence. (It won't do to object that the essence is abstract while God is concrete, since modes are best taken to be concrete.)

Consider, after all, a plausible functional characterization of modes: an immediate *F*ness mode μ of *x* is an object μ such that μ, or at least *x* and μ together,[9] fully grounds *x*'s being *F*. But when *x* itself fully grounds *x* being *F*, then the functional characterization makes *x* be its own *F*ness trope. And on the Oppy-Brower-Pruss account of divine simplicity (Oppy 2003; Brower 2008; Pruss 2008), God truthmakes, and hence grounds, all the intrinsic attributes of God. (The question of the extension of that account to non-intrinsic attributes, while indeed a challenge to the proponent of simplicity, is not a *special* difficulty for our nested mode account of the incarnation.)

Alternately, one might simply posit the identity between God and his essence, as typical proponents of divine simplicity do, and then apply the suggestion of Section 5 so that one can do without nested modes. Or at least without them on the side of the divine essence. Perhaps, for instance, something like trope theory is the right account of predication in the case of creaturely subjects but ostrich nominalism is the right account for God.

[9] This option or some further refinement may be needed if we are to leave open an understanding of transsubstantiation on which the modes of bread and wine persist even when there no longer is any bread or wine, but only Christ's body and blood. For in such a case, we probably wouldn't want to say that the bread's persistent whiteness trope grounds the bread's being white, since the bread is presumably not white when it doesn't exist (maybe, though, we could take it to ground the past whiteness of the bread?).

VII. CONCLUSIONS

Nested mode facts are naturally expressed with '*qua*' sentences such that when the '*qua*' clause is left off, some of the sentences will look contradictory. A start of a metaphysical semantics for Christological predication in terms of modes of the human and divine essence can be given.

The account has the very significant advantage that ordinary predicates like 'is sitting' to Christ can apply to Christ for exactly the same reasons that they can apply to us: the subject's human essence has the relevant mode. It is indeed a desideratum on a theory of the incarnation that there be such a uniformity between the application of ordinary creaturely predicates to Christ and to us – Christ is a human being like us, except in sin. In this way, the present account will be superior to property transference accounts like Leftow's (2002) and Stump's (2002) on which ordinary predicates apply to Christ because they apply to a proper part of Christ, say his body and soul composite, since it does not seem that in our own case the ordinary predicates apply to us because they apply to a proper part of us (for a discussion of such property borrowing accounts, see Jeffrey 2014).[10]

REFERENCES

Aquinas, Thomas. 1920. *The Summa Theologica of St. Thomas Aquinas.* Trans. the Fathers of the English Dominican Province.

Brower, Jeffrey. 2008. 'Making Sense of Divine Simplicity', *Faith and Philosophy*, 25: 3-30

Fine, Kit. 1994. 'Essence and Modality', in J. Tomberlin (ed.), *Philosophical Perspectives VIII: Logic and Language* (Atascadero, CA: Ridgeview), pp. 1-16

Fine, Kit. 1995. 'Senses of Essence', in W. Sinnott-Armstrong *et al.* (eds.), *Modality, Morality, and Belief* (Cambridge: Cambridge University Press), pp. 53-73

Geach, Peter T. 1956. 'Good and Evil', in S. M. Cahn and J. G. Haber (eds.), *20th Century Ethical Theory* (Upper Saddle River, NJ: Prentice Hall), pp. 300-306

Gorman, Michael. 2005. 'The Essential and the Accidental', *Ratio*, 18: 276-289

[10] I am grateful to Jared Brandt, Rob Elisher, Gideon Jeffrey, Matt Wilson, and the participants of the conference *Analytic Theology: Faith, Knowledge, and the Trinity* held in Prague on 19-20 September 2013, for a number of comments on an earlier version that have greatly improved this work. The Prague conference was generously funded by the John Templeton Foundation (grant #15571).

Jeffrey, Gideon. 2014. 'Property Borrowing and Compositional Christology', Society of Christian Philosophers Midwest Regional Conference (Palos Heights, IL) March 27, 2014

Leftow, Brian. 2002. 'A Timeless God Incarnate', in S. T. Davis, *et al.* (eds.), *The Incarnation: An Interdisciplinary Symposium on the Incarnation of the Son of God* (Oxford: Oxford University Press), pp. 273-302

Morris, Thomas V. 1987. *The Logic of God Incarnate* (Ithaca: Cornell University Press)

Oppy, Graham. 2003. 'The Devilish Complexities of Divine Simplicity', *Philo*, 6: 10-22

Pawl, Timothy. 2015. 'Temporary Intrinsics and Christological Predication', in J. Kvanvig (ed.), *Oxford Studies in Philosophy of Religion* (Oxford: Oxford University Press)

Pruss, Alexander R. Forthcoming. 'Aristotelian Forms and Laws of Nature', *Analysis and Existence*

Pruss, Alexander R. 2008. 'On Two Problems of Divine Simplicity', in J. Kvanvig (ed.), *Oxford Studies in Philosophy of Religion*, Vol. 1 (Oxford: Oxford University Press), pp. 150-167

Sider, Theodore. 2011. *Writing the Book of the World* (Oxford: Clarendon)

Stump, Eleonore. 'Aquinas' Metaphysics of the Incarnation', in S. T. Davis, *et al.* (eds.), *The Incarnation: An Interdisciplinary Symposium on the Incarnation of the Son of God* (Oxford: Oxford University Press), pp. 197–218

van Inwagen, Peter. 1995. 'And Yet They Are Not Three Gods But One God', in *God, Knowledge, and Mystery* (Ithaca: Cornell University Press), pp. 222-259

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

[5] J. Roland and J. Cogburn, 'Anti-luck Epistemology and Necessary Truths', *Philosophia*, 39 (2011), 547-561. Safety theorists do often mention the 'illusory', as Pritchard puts it at one point, problem safety has in dealing with necessity, but – as the term 'illusory' suggests – they universally think it can be sidestepped in the manner discussed in the main text; no extended discussion is deemed necessary.

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.

Acknowledgment. I am grateful for the comments of Duncan Pritchard, Dani Rabinowitz and Chris Tucker on an early draft of this paper. I am also grateful for the comments of Charity Anderson, Matthew Benton, Julien Dutant, John Hawthorne, and Brian Heddon on a later version of it. Finally, I did most of the preliminary work for this whilst on a period of special leave financed by the 'New Insights in Religious Epistemology' project at Oxford. This project was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation.

DIVINE LOVE AND THE ARGUMENT FROM DIVINE HIDDENNESS

EBRAHIM AZADEGAN

Sharif University of Technology, Teheran

Abstract. This paper criticizes one of the premises of Schellenberg's atheistic argument from divine hiddenness. This premise, which can be considered as the foundation of his proposed argument, is based on a specific interpretation of divine love as eros. In this paper I first categorize several concepts of divine love under two main categories, eros and agape; I then answer some main objections to the ascription of eros to God; and in the last part I show that neither on a reading of divine love as agape nor as eros can Schellenberg's argument be construed as sound. My aim is to show that even if – *contra* Nygren for example – we accept that divine love can be interpreted as eros, Schellenberg's argument still doesn't work

I. THE ARGUMENT FROM DIVINE HIDDENNESS

Many people are perplexed that God should permit a situation in which many human beings live in incomprehension and bewilderment regarding His existence, while all the time He could save humanity from such a predicament. The problem of 'divine hiddenness' refers to the problem for the theist in holding that we human beings live in a world in which we have limited cognitive faculties; knowledge of God is essential for our flourishing in this-worldly and otherworldly life; and where, in addition, God, the omniscient, omnipotent, and perfectly loving, has permitted us to live in this bewilderment and perplexity regarding His attributes and existence, all the while knowing that it is essential for our well-being.

EUROPEAN JOURNAL FOR PHILOSOPHY OF RELIGION 6/2 (SUMMER 2014), PP. 101-116

More than anyone else, John Schellenberg is responsible for renewing and framing the contemporary debate about divine hiddenness.[1] Roughly speaking, he sees the fact that there are some non-resistant non-believers as being incompatible with the existence of a perfectly loving God.

The final version of his argument is as follows:

(1) Necessarily, if God exists, anyone who is (i) not resisting God and (ii) capable of meaningful conscious relationship with God is also (iii) in a position to participate in such relationship (able to do so just by trying).
(2) Necessarily, one is at a time in a position to participate in meaningful conscious relationship with God only if at that time one believes that God exists.
(3) Thus (from 1 and 2), necessarily, if God exists, anyone who is (i) not resisting God and (ii) capable of meaningful conscious relationship with God also (iii) believes that God exists.
(4) There are (and often have been) people who are (i) not resisting God and (ii) capable of meaningful conscious relationship with God *without* also (iii) believing that God exists.
(5) Thus (from conjunction of 3 and 4) God does not exist.[2]

Premise 1 of his argument is the most important premise, and of course requires a separate argument. To support premise 1 Schellenberg argues that if God is loving, He seeks an explicit reciprocal relationship with us, involving not only such things as divine guidance, support, and forgiveness, but also what leads to the faith, trust, obedience, and worship that would contribute to human well-being. He therefore concludes that if God seeks a reciprocal relationship, then for any human person H at any time *t*, if H is at *t capable of* relating personally with God, H at *t* is in a position to do so just by choosing.

---

[1] Schellenberg's main discussions of this subject can be found in John L. Schellenberg, *Divine Hiddenness and Human Reason* (Ithaca: Cornell University Press, 1993), and 'What the Hiddenness of God Reveals: A Collaborative Discussion', in *Divine Hiddenness: New Essays*, ed. D. Howard-Snyder and P.K. Moser (Cambridge: Cambridge University Press, 2002), pp. 33-61. He responds to several objections to his argument in 'The Hiddenness Argument Revisited (I)', *Religious Studies*, 41 (2005), 201–15, and 'The Hiddenness Argument Revisited (II)', *Religious Studies*, 41 (2005), 287–303. For his recent and most developed version of the argument see his *The Wisdom to Doubt* (Ithaca: Cornell University Press, 2007), Chaps. 9 and 10.

[2] Schellenberg *The Wisdom to Doubt*, pp. 204–206.

In this paper I focus on the notion of the *capability of a person to be the object of divine love and to engage in personal relational love with God* in order to show that Schellenberg has incorrect expectations of divine love and the human–divine relationship. He has dismissed the vital role that God expects us (human beings) to play in our relationship with God. In this regard, I shall try to answer the following three questions:

(1) What type of love is required in order for us to have a *personal reciprocal* relationship with God?
(2) Is it *legitimate* to ascribe this type of love to God? And under what conditions?
(3) If it is legitimate to ascribe to God a type of love that requires a personal reciprocal relationship, then what would be the conditions of the object of this type of love? In other words, who is *capable of* being in the reciprocal love relationship with God?

The paper is organized as follows. First, I shall briefly explain different concepts of divine love, and then I shall show that, drawing from Nygren's account, one can distinguish two categories of divine love: eros and agape.[3] I shall show that eros, but not agape, as divine love, requires a personal reciprocal relationship and union with God. Then I shall answer some of the main objections to attributing eros to God, namely the Aristotelian and Platonic objections. Finally, I shall interpret Schellenberg's argument according to both types of divine love (eros and agape) in order to show that in neither of these forms is his first premise compelling, because agape does not require a reciprocal relationship and eros requires a high standard of capability for the object of love, which normal human beings lack unless they make sufficient efforts to obtain it. In other words, merely being non-resistant to God is not enough.

[3] Anders Nygren, *Agape and Eros*, trans. Philip S. Watson (London: Society for Promoting Christian Knowledge [SPCK], 1982). It is worth noting that Nygren here distinguishes three types of divine love: eros (the Hellenistic concept of divine love), nomos (the Old Testament concept of divine love), and agape (the truly Christian concept of divine love). He tries to show that both eros and nomos require some special capabilities and admirable values on the part of the object of love, while agape is a type of God's love that cannot be contained within the framework of the Law (like nomos), and is not justified on the ground of merit (like eros), but is based on free grace and groundless, spontaneous, and unmotivated divine love. In this paper I use the important difference between eros and agape to construct an objection to Schellenberg's argument; I think that nomos can be regarded as a token of the eros-type of divine love, broadly speaking because it is based on a sort of value inherent in the object of love as well.

II. DIVINE LOVE

It is not surprising that Schellenberg constructs his argument based on certain expectations that people may have of a loving God given the immense importance of this subject in the Christian tradition. So essential to Christianity is the notion of divine love that it seems no exaggeration to describe it as the heart of the Christian faith. 'God is love and whoever lives in love lives in God, and God in them.'[4] These words, from the first letter of John, powerfully express the Christian image of the relationship between God and humanity. In the same chapter, John also offers a kind of summary of the Christian concept of the human–human relationship on that basis: 'Dear friends, let us love one another, for love comes from God. Everyone who loves has been born of God and knows God. Whoever does not love does not know God, because God is love.'[5]

The claim that *God is love* can be interpreted in at least two different ways. It can be regarded as expressing the idea that (a) God is the source or origin of love from which all lovers derive their love; love is a gift *from* God to human beings, and God causes, gives and inspires love in us, but 'lover' is not a proper attribute of God, or (b) God Himself is the one who loves and He loves us, as well.[6] The first interpretation is not our concern here because it deals with the cause of human beings' love towards God or other beings, while here we are dealing with God's love toward human beings.

The New Testament word for God's generous love is 'agape'. This comes from the Greek translation of the Hebrew word '*ahaba*' in the *Song of Songs*, and subsequently became a typical expression for the Biblical notion of love. The term 'agape' refers to God's love for all humans and all creatures regardless of their righteousness or depravity. Agape is God's disinterested love; God loves in an agapeic manner because of His nature, for He is essentially loving – for *He is love*. It is not the case that love is merely one of the activities of God, such that God might act lovingly in some cases and not in others; rather, love is the essential activity of God, in the sense that all of His actions should be interpreted according to His love. Every activity of God is loving activity, and all His relations with men are characterized by His love. God loves, in the sense of agape,

[4] I John 4: 16.

[5] I John 4: 7–8.

[6] For more details on these interpretations and the relation between them, see Catherine Osborne, *Eros Unveiled* (Oxford: Clarendon Press, 1996), pp. 41–5.

all of His creatures, indifferent to their goodness or badness.[7] Agape is generous, spontaneous, and unmotivated love for all creatures.

However, many believe that we can also ascribe what some have called an 'acquisitive' type of love to God;[8] I shall use the familiar historic term 'eros' to refer to this type of divine love, by drawing a contrast with the other type of divine love that is generous and unconditional. Historically, the word 'eros' refers to attentive, mutual, reciprocal (that is: seeking reciprocity), and even indulgent or egocentric love. The term comes from the ancient Greek notion of erotic love, in which a man longs to acquire what is valuable in his eyes, something that he does not possess but desires to have. Eros, so defined, as the desire for union with the beloved, is egocentric love and is rooted in our needs and weakness. It seems that eros is a result of valuing and appreciating what the lover already lacks or needs.[9]

Based on this narrow interpretation, eros could potentially include a defective element – and, indeed, because of the distinctively human flaws that undergird it, Nygren explains that the notion of divine love cannot possibly be interpreted in the sense of eros. On a properly Christian conception, he contends, agape is the only acceptable interpretation of God's love towards His creatures.[10]

Nygren goes further than criticizing the idea that divine love be considered as eros: he also proposes that should the human desire for union with God be interpreted in such a manner, it would be infected with egocentricity. Such an interpretation, he maintains, should therefore

---

[7] See Nygren, *Agape and Eros*, pp. 108–117.

[8] It seems noteworthy that if we try to place God and humanity in a relationship of love where God is the lover and humanity the beloved, we are likely to start by comparing the two in terms of superiority and inferiority, or bestowal and reception. It seems that there is nothing that we, as humans, can be in a position to supply for God; nor is there anything that we have which is not bestowed by God. So, if we accept the dichotomy between acquisitive and generous love, we might find ourselves committed to the conclusion that since God is perfect and merciful, His love could only be 'generous'; and, conversely, our love toward God can only be 'acquisitive'. In this paper I am concerned with the dichotomy between these two different types of love: the acquisitive (eros) and generous (agape) types of love. In due course I shall refer to some of the proponents of the 'acquisitive' interpretation of God's love.

[9] Nygren, *Agape and Eros*, pp. 158–182.

[10] For more on this dichotomy and the concept of divine love in other monotheistic traditions, see Benyamin Abrahamov, *Divine Love in Islamic Mysticism: The Teachings of al-Ghazâlî and al-Dabbâgh* (London: Routledge, 2011), ch. 1.

be dismissed. On this basis, Nygren criticizes all kinds of Gnostic and mystical practices from the history of religion, claiming that they are ill-formed.[11] Agape, he emphasizes, is the only sense in which the New Testament concept of love can be interpreted.

However, on the basis of an examination of all the uses of the word 'agape' in the New Testament and early Christian texts, Catherine Osborne tries to show that in many cases divine love *can* be interpreted as eros. She criticizes Nygren's concept of the pure Christian concept of love, showing how the authors of the New Testament were influenced by Greek and Hebrew terminology and background knowledge inherited along with it.[12]

The possibility that Christian Scripture provides a basis for interpreting divine love as *both* eros and agape is also raised by Alvin Plantinga.[13] He emphasizes that there are some scriptural images that permit us to ascribe eros to God, and that his love is not exclusively agapeic. He argues that Scripture presents this kind of divine eros as existing within the subjects of the Trinity, especially Jesus. Plantinga cites Isaiah 62: 5: 'As a bridegroom rejoices over his bride God rejoices over you.'[14] The bridegroom rejoicing over his bride does not love her with a merely agapeic love, Plantinga explains: 'He desires and longs for something outside himself, namely union with His beloved.'[15]

Despite these theological responses, I think Nygren's objections to attributing eros to God are worthy of further philosophical investigation. To reach a better understanding of how eros can be attributed to God we shall scrutinize the notion of eros, its roots, the definition of it, and the criterion under which it can be attributed to God. The first questions that should be addressed in this regard are: How can eros, whose ultimate aim is the union with the beloved, be attributed to God? And what criteria would something have to meet to be an object of eros, if the lover is God? In what follows, I shall explore Aquinas' answer to this question; I shall then try to answer some of the objections to viewing eros as a form of divine love.

---

[11] Nygren, *Agape and Eros*, pp. 190–220.

[12] Catherine Osborne, *Eros Unveiled*, chs. 1 and 2.

[13] Alvin Plantinga, *Warranted Christian Belief* (Oxford: Oxford University Press, 2000), pp. 311–323.

[14] *Isaiah* 62:5.

[15] Plantinga, *Warranted Christian Belief*, p. 320.

To be clear: what I am going to suggest regarding the sketched dichotomy between the agape and eros interpretations of divine love, is that even if, *contra* Nygren, we accept that divine love can be interpreted as eros, Schellenberg's argument still doesn't work. If divine love is purely agape, then Schellenberg's argument is in trouble, for God's capacity to love us agapeically is obviously entirely independent of anything about us, including the level of hiddenness or non-hiddenness of God from us. If God's love were purely agape, as Nygren suggests, Schellenberg's argument would fail. But it seems to me that it is plausible to think of God's love as eros in at least some senses. I shall now go on to consider these and show that even if God's love is eros in these senses, Schellenberg's argument still fails.

III. DIVINE LOVE AIMS AT UNION WITH THE BELOVED

In addition to the generous type of love that is love for the sake of the beloved regardless of her intrinsic values, Aquinas distinguishes a type of love that aims at union with the beloved when a lover not only wishes some good for the beloved but also loves the relationship with the beloved to whom he wishes good.[16] For Aquinas, the desire for this love is rooted in the joy of union with the beloved, because 'every lover rejoices at being united to the beloved'.[17] However, Aquinas focuses primarily on the union between man and God as the result of our love towards God. He thinks that the ultimate good for any human person is union with God.[18] This kind of idea should, in fact, be considered the motivation for various mystical practices in the Christian tradition. However, when we talk about God's desire for union with human beings – and in particular when we attribute eros to God – the idea of union seems problematic. What does it mean to say that *God* desires union with *us*? When we understand divine love as agape, and thereby dwell on God's benevolence and His desire for the good of human beings regardless of our goodness or badness (what Aquinas named the simple view of love), there seems to be no problem. Nonetheless, the exact question that I shall pursue is: If it were legitimate to ascribe eros to God, what would be the nature of the union desired and the conditions for bringing it about?

[16] ST I-II q.26 a.4 (NY: Cosimo, trans. Fathers of English Dominican Province, 2007, originally published 1912).

[17] ST I-II q.70 a.3.

[18] See, for example, ST I-II q.1 a.8.

As mentioned, Aquinas distinguishes two different sorts of love, the desire for the good of the beloved and the desire for union with the beloved, which in the case of divine love I have classified under the categories of agape and eros respectively. Here I focus on the more problematic part of this account, which interprets the erotic element of divine love as God desiring union with us, human beings.

Reflecting on this form of love, Eleonore Stump explains that two things seem to be required for union between persons: *personal presence* and *mutual closeness*.[19] Personal presence includes not only causal and cognitive contact between the lover and the beloved, but also their conscious awareness of and shared attention to each other. Mutual closeness is what strengthens the required shared attention and mutual awareness, in a sense that the lover and the beloved become completely engaged with each other. Mutual closeness in the case of two persons can hardly be reduced to propinquity or even lively conversational relations: it seems to include sharing thoughts and feelings as well. And not just any thoughts and feelings will do here: as Stump says, 'For Paula to be close to Jerome, Jerome has to share with Paula those thoughts and feelings of his that he cares about and that are revelatory of him. Jerome would not be close to Paula if he shared with her a great many of his most trivial thoughts but nothing of what was important to him or revealing of him.'[20]

Suppose, however, that Jerome does reveal enough of his thoughts and feelings to Paula; if Paula does not will to understand Jerome's self-revelation, or is not able to understand what he is trying to reveal, she will not be close to him. So, mutual closeness requires self-revelation on the part of the lover *and* the will for reception and the ability of comprehension of the received material on the part of the beloved.[21]

In the case of divine love toward human beings, it seems that according to Theistic religions, God meets both of the conditions for allowing human beings to be in union with Him: God is omnipresent, and has causal relations with and direct and unmediated cognitive access to everything. In addition, God is always and everywhere in a position to pay attention to any creature able and willing to share attention with

[19] Eleonore Stump, *Wandering in Darkness: Narrative and the Problem of Suffering* (Oxford: Oxford University Press, 2010), ch. 6.

[20] Ibid., p. 119.

[21] Ibid.

Him. God also has self-disclosed Himself to be known by us in revelation and creation. He has revealed certain truths about Himself through His prophets and messengers in history, and has revealed (some truths about) His mind and will to human beings. For these reasons, God has disclosed Himself to fulfil His desire for union with human beings.

But the crucial question that arises is whether God's desire for union with human beings is sufficient for our being in a reciprocal relationship of love and union with God. Contra Schellenberg, my answer is 'No', because, as mentioned, in addition to divine omnipresence and self-disclosure, union requires three things on the part of the beloved. First, it requires the beloved's attention to God's presence. This I shall term *praying*: praying seems to be a way in which we can achieve shared attention with God in a way that man wholeheartedly pays attention to God and asks Him for whatever he needs and has a sense that God hears and sees him and pays attention to him. Second, union requires the beloved's will to understand the divine presence in His creatures through His revelations. And third, it requires the beloved's ability to comprehend divine self-disclosure in His revelation. The second and third conditions require thinking and reflection on revelation and Scripture, in addition to continuous repentance for the sins which lead us astray. So it seems that praying, repentance, and reflection on divine revelation are three prerequisites for us to respond to God's eros. These are three conditions which must be met by the beloved man if God's desire for union with him is to be satisfied.

Furthermore, these things may be taken to be a matter of degree. The more a human being becomes purified and sanctified, the stronger the relation of eros he can have with God. How close God becomes to any one person depends on how much that person tries to pray, repent, and be open to the divine free gift. God's grace is free for all of us, but our success in availing ourselves of it depends on us.

In this way, it seems that eros can be attributed to God and yet Schellenberg's argument still not work. What we find, then, through reasoning and consideration of the concept of love in Christianity, is that while every human being is the potential object of God's eros, in fact only meritorious human beings can reach the state of union with their Lord.[22]

[22] There is a long-running debate in the literature regarding the notion of divine grace and whether humanity merits the gifts of God. Nevertheless, my contention here is that

I have argued, then, that whilst it is plausible to consider God's love as eros as well as agape, and thus that the 'quick' way to sidestep Schellenberg's argument (that of simply denying any erotic element in the divine love towards His creatures) is blocked, the sort of erotic element supposed to be present in God's love towards His creatures is nevertheless not sufficient to support Schellenberg. In the next part I shall therefore consider and respond to some Aristotelian and Platonic objections to construing divine love as eros in any sense, in order to shore up my implicit suggestion that the Christian is in fact best advised to allow an element of eros into his or her conception of divine love, rather than simply take the 'quick' way with Schellenberg's argument.

IV. THE PLATONIC OBJECTION TO ATTRIBUTING EROS TO GOD

As defined, eros involves desire for union with the beloved, and is a matter of longing, perceiving, noting, appreciating, delighting in, and relishing. A primary question then, is: May eros be ascribed to God? As I shall explain, from a Platonic point of view, it seems that eros may hardly be supposed a suitable interpretation of divine love.[23]

Traditionally, every discussion of love – whether celestial, romantic, ascending, or descending love – begins with Plato. His ideas on love are expressed mainly in the *Symposium*, *Lysis*, and the *Phaedo*. While in the *Symposium* Socrates reports the views of Diotima of Mantinea, a wise woman, philosopher, and mystic, in the two other books Socrates himself is the main interlocutor. In the *Symposium* Socrates reports his dialogue with Diotima regarding the nature of love and its effects. Diotima, we are told, declares that eros is not a god but at most a daemon or spirit, an intermediary between gods and mortals. According to Plato, Eros is a son of Poros (the symbol of resources, wealth, and possessions) and Penia (the symbol of poverty).

> So because Eros is the son of Poros and Penia, his situation is in some such case as this. First of all, he is always poor; and he is far from being

only meritorious human beings are capable of being in union with God, regardless of the way a human being may achieve such a meritorious character, and no matter whether it is itself through grace of God or not.

[23] Catherine Osborne examines this claim in detail and concludes that because Plato intentionally uses direct quotes from Diotima it is better to ascribe this claim to her rather than Plato or Socrates. See Osborne, *Eros Unveiled*, pp. 86–116.

> tender and beautiful, as many believe, but is tough, squalid, shoeless, and homeless, [...] he has the nature of his mother, always dwelling with neediness. But in accordance with his father he plots to trap the beautiful and the good, and is courageous, stout, [...] desirous of practical wisdom and inventive.[24]

But of course deficiencies and shortages cannot be attributed to God. 'How then could he who is without a share in the beautiful and good things be a god?' Diotima asks; and Socrates replies: 'In no way it seems.'[25]

From this passage one may conclude that for Plato, eros cannot be an attribute of God.[26] Following Robert Adams, I will summarize Plato's argument (to the effect that eros is not an attribute of God) as follows:[27]

(1) Eros involves desire for its object.
(2) One desires what one needs and does not have.
(3) Thus, the object of eros is what one needs or lacks.
(4) Eros, especially the gods' eros if they have any, will have things beautiful and good as its object.
(5) God cannot lack such things or be needy in respect of goodness and beauty.
(6) Thus, eros is incompatible with deity.

Adams, responding to Plato's argument, argues that eros can be regarded as a model of divine love concerning finite things. He objects that neediness and desire is not essential to eros because its central feature is *valuing*, and desire is only one possible expression of valuing. According to Adams, the essential content of eros would be a kind of liking, admiration, and appreciation, and not necessarily the desire or longing that implies the absence of the desirable/longed-for element. Then eros would be a model of the divine love God expresses when He admires and appreciates the good acts and character of His servants. Thus eros can be attributed to God without attributing any deficiency or neediness to Him.

I think this reply to Plato's objection is inconclusive. The 'value-appraisal' account of love, which seems to be what Adams has in mind,

---

[24] *Symposium* (Cambridge: Cambridge University Press, trans. Benjamin Jowett, 1980), 203 b–d.

[25] Ibid., 202 c5–d7.

[26] Here it should be noted that eros as an emotion or state of mind takes its name from Eros, the person or a demon. In some phrases these two are used indistinguishably.

[27] Robert M. Adams, *Finite and Infinite Goods* (Oxford: Oxford University Press, 1999), p. 133.

is just one sort of eros alongside the union account. I agree with Adams that we can ascribe the value-appraisal type of eros to God, because it is certain that God admires and appreciates good and admirable beings. However, if this was all that there is to God's erotic love, eros as divine love toward human beings reduces to God's appreciation of good human characters and attributes. And consequently the objects of divine eros are only those creatures who instantiate these virtues, for they alone are worthy of God's appreciation and admiration. It seems that God may be pictured as having the desire for being in union with *all* of us, independently of whether or not and to what extent He appreciates us.

Perhaps we can respond to Plato's objection in a different way, by examining the relation between neediness and desire. We can, perhaps, distinguish between two kinds of desire: one may desire something as a means to one's own ends (as a sort of neediness); or one may desire something as a means to the beloved's end.[28] It is in the second sense that we should understand God's desire for union with human beings. If, as Aquinas proposes, the goal of human creation is union with God, then it seems plausible that the benevolent God also desires that we achieve this union with Him. In this case it is not required to ascribe any neediness to God. Although it is true that needs and desires are commonly associated, they are nonetheless not correlative. It is possible to need something (for example, a reform in one's eating habits) without desiring it; but it is also possible to desire something only for the fulfilment of the desire for the good of the object of desire (for example, that those one hears on the news are suffering from famine in a faraway country receive food).

There is another objection to the attribution of eros to God, however, this one based on ideas from Aristotle. In the following section I shall examine this objection.

V. THE ARISTOTELIAN OBJECTION TO DIVINE LOVE FOR US

Aristotle seems to have thought of God as so absorbed in self-contemplation as to have no concern for lesser beings.[29] For Aristotle, God's self-contemplation and self-love is the origin of creation of other

[28] There may also be a type of desire without caring about the end of desire – desire for the sake of desire itself. But it seems eccentric to attribute this to God.

[29] This is a common interpretation of Book Lambda of his *Metaphysics* (Penguin Classics, trans. Hough Lawson-Tancred, 1999), 1074b15–1075a11.

beings. But are we to understand his view as the view that God is so absorbed in self-contemplation and self-love that He has no concern or thought *at all* for lesser beings? Why should divine self-contemplation block God from bestowing attention and love on other creatures? It seems that we have no reason to conclude from the Aristotelian arguments that, while God always engages in self-contemplation, He cannot pay attention to other beings.

The proponent of the Aristotelian argument can improve his objection, however, by saying that what blocks God from desiring union with lesser beings is the possibility of vulnerability, which is associated with love. God would put Himself at risk if He came to love other beings. Vulnerability and risk are associated with the sort of desire the satisfaction of which is to some extent out of the control of the lover. Insofar as the fulfilment of God's desire to be in union with human beings is at least in part within human beings' control, God's having a desire for union with us makes God vulnerable to us. But, the objection continues, God is self-sufficient and sovereign, and thus nothing that matters to God is dependent on anything human beings do. Therefore, it is not possible to attribute eros to God, at least in the sense that eros includes God's desire for union with us.[30]

I think this objection is not yet compelling. It seems plausible that the sovereign God can freely choose to create a creature that has free will, and that the existence of creatures with free will then necessitates their having free choices and actions. Thus far, their free actions seem not to be in contradiction with divine sovereignty; however, if God desires that the free creatures freely choose something that is good for them, then this divine desire does indeed makes God vulnerable to their actions. It seems true then that God, if He desires union with us, is vulnerable to our actions; but this can hardly be considered as incompatible with divine sovereignty; it was His free choice to create free creatures in the first place and at any stage, should He so choose, He could remove their freedom. At least this much seems to be conceded even by many of those who argue against the free-will defence in response to the problem of evil. God's 'vulnerability' could be interpreted as something like God's disappointment in or concern for evildoers.

As a result, for our purpose in this paper it can be concluded that eros – including both God's admiration for good persons and God's

[30] See Stump, *Wandering in Darkness*, pp. 122–3.

desire for union with all of us – is a possible interpretation of an aspect of divine love. It is not then that God loves us solely agapeically. However, as I have noted, according to the union account of eros, the two conditions of personal presence and closeness must be fulfilled by human beings as well. And this is sufficient to block Schellenbeg's argument.

VI. REASSESSMENT OF SCHELLENBERG'S ARGUMENT

In previous sections, I have accepted the idea that divine love has two main aspects, agape and eros; while agape is generous and unmotivated love, and involves benevolence, mercy, and compassion for all creatures regardless of their value, eros is attentive and responsive love, and involves liking, longing, admiration, and desire for union with the beloved. However, union with the beloved requires mutual closeness and personal presence. Mutual closeness and personal presence can be obtained only if the lover and the beloved do what is necessary for this. Then, the responsibility of human beings to respond to God's desire for union must necessitate not only his *non-resistance* to God but also his reflection on divine revelation, repentance, and praying.

Based on these premises, and by distinguishing eros and agape as two conceptions of divine love, I am now in a position to go through the formalization of Schellenberg's premise 1, first replacing love with eros, and then replacing love with agape, showing more formally than I have hitherto how the argument fails in both cases. Schellenberg's premise 1 can be summarized in the following statement:

> S(a): If the loving God of theism exists, then, necessarily, His agapeic love ensures that all persons who are not resisting a relationship with Him will be in a position to participate in developing *relationship with Him*.

Obviously S(a) cannot be demonstrated, because God's love as agape does not necessarily entail any personal relationship whatsoever. Divine love as agape is like rain for all creatures: it falls on the just and unjust alike; it is not responsive to anything in the beloved and does not require any reciprocal relationship. If we accept agape as the interpretation of divine love it can straightforwardly be seen that Schellenberg's argument does not work. Of course, one may claim that divine agape necessitates divine eros in the sense that God, by His generous love toward all of us, bestows upon us what is required on our part to establish a relationship of reciprocal love with Him and to come into union with Him. One

may thus expect that God, the perfectly loving, would provide us with whatever we need for our flourishing and growth toward our aim of union with Him, including the belief that He exists. But my answer to this is twofold. First, I suggest, there is an ambiguity in this sort of expectation. We should distinguish between two kinds of expectation we may have of God's agapeic love: one may expect God to provide human beings with whatever they need to *reach* their ultimate goal; or one may expect God to provide whatever they need *to be in* their ultimate good state. The former expectation – which considers God's *purpose* for the creation of human beings – is the only one that seems tenable, since it is a common claim of theists that God loves perfectly, but that God is also perfectly just and perfectly rational. God's love, then, would have to be calibrated to that degree compatible with the other properties essential to divine perfection. Divine love will not have the consequence the objector expects if that consequence is incompatible with divine rationality and justice. In theological terms, there is a venerable tradition which holds that grace is necessary for one to appreciate the evidence in support of theism, but (and this is an important but), grace is a divine gift of which justice precludes a universal distribution. God could give everybody the belief in His existence (as perhaps He did for angels); but His purpose for our creation is, perhaps, that He expects us to acquire this belief in other ways.

Second, and undeniably, the best sort of love respects the choices of its object. Thus, God, who is perfect in love, would prefer not to coerce us to return His love. Moreover, were God to override our wills in this regard, our love of Him would not be freely chosen. So God would prefer us to come to believe in Him and then reciprocate His love on our own.

We can see, then, that if we think of love in Schellenberg's argument as agape, his argument fails. By substituting eros for divine love in premise S we have:

> S(e): If the loving God of theism exists, then, necessarily, His eros ensures that all persons who are not resisting a relationship with Him will be in a position to participate in a developing *relationship with Him*.

My claim now is that even though it is in itself reasonable (contra Plato and Aristotle) to think of divine love as having an erotic element, S(e) cannot be defended. For whilst it is true that eros requires a reciprocal relationship, it is not the case that every normal human being who is non-resistant can be engaged in divine eros. If my argument is correct,

it is not the case that only nonresistance or even non-culpability is sufficient for being engaged in union with God. One actually has to *do* something in addition, and not everyone who is nonresistant and non-culpable does do this thing – repent, pray, and reflect on revelation.

Therefore, Schellenberg's main premise is defeated.

VII. CONCLUSION

Schellenberg argues that God does not exist because God's love requires a mutual relationship in which He cannot be hidden to those who are nonresistant, and yet it seems that God is hidden to some nonresistant nonbelievers. In this paper this argument is undermined by showing that God's love does not require that He be unhidden. If God's love is agape, God's love does not require Him to be unhidden from anyone. If His love is eros, it will only require Him to be unhidden from those who satisfy certain further conditions.

My objection was presented through an examination of Schellenberg's use of the notion of divine love in his argument. First, he uses the notion of divine love without paying sufficient attention to its different aspects and dimensions. Second, he ignores the point that it is the eros interpretation of divine love that requires reciprocal relationship, and that eros requires certain qualifications on the part of the creatures who are its object. Third, and most importantly, he treats the *capability* to have a relationship to God as a very simple capability that everybody who is capable of having normal human relationships possesses. He pays no attention to the sophisticated, long, and hard path of the mystical journey toward perfection that, in the history of mankind, only a very few men and women have properly traversed.[31]

[31] Thanks for wise counsel and penetrating comments to Muhammad Legenhausen, Tim Mawson, Mahmoud Morvarid, Richard Swinburne and Hamid Vahid.

A CRITICAL EVALUATION OF REA'S RESPONSE TO THE PROBLEM OF DIVINE HIDDENNESS

ROSS PARKER

Charleston Southern University

Abstract. In an important discussion of the problem of hiddenness, Michael Rea briefly presents and defends an argument from divine hiddenness which he thinks encapsulates the problem of divine hiddenness, and then develops a detailed and nuanced response to this argument. Importantly, Rea claims that his response does *not* depend on the commonly held theistic view that God allows hiddenness to secure human goods. In this paper I offer a detailed criticism of Rea's account of what justifies God in allowing divine hiddenness, arguing that Rea's response to the argument from divine hiddenness is unsuccessful.

The problem of divine hiddenness is one of the most significant objections to belief in God. More precisely, the problem of hiddenness can be embodied in an *argument* from hiddenness which concludes that God doesn't, or likely doesn't, exist. While the literature discussing arguments from divine hiddenness has focused on the argument from inculpable nonbelief as defended by J. L. Schellenberg,[1] there have also been other important presentations of arguments from divine hiddenness.[2] The

[1] J. L. Schellenberg, *Divine Hiddenness and Human Reason* (Ithaca, NY: Cornell University Press, 1993); J. L. Schellenberg, 'Divine Hiddenness Justifies Atheism', in *Contemporary Debates in Philosophy of Religion*, ed. Michael L. Peterson and Raymond J. VanArragon (Malden, MA: Wiley-Blackwell, 2004), pp. 30–41; J. L. Schellenberg, *The Wisdom to Doubt: A Justification of Religious Skepticism* (Ithaca, NY: Cornell University Press, 2007).

[2] For examples, see Theodore M. Drange, 'The Argument from Non-Belief', *Religious Studies*, 29, no. 4 (1993), 417–432; Theodore M. Drange, *Nonbelief & Evil: Two Arguments*

common core of these arguments is that there is an incompatibility, or evidential tension, between the existence of a perfectly good God and the existence of cases where significant truths about God lack epistemic support for certain persons.

In a recent discussion of the problem of hiddenness, Michael Rea briefly presents and defends an argument from divine hiddenness which he thinks encapsulates the problem of divine hiddenness, and then develops a detailed and nuanced response to this argument.[3] Importantly, Rea claims that his response to his argument from hiddenness does *not* depend on the commonly held theistic view that God allows hiddenness to secure human goods. Rea proposes that what justifies God in allowing hiddenness is the good of God acting in accord with the divine personality. Further, Rea contends that hiddenness is compatible with God's concern for all people because God has provided a widely and readily accessible way to experience his presence despite divine hiddenness – humans can have *mediated* experiences of God made available via Christian scripture and liturgy.

In this paper I argue that Rea's response to the argument from divine hiddenness is unsuccessful. In order to do this, in §1 I outline Rea's presentation of his version of the argument from divine hiddenness and then in §2 I summarize Rea's response to this argument from hiddenness. My critical evaluation is given in §3-§6. In §3, I argue that Rea's stated understanding of 'divine hiddenness' (or as he prefers 'divine silence') makes it such that the argument he presents fails to embody a plausible problem of hiddenness. Accordingly, I modify his account of divine silence, attempting to stay as true to his stipulative definition as is feasible while providing a definition that grounds a plausible argument from hiddenness. In §4 I critique Rea's appeal to the divine personality to justify divine hiddenness. In §5 I argue against Rea's claim that God provides a way for persons who experience divine silence to encounter him through mediated experiences made available in scripture and

for the Nonexistence of God (Amherst, NY: Prometheus Books, 1998); James A. Keller, 'The Hiddenness of God and the Problem of Evil', *International Journal for Philosophy of Religion*, 37, no. 1 (1995), 13–24; Stephen Maitzen, 'Divine Hiddenness and the Demographics of Theism', *Religious Studies*, 42 (2006), 177–191.

[3] Michael Rea, 'Narrative, Liturgy, and the Hiddenness of God', in *Metaphysics and God: Essays in Honor of Eleonore Stump*, ed. Kevin Timpe (New York: Routledge, 2009), pp. 76–96. References to this work in the remainder of this paper will be made parenthetically in the text.

liturgy. Finally, in §6 I challenge Rea's claim that the availability of mediated experiences of God makes divine silence compatible with God's concern for the well-being of all people.

§1. REA'S ARGUMENT FROM HIDDENNESS

In order to offer a response to the problem of hiddenness, Rea begins by stating and defending an argument from hiddenness. His argument is presented as a *reductio*. If we assume that God exists, then the following mutually inconsistent claims also seem true:

> P1: God has allowed himself to remain hidden from many people.
>
> P2: It would be bad for an omnipotent, omniscient God to remain hidden from anyone.
>
> P3: God, being perfectly good, cannot do anything that is bad.

I refer to this argument as *Rea's inconsistent triad* (RIT). (Unless I indicate otherwise, when I refer to the argument from hiddenness, I am referring to RIT.)

Rea doesn't give any defence of P3, nor does he need to, since neither proponents nor opponents of arguments from hiddenness question the claim that if God exists, he would be perfectly good, nor that God's goodness would entail that he cannot do anything that is bad.

Regarding support for P1, Rea claims that defences of the existence of hiddenness standardly cite something like the truth of the following two facts (p. 76):

> Inconclusive Evidence: For many people, the available a priori and empirical evidence in support of God's existence is inconclusive: one can be fully aware of it and at the same time rationally believe that God does not exist.
>
> Absence of Religious Experience: Many people – believers and unbelievers alike – have never had an experience that seems to them to be a direct experience or awareness of the love or presence of God; and those who do have such experiences have them rarely.

Rea grants that these two facts are true. He argues, however, that the fact that inconclusive evidence and absence of religious experience both obtain is better understood as 'divine silence' rather than 'divine hiddenness' (see pp. 78-81). Rea then stipulates that when he speaks of

divine silence he 'will be speaking simply of the fact that INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE both obtain' (p. 81). Though Rea doesn't update RIT after arguing for the terminological change from 'divine hiddenness' to 'divine silence', it will aid in understanding the argument to update RIT as follows:[4]

> P1*: God has allowed INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE.
>
> P2*: It would be bad for an omnipotent, omniscient God to allow INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE.
>
> P3: God, being perfectly good, cannot do anything that is bad.

(Since I don't find the terminological preference for divine silence rather than divine hiddenness as important as Rea does, I use divine silence and divine hiddenness interchangeably. Unless I indicate otherwise, I use these terms in the way that Rea has stipulated.)

Concerning support for P2*, Rea claims that the basic problem with divine silence is that it seems inconsistent with the following thesis (a thesis that seems entailed by God's perfect goodness):

> DIVINE CONCERN: God strongly desires to promote the well-being of all his rational creatures, both now and in the afterlife. (p. 77)

Rea rightly notes, however, that P2* is true only if divine hiddenness does *not* promote any good the promotion of which would justify God in permitting hiddenness. So the proponent of the argument from hiddenness claims that divine silence is all-things-considered bad for God to allow, because it is inconsistent with DIVINE CONCERN. The opponent of the argument (assuming that he grants the existence of divine silence) must defend a God-justifying good which 'would justify God in permitting whatever bad things come from divine hiddenness' (p. 77).

§2. REA'S RESPONSE TO RIT

Rea's intentions in responding to the RIT can be summarized as follows. Rea notes that most attempts to identify God-justifying goods assume

[4] This change is also supported by Rea's statement that the proponent of the hiddenness argument can 'replace talk about divine hiddenness with talk about the obtaining of INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE' and the argument 'would then proceed with much the same force as the original' (p. 77).

that the good (or goods) which justify hiddenness must be a *human* good. Rea aims to buck this trend, intending 'to defend a response to the problem of divine hiddenness that is consistent with the following claim' (p. 78):

> NO HUMAN GOOD: It is not the case that God permits INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE in order to secure human goods.

Rea will then argue that 'even if NO HUMAN GOOD is true, divine hiddenness does not cast doubt on DIVINE CONCERN' (p. 78). He thinks this is so if the following is true:

> DIVINE SELF-DISCLOSURE: God has provided some widely and readily accessible way of finding him and experiencing his presence despite silence.

Rea will then argue that DIVINE SELF-DISCLOSURE is in fact true.

I think that Rea's response can be stated as a defence of the following three separate theses (these are my statements of Rea's theses, not his):

> T1: The good of divine acts which express God's personality justifies divine silence (making Rea's greater good account compatible with NO HUMAN GOOD).
>
> T2: Even if God does not permit divine silence in order to secure human goods, divine silence is compatible with DIVINE CONCERN so long as DIVINE SELF-DISCLOSURE is true.
>
> T3: DIVINE SELF-DISCLOSURE is true: Biblical narratives and liturgical acts are means by which we might find and experience the presence of God in the midst of divine silence.

§2.1. Defending T1

Rea recognizes that 'it is easy to see why one might think' divine silence is incompatible with DIVINE CONCERN (p. 81). Consider the experience of Mother Teresa, who longed to experience God's presence but (according to her private writings) did not experience God's presence at many times in her life. How could a compassionate God refrain from answering the cries of Mother Teresa and others like her? (pp. 81-2) A human parent would surely draw near to his child and offer words of comfort and affection; shouldn't we expect at least as much from a perfectly loving divine being?

In response, Rea claims that inferring that divine silence is incompatible with DIVINE CONCERN commits an error often committed in complaints about the behaviour of humans; it 'depends on a particular interpretation of behaviour that can in fact be interpreted in any of a number of different ways, depending upon what assumptions we make about the person's beliefs, desires, motives, and overall personality' (p. 82). To make an accurate interpretation of someone's behaviour, one must know things like the person's cultural background, what sort of social norms she is likely to know and respect, her views about what various kinds of behaviours communicate, and so on. And if we need this kind of information to interpret *human behaviour*, then we should be slow to interpret the behaviour of the God of the universe. Even with the witness of Christian Scripture, Rea thinks 'we have precious little by way of clear and reliable information about God's personality and about his general style of interacting with others' (p. 82).

Since we don't know these key facts about God necessary for interpreting his behaviour, divine silence may simply be the result of an expression of God's unique personality and/or God's preferences about how to interact with creatures like us. This possibility becomes Rea's proposed God-justifying good: ' ... divine silence is an outgrowth of the divine personality or of God's preferences about how to interact with creatures like us ... ' (p. 86) Since it is intrinsically good for God to live out God's personality, this intrinsic good justifies God allowing divine silence. Further, the good of God acting according to his personality is not a human good, so Rea's proposed God-justifying good is compatible with NO HUMAN GOOD.

Rea acknowledges that God acting in accord with his personality does result in suffering, which seems to indicate that God is unconcerned for those who experience divine silence. But he defends as plausible the following claim:

> It might be that our suffering in the face of divine silence is unreasonable, due more to our own immaturity or dysfunction than to any lack of kindness on God's part. Perhaps it results from our own untrusting, uncharitable interpretations of divine silence, or from an inappropriate refusal to accept God for who God is and to accept God's preferences about when and in what ways to communicate with us.[5]

[5] Rea's defends this claim at length (pp. 84-85, 87).

§2.2. Defending T2

Rea acknowledges, however, that if 'God had provided no way for us to find him or experience his presence in the midst of his silence', divine silence would be incompatible with DIVINE CONCERN (p. 83). Rea thinks, however, that DIVINE SELF-DISCLOSURE is true.

> DIVINE SELF-DISCLOSURE: God has provided some widely and readily accessible way of finding him and experiencing his presence despite silence.

Rea seems to take it as obvious that T2 is true: *if* there is a 'widely and readily accessible way' of experiencing God despite divine silence, then divine silence would not conflict with DIVINE CONCERN.

§2.3. Defending T3

Rea avers that most discussions of divine hiddenness assume that God has provided some widely and readily accessible way of finding him only if one or both of INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE are false. This position, however, 'ignores ... the possibility of *mediated experiences* of the presence of God through media that are themselves widely and readily available' (p. 88).

Very roughly, Rea's mediated experiences are not direct experiences of the object experienced, but provide a person with much if not all of the propositional and non-propositional information one would get via direct experience. Rea builds on the work of Eleonore Stump to argue that mediated experiences of God are available through the biblical narrative.[6] Stump's account is summarized by Rea as follows. A *second-person experience* is a conscious experience of another person *as a person,* rather than as a mere object. Turning to the biblical narrative, Stump claims that many biblical narratives are *second-person accounts,* by which she means they are narratives that communicate the content of a second person experience. Most importantly for Rea's purposes, Stump claims that second-person accounts can communicate roughly the same kind of information that one gets from a second-person experience of God

[6] Rea cites Eleonore Stump, 'The Problem of Evil: Analytic Philosophy and Narrative', in *Analytic Theology: New Essays in the Philosophy of Theology*, ed. Oliver D. Crisp and Michael C. Rea (Oxford: Oxford University Press, 2009), pp. 251–264; Eleonore Stump, *Wandering in Darkness: Narrative and the Problem of Suffering* (Oxford: Oxford University Press, 2010).

by making the original experience available through the narrative to the reader. Put in his own terms, Rea claims that the biblical second-person accounts *mediate* second-person experiences of God. The upshot for Rea is that if this account is right, through reading the biblical narrative, a person can have a *mediated experience* of God.

Rea also thinks that something similar can be said about liturgical actions. The Eucharist and baptism are commemorative of past events, and can possibly be ways of making present the things commemorated, perhaps offering mediated experience of God. (Rea has less to say about how this works in liturgical actions, and seems less confident that mediated experiences are made available through liturgical actions.)

Rea concludes that if his account of mediated experiences is correct, 'then (given that Biblical narrative and the right sorts of liturgical forms ... are readily available) DIVINE SELF-DISCLOSURE is true.' He then reiterates that 'if DIVINE SELF-DISCLOSURE is true, then ... divine silence is unproblematic' (p. 93).

§3. EVALUATING REA'S PRESENTATION OF THE ARGUMENT FROM DIVINE HIDDENNESS

Below I will offer critical evaluations of T1, T2, and T3. But before doing that, I must evaluate Rea's presentation of the argument from hiddenness. Rea's presentation of the argument from hiddenness is not dependent on any extant version of an argument from hiddenness in the literature, so it's important to consider whether the argument he's presented is plausible. To begin, I overview the support Rea gives for the key premise of the argument. I then argue that Rea's stipulative definition of divine silence results in his presentation of the argument from hiddenness failing to embody a plausible problem of hiddenness for theism.

Recall the updated version of RIT:

> P1*: God has allowed INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE.
>
> P2*: It would be bad for an omnipotent, omniscient God to allow INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE.
>
> P3: God, being perfectly good, cannot do anything that is bad.

According to Rea, the key premise P2* is supported by the apparent inconsistency of INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS

EXPERIENCE with DIVINE CONCERN. Rea gives two lines of reasoning for thinking that divine silence is incompatible with divine concern. I'll refer to these as A and B:

> A: All of the major theistic religions agree that belief in God is vital for our present and future well-being. But a world in which God is hidden is one in which God is doing far less than he could (if he is omnipotent and omniscient) to promote rational theistic belief. Hence, it is one in which God is doing far less than he could to promote our well-being. (p. 77)

> B: Divine hiddenness is a source of suffering in believers, who often feel abandoned, neglected, unloved, or rejected by the being to whom they have devoted their lives and whom they have been taught to regard as their loving heavenly Father. (p. 77)

As an initial comment, B seems unduly narrow. It should include not just believers, but also nonbelievers who desire to believe in God but experience divine silence. Their suffering can be as serious as the suffering of believers who experience divine silence.[7] So this line of support can be made more precise as:

> B*: Divine silence is a source of suffering in believers who feel abandoned, neglected, unloved, or rejected by God, and is a source of suffering for unbelievers because they lack the belief in God they desire.

Having clarified the support for P2*, I now turn to a problem for that premise. The problem with P2* is this: it is possible that INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE both obtain without there being any clear incompatibility or even evidential tension with DIVINE CONCERN. Here's why. INCONCLUSIVE EVIDENCE states that *many people* lack a priori and empirical evidence sufficient to justify belief in God; ABSENCE OF RELIGIOUS EXPERIENCE asserts that *many people* have never had a direct religious experience which could justify belief in God. But the truth of these two facts is compatible with every person in the world having evidence sufficient for knowledge-level justification for belief in God and knowledge-level justification for belief that God is concerned for their well-being and actively involved in their lives.

[7] For a good description of the suffering that some nonbelievers seem to experience because of divine silence, see Schellenberg, *The Wisdom to Doubt*, pp. 227–235.

Consider the following possible scenario. The many people who lack a priori and empirical evidence to justify belief that God exists (i.e., who make INCONCLUSIVE EVIDENCE true) have frequent experiences of the loving presence of God. These people have direct evidence for God's existence via their direct experiences of God, and because of their frequent religious experiences they could not reasonably feel neglected or abandoned by God. Additionally, the many people who fail to have direct experiences of God (i.e., who make ABSENCE OF RELIGIOUS EXPERIENCE true) have knowledge-level justification via their available a priori and empirical evidence. Further, their a priori and empirical evidence provides knowledge-level justification for beliefs about God's love for them and continual involvement in their life.

If it is possible that divine silence is true *while* every person has knowledge-level justification for belief that God exists and that God loves them and is active in their life, then the truth of INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE in itself is not in any way in tension with DIVINE CONCERN. So the problem with Rea's definition of divine silence is that it does not entail any problem of hiddenness. Since this is the case, a response to RIT, with divine silence understood as Rea stipulates, doesn't require the account of the God-justifying good that Rea develops, or any account of a God-justifying good. Rather, one can simply deny P2*.

If Rea's presentation of the argument from hiddenness is to embody the problem of hiddenness in a plausible way, his account of divine silence needs to be modified such that it's plausible to think that silence is *pro tanto* bad because of its tension with DIVINE CONCERN. Seeking to modify Rea's account as little as possible, I propose the following revised account of divine silence:

> INCONCLUSIVE EVIDENCE & LACK OF RELIGIOUS EXPERIENCE*: For many people, their available a priori and empirical evidence in support of God's existence is inconclusive (they can be fully aware of it and at the same time rationally believe that God does not exist), *and* they have never had an experience that seems to them to be a direct experience or awareness of the love or presence of God.

If divine silence is understood as INCONCLUSIVE EVIDENCE & LACK OF RELIGIOUS EXPERIENCE*, then there is a plausible case to be made for the inconsistency of divine silence and DIVINE CONCERN. Additionally, while it's the case that this modified account of divine silence asserts

more than Rea's initial account, it would be somewhat surprising to think that Rea's INCONCLUSIVE EVIDENCE and ABSENCE OF RELIGIOUS EXPERIENCE both obtain but deny that INCONCLUSIVE EVIDENCE & LACK OF RELIGIOUS EXPERIENCE* obtains. So hereafter, to make plausible Rea's argument from divine hiddenness, divine silence will be understood as INCONCLUSIVE EVIDENCE & LACK OF RELIGIOUS EXPERIENCE*. Because of this, RIT must be modified as:

> P1#: God has allowed INCONCLUSIVE EVIDENCE & ABSENCE OF RELIGIOUS EXPERIENCE*.
>
> P2#: It would be bad for an omniscient, omnipotent God to allow INCONCLUSIVE EVIDENCE & ABSENCE OF RELIGIOUS EXPERIENCE*.
>
> P3: God, being perfectly good, cannot do anything that is bad.

§4. EVALUATING T1

Recall that Rea intends for his account of what justifies hiddenness to be compatible with NO HUMAN GOOD, which says that *it is not the case that God permits divine silence in order to secure human goods.* I summarized Rea's account of the non-human good that justifies divine silence as

> T1: The good of divine acts which express God's personality justifies divine silence (making Rea's greater good account compatible with NO HUMAN GOOD).

Rea's argument for T1 has two components. First, Rea attempts to undercut the inference from divine silence to the conclusion that God is unloving or unconcerned for humanity. Second, Rea puts forward the hypothesis that divine silence is a result of divine action that expresses God's personality, and since God expressing his personality is intrinsically good, this good might justify divine silence without any reference to human goods. Here I offer brief criticism of both steps (§4.1 and §4.2) and then point out a further problem for Rea's account of what justifies God in allowing divine silence (§4.3).

§4.1. Interpreting divine silence

Rea's case against inferring from divine hiddenness that God is unloving or unconcerned can be summarized as follows (hereafter the INTERPRETATION SCEPTICISM ARGUMENT):

IS1: In order to understand someone's behaviour, one must know substantial information about the person's beliefs, desires, motives, dispositions, and overall personality.

IS2: But when it comes to God, we know very little about these factors.

IS3: So we cannot understand why God is silent.

IS4: So we shouldn't interpret divine silence as expressing lack of love or concern.

Rea motivates IS1 with a couple of examples and explanations of our inability to understand human behaviour. Here's one of his examples, which I will refer to as 'Hallway Silence':

> Hallway Silence: A senior member of your department doesn't greet you in the hallway. Is he offended by you? Does he think you're beneath him? Is he depressed and having a bad day? Or is that *just him,* a little preoccupied and not really noticing his surroundings? (p. 82)

Rea is correct that on the evidence that I have here, any of the options explaining my colleague's behaviour are viable. So there is something that I cannot understand about my colleague's behaviour. But there are still reasonable interpretations of my colleague's behaviour that I can make in this case. We can have a fair bit of knowledge upon which to build interpretations of a person's actions by knowing that a person is a normally functioning human being. Human beings are motivated by love, fear, physical pleasure, moral considerations, and a host of other psychological factors, and they have practical reasoning skills that enable them to act on their motivational states. So concerning HALLWAY SILENCE, for example, I seem perfectly justified in interpreting my colleague's behaviour as *not* intended to ascertain how my day is going.

Further, the reasons for not interpreting the silence in HALLWAY SILENCE as an expression of unconcern or disdain go beyond mere lack of information about the other person. It's not simply that I don't have enough information about my colleague's personality, beliefs, or cultural expectations; rather, it's reasonable to withhold interpreting his behaviour in part because we can think of multiple *plausible* explanations for his behaviour (as shown in the options Rea gives in the case). We know that people are sometimes preoccupied when walking down the hall. We know that some people are introverted, or that some people aren't comfortable talking with people in passing, etc. Contrast this with a case like the following:

> Non-Intervening Colleague: You meet a colleague in the hall, look him in the eye, and say 'I have an important question I need to ask you'. He ignores you and looks the other way. Just then, you feel severe pain in your chest, and collapse to the floor. You look at your colleague again and ask 'please call for help!' but again he ignores you and looks the other way.

Assume that you know this person has a conversational knowledge of English, and you have good reason to believe that he hears you and sees you. Perhaps you shouldn't attribute this behaviour to a *particular* malicious intention. But no matter how little I know about this person, I am justified in believing that his behaviour is incompatible with his concern for me. This is so because, whatever the details of a person's background and personality, actions that are hurtful or undermine another's well being are not loving, unless we have a reason to think that the action (though undermining my well being) is necessary for a good purpose.

It's true that substantial information about the person's beliefs, desires, motives, dispositions, and overall personality is necessary for a *full* understanding of a person's behaviour. But we can make reasonable interpretations providing a *partial* understanding of a person's behaviour even if we lack some information about these characteristics of the person. The parallel of this point with interpreting divine hiddenness is important. The proponent of the argument from hiddenness claims that divine silence is incompatible with DIVINE CONCERN. She can recognize that to give a *specific* interpretation of a person's action we may need to know a good bit about a person, but also think that it's reasonable to believe an action is incompatible with promoting the well-being of another without knowing much about the person. And I take it that the proponent of the hiddenness argument sees divine silence as more like the case of NON-INTERVENING COLLEAGUE rather than the case of HALLWAY SILENCE.

Perhaps what does the heavy lifting in Rea's argument is IS2. Rea stresses that we know very little about the factors of God's personality that would allow us to understand his interaction with us. Consider the following statements:

> Even granting the complete reliability and transparency of Biblical testimony about God, we have precious little by way of clear and reliable information about God's personality and about his general 'style' of

> interacting with others; and to ask about God's 'culture' or about what sorts of social norms God would likely recognize and respect seems to border on the overly anthropomorphic. God is as alien and 'wholly other' from as it is possible for another person to be. (p. 83)
>
> ... the most enigmatic, eccentric, and complicated people we might ever encounter in literature or in real life are, by comparison with God, utterly familiar and mundane. (p. 85)

While what Rea says here is in part correct, the case seems overstated. Consider first that Rea acknowledges that we *do* know an important truth about God's personality; namely, DIVINE CONCERN:

> DIVINE CONCERN: God strongly desires to promote the well-being of all his rational creatures, both now and in the afterlife.

If DIVINE CONCERN is true, then we have a solid understanding of one of God's fundamental desires and motives. So when we attempt to interpret divine silence (or other of God's actions) it's not the case that we know little of importance concerning God's beliefs and desires. Rather we interpret divine silence in light of the truth that if there is a God, he has perfect love for all humans and desires their ultimate well-being. Even if there is much we don't know about the beliefs, desires, motives, and dispositions of a perfect Divine being, if divine silence seems to clearly not promote human well-being, then we have some reason to interpret it as incompatible with DIVINE CONCERN.[8]

Rea thinks that the considerations he gives against interpreting the actions of persons support a strong agnosticism concerning the interpretation of divine silence:

[8] Additionally, Rea overstates his case when he claims that even with Christian Scripture, there is little we can know about the way God would interact with humans. The Christian Scriptures give many affirmations from which one would be able to know something about God's beliefs, desires, and overall personality. To take just one example, Scripture commends that God's personality and interaction with humans can be understood analogously with what a human parent would do in dealing with a child. Consider this statement of Jesus from the Sermon on the Mount: 'Ask, and it will be given to you; seek, and you will find; knock, and it will be opened to you. For everyone who asks receives, and the one who seeks finds, and to the one who knocks it will be opened. Or which one of you, if his son asks him for bread, will give him a stone? Or if he asks for a fish, will give him a serpent? If you then, who are evil, know how to give good gifts to your children, how much more will your Father who is in heaven give good things to those who ask him!' (Matthew 7:7-11 ESV)

> Silence is an interpretable kind of behaviour; and, as with any other person, God's behaviour doesn't wear its interpretation on its sleeve – it can be understood only in the light of substantial background information. To be sure, divine silence could be an indication of divine rejection or lack of concern. But *that interpretation is entirely optional, given our evidence*. (p. 83, my emphasis)

This is too strong a conclusion. It doesn't follow from the truth that the behaviour of persons is interpretable and our lack of robust knowledge of the divine personality that there is no good reason to interpret divine silence as incompatible with DIVINE CONCERN. Rea's considerations concerning interpreting agential behaviour should remind us to be careful and circumspect in inferring God's intentions in allowing divine silence. But it seems clear that we know enough about what a perfect being would do such that divine silence is a *pro tanto* bad state of affairs, such that if one wants to deny P2# (as Rea is attempting to do), then one needs to provide a reason to think that divine silence promotes a good 'the promotion of which would justify God in permitting whatever bad things come from divine hiddenness' (p. 77). As Richard Swinburne (an influential *opponent* of the argument from hiddenness) states

> All knowledge is good, but especially knowledge of deep truths about the Universe and our place in it, who or what is the source of our being, and more truths about our duties and the good actions beyond duty which we can do ... As our creator, God will seek to interact with us. He will want us to feel his presence, to tell him things and ask him to do things; and he will want to tell us things ... and to do good things with us, to cooperate with us in producing further goods.[9]

Swinburne thinks it's clear that 'God will, for these and other reasons ... want us to know that he exists'.[10] This is why theists like Swinburne give accounts of plausible goods which could justify God allowing divine silence – they recognize that God would not withhold evidence of his existence unless he had a good reason to do so. So I now turn to consider Rea's account of what justifies God in allowing hiddenness.

[9] Richard Swinburne, *Providence and the Problem of Evil* (Oxford: Clarendon Press, 1998), pp. 112–113.

[10] Ibid., p. 113.

§4.2. Rea's Defence of his God-justifying Good

Recall that Rea states that he intends to 'defend a response to the problem of divine hiddenness that is consistent with [NO HUMAN GOOD]' (p. 78). This can be understood in a couple of different ways. One might think that Rea intends to provide an account of a God-justifying good and argue that this account is plausibly true. Rea does not, however, give any argument to support the conclusion that his God-justifying good is *true*, or even reasonable to believe. Rather, the structure of his defence for his God-justifying good is given in the following passage:

> To be sure, divine silence could be an indication of divine rejection or lack of concern. But that interpretation is entirely optional, given our evidence. Divine silence might instead simply be a reflection of the fact that God prefers to communicate with us and to draw us into his presence in ways other than ones that would render [divine silence] false. It might just be a reflection of God's personality, so to speak. (p. 83)

Another point concerning Rea's defence of his God-justifying good is important to note. Rea nowhere gives an account of what divine belief, desire, or motive might motivate God acting in such a way that divine silence occurs. His appeal is to an *unknown* aspect of the divine personality.

Here's why I think this is inadequate. Proponents of arguments from hiddenness have provided reasons to think that P2# is true. They argue that God allowing divine silence is bad. If the interpretation of divine silence as bad was 'entirely optional, given our evidence', then putting forward a God-justifying good that's merely possible may be sufficient for thinking that P2# is false. But I've argued that though divine silence is an interpretable behaviour, this doesn't undermine the case that divine silence is a *pro tanto* bad state of affairs. Rea needs to give a reason to think that divine silence, though *pro tanto* bad, is not *all-things-considered* bad by developing a plausible account of why God allows divine silence. Rea's claim – that possibly God has an unknown part of his personality that is such that acting from his personality leads to divine silence – doesn't give a reason to think that divine silence is not all-things-considered bad. In other words, since reasons have been given for the claim that divine silence is incompatible with DIVINE CONCERN (see A and B* above), and because Rea's INTERPRETATION SCEPTICISM ARGUMENT is not successful, the onus is on the opponent of the argument from hiddenness to give some reason to think that there is a good that justifies God in allowing

hiddenness. Simply appealing to the possibility of an unknown divine motivation fails to count as a defence of a God-justifying good.[11]

§4.3. Another Problem for Rea's God-justifying Good

I have argued that Rea's case against inferring that God is unloving/unconcerned from divine silence and his positing of God's actions expressing the divine personality as his account of the God-justifying good both have problems. However, even if my account here is unsuccessful, there is another problem for T1. Rea admits that the good of God acting from the divine personality justifies hiddenness *only if* DIVINE SELF-DISCLOSURE is also true:

> DIVINE SELF-DISCLOSURE: God has provided some widely and readily accessible way of finding him and experiencing his presence despite silence.

But if DIVINE SELF-DISCLOSURE is a necessary condition for justifying divine silence, then it seems inaccurate to claim that what justifies divine silence is the good of God living out the divine personality. Rather, for Rea what justifies divine silence is the good of God living out the divine personality *in conjunction with* the availability of means to find God and experience him despite divine silence. So even if my criticism of T1 is unsuccessful, Rea's account of what justifies God in allowing divine silence is only as good as his case for T2 and T3. I accordingly evaluate these two theses in the next two sections.

§5. AGAINST T3

Recall T3:

> T3: DIVINE SELF-DISCLOSURE is true: Biblical narratives and liturgical acts are means by which we might find and experience the presence of God in the midst of divine silence.

Two preliminary points. First, I limit my discussion of Rea's case for DIVINE SELF-DISCLOSURE being met through mediated experiences of

[11] There are clear parallels between Rea's appeal to an unknown component of the divine personality and the sceptical theist's response to the argument from evil. The issue of sceptical theism has been discussed extensively in the recent literature, and I can't do justice to that discussion here.

God made available through biblical narratives. My criticisms, however, could be applied, *mutatis mutandis*, to what Rea says about mediated experiences via liturgy.

Second, in order for Rea's account of mediated experiences of God to be plausible, Stump's position that second person accounts of God given in certain biblical narratives make available second-person experiences of God must be true. Rea gives no argument for the truth of this position; he simply assumes it in developing his account of how DIVINE SELF-DISCLOSURE is true. To evaluate Stump's position, however, is beyond the scope of this paper, so I will grant this part of Rea's account.

The fundamental problem for T3 is that people who experience divine silence seem to not be in the epistemic position necessary to have mediated experiences of God via biblical narratives. Concerning what is required to have a mediated experience of God through biblical narratives, Rea states:

> A certain kind of 'seeing as' is a necessary condition for experiencing the presence of another person as such: one has to consciously regard the other *as a person* ... Likewise, then, one would expect that a similar sort of 'seeing as' would be involved in having mediated experiences of the presence of another person. Thus, for example, if one were to read a story about Fred's second-person experiences of Wilma while failing to see Wilma as a (real) person ... the experiences conveyed by the narrative would be different and, in that event, there would be no reason to think that the narrative would in any sense be mediating Wilma's presence. If this is right, then *whether Biblical narratives mediate the presence of God will depend importantly upon whether one takes those narratives to be reporting real experiences of God.* (p. 91, emphasis mine)

So in order to have a mediated experience of God through the biblical narrative, one must take those narratives to be reporting real experiences of God. The most plausible way to understand this requirement is that one would have to believe that the biblical narrative reports real experiences of God. But I argued in §3 that in order for Rea's argument from hiddenness to be plausible, divine silence must be understood as INCONCLUSIVE EVIDENCE & LACK OF RELIGIOUS EXPERIENCE*. Those who experience divine silence are such that 'their available a priori and empirical evidence in support of God's existence is inconclusive (they can be fully aware of it and at the same time rationally believe that God does not exist), and they have never had an experience that seems to them

to be a direct experience or awareness of the love or presence of God'. So people who experience divine silence do not have epistemic reason to justify belief that God exists, which means they don't have reason to take the biblical narratives as reporting real experiences of God. Since this is so, those who experience divine silence should not believe that the biblical narratives are reporting real experiences of God. And this, according to Rea's own account, would keep them from being able to have a mediated experience of God via the biblical narratives.

Rea has responded to this criticism in correspondence as follows.[12] Consider a person who lacks evidence for God's existence *because* he accepts an argument from hiddenness against God's existence; i.e., if he did not accept the argument from hiddenness he would justifiably believe in God. This person then comes to believe that biblical narratives (if true) could provide mediated experiences of God. This could provide a defeater for the reasons he had for accepting the argument from hiddenness, such that he would be open to experiencing the presence of God via the biblical narrative.

This is an interesting response, but it fails to undercut the general point I'm trying to make. First, this reply only shows that those who (i) experience divine silence because of hiddenness and then (ii) come to see that it's possibly true that one can have a mediated experience of God have reason to believe the biblical narratives are true, and thus be open to mediated experiences of God. But it seems implausible to think that most people who experience divine silence fail to believe in God because of the argument from hiddenness. So for those who lack a priori, empirical, or experiential evidence for God's existence for reasons other than an argument from hiddenness, they still lack a reason to believe that the biblical narratives are true and thus fail to be in an epistemic position to have a mediated experience of God via scripture. Second, for the person whose reason for unbelief is defeated by considering the possibility of mediated experiences of God, it is not true that this person has the possibility of experiencing God despite silence (which is the claim of T3). Rather, this person now has the reason that was keeping him from justifiably believing in God defeated. So he no longer experiences divine silence – he has a justified belief in God.

Another possible response that Rea could make is as follows: while the person who experiences divine silence *epistemically* ought not

[12] Personal e-mail correspondence, June 26, 2013.

believe that the biblical narratives are reporting real experiences of God, he could have *pragmatic* reasons to believe that the biblical narratives are reporting real experiences of God. In response, believing that the biblical narratives are reporting real experiences of God for pragmatic reasons when one has insufficient evidence to justify this belief seems psychologically implausible. Most accounts of belief agree that belief is *synchronically* involuntary. As Swinburne states, 'In general, a person cannot choose what to believe there and then. Believing is something that happens to someone, not something that he does.' [13] If this is the case, one cannot *choose* to believe that the biblical narratives are reporting real experiences of God at time *t* on the basis of pragmatic reasons. So in order to believe that biblical narratives are reporting real experiences of God, one would have to act in such a way that over time (a) one convinces oneself that the biblical narratives are true when one doesn't have good evidence for this, or (b) one comes to have evidence that supports this proposition. If (a), then it seems like the person has done something irrational, and it seems problematic that a perfectly loving God would require people who experience divine silence to be irrational in order to be able to experience God despite divine silence. If (b), then by having evidence for believing that the biblical narratives are true the person would have evidence that supports belief in God, which would mean that she would no longer experience divine silence. So I conclude that even if Rea's account of mediated experiences of God is correct, only those for whom God is *not* hidden will be able to have mediated experiences of God, which means that we should deny T3.

§6. AGAINST T2

In the previous sections I argued against T1 and T3. In this section I argue that even if we grant these two theses (which again, I have argued against), there is a fundamental problem for T2:

> T2: Even if God does not permit divine silence in order to secure human goods, divine silence is compatible with DIVINE CONCERN so long as DIVINE SELF-DISCLOSURE is true.

The problem is that a gap exists between DIVINE CONCERN and DIVINE SELF-DISCLOSURE. DIVINE CONCERN claims that 'God strongly desires to

[13] Richard Swinburne, *Faith and Reason*, 2nd ed. (Oxford: Oxford University Press, 2005), p. 24.

promote the well-being of *all* of his rational creatures'. But DIVINE SELF-DISCLOSURE only claims that God has made a way to experience him that is 'widely and readily accessible'.

In order for DIVINE CONCERN to be compatible with divine silence, the way made available to experience God needs to make God accessible to *all* those who experience divine silence. Rea seems to acknowledge this, stating that

> If we have been entirely cut off from God's presence, God then has done or permitted something that is both devastatingly harmful to us and totally out of our control, and it is much harder to make plausible the suggestion that God has taken reasonable steps to be compassionate towards us in the midst of our suffering. (p. 88)

If God desires the well-being of *all* his rational creatures (including those who experience divine silence), then it seems that he would provide *all* who experience divine silence access to mediated experiences of God. Rea's position is that the widely and readily accessible way of finding God and experiencing his presence despite silence is mediated experiences of God made available by Christian Scripture and liturgy. It seems implausible, however, to think that *all people* for whom God is hidden currently have access to Christian Scripture and liturgy (to say nothing of people in the past). It is an uncontroversial fact that *many* people have no access to Christian Scripture or the Christian church. Further, many instances of divine hiddenness cited by proponents of arguments from hiddenness are people who do not have access to these means of indirect experience of God. In Schellenberg's initial defence of the argument from hiddenness, for example, he puts forward as examples of divine hiddenness 'individuals – primarily from non-Western cultures – who have never so much as entertained the proposition "God exists"'.[14]

So even if people who experience divine silence can find God and have mediated experiences of his presence via Scripture and liturgy (contra my case above), divine silence is not thereby shown to be compatible with DIVINE CONCERN if there are people who experience divine silence and who don't have access to Christian Scripture or liturgy. For Rea to defend the compatibility of divine silence and DIVINE CONCERN, the means of mediated experiences of God must be available for *all* who experience hiddenness; otherwise there will be individuals who 'have been entirely cut off from God's presence', and for whom God 'has done or permitted

[14] Schellenberg, *Divine Hiddenness and Human Reason*, p. 58.

something that is both devastatingly harmful ... and totally out of [their] control'. Rea has not, however, given any defence of the claim that all who experience hiddenness have access to Christian scripture and liturgy; further, to defend this claim would be a difficult task in light of the experience of many who have no access to Christian Scripture and liturgy and seem to be exemplar cases of divine hiddenness.

§7. CONCLUSION

I have attempted to show that the response to the problem of hiddenness given by Rea is unsuccessful for a number of reasons. This is not to say that the argument from hiddenness Rea develops gives evidential support for the conclusion that a perfectly good God doesn't exist. There are a number of other theistic responses to the problem of hiddenness on offer.[15] But theists looking for an adequate response to the problem of hiddenness will need to look to one or more of these other accounts, rather than the one critiqued here.[16]

[15] To begin looking at theistic responses to the problem of hiddenness, here are some suggestions (this list is not meant to be exhaustive): Thomas V. Morris, 'The Hidden God', *Philosophical Topics*, XVI, no. 2 (1988), 5–21; Michael Murray, 'Coercion and the Hiddenness of God', *American Philosophical Quarterly*, 30, no. 1 (1993), 27–38; Daniel Howard-Snyder, 'The Argument from Divine Hiddenness', *Canadian Journal of Philosophy*, 26, no. 3 (1996), 433–453; Swinburne, *Providence and the Problem of Evil*; Daniel Howard-Snyder and Paul Moser, eds., *Divine Hiddenness: New Essays* (Cambridge: Cambridge University Press, 2002); C. Stephen Evans, 'Can God Be Hidden and Evident at the Same Time? Some Kierkegaardian Reflections', *Faith and Philosophy*, 23, no. 3 (2006), 241–253; Ted Poston and Trent Dougherty, 'Divine Hiddenness and the Nature of Belief', *Religious Studies*, 43 (2007), 183–198; Andrew Cullison, 'Two Solutions to the Problem of Divine Hiddenness', *American Philosophical Quarterly*, 47, no. 2 (2010), 119–134; Travis Dumsday, 'Divine Hiddenness as Divine Mercy', *Religious Studies*, 48 (2012), 183–198; Travis Dumsday, 'Divine Hiddenness and Creaturely Resentment', *International Journal for Philosophy of Religion*, 77 (2012), 41–51; Justin McBrayer and Philip Swenson, 'Scepticism About the Argument from Divine Hiddenness', *Religious Studies*, 48 (2012), 129–150.

[16] I would like to thank Michael Rea for helpful comments on an earlier draft of this paper.

INHERENCE AND DENOMINATION IN THE TRINITY

PAUL THOM

The University of Sydney

Abstract. The present paper describes an 'ontological square' mapping possible ways of combining the domains and converse domains of the relations of inherence and denomination. In the context of expounding and extending medieval appropriations of elements drawn from Aristotle's *Categories* for theological purposes, the paper uses this square to examine different ways of defining Substance-terms and Accident-terms by reference to inherence and denomination within the constraints imposed by the doctrine of the Trinity. These different approaches are related to particular texts of thinkers including Bonaventure and Gilbert of Poitiers.

Given that the doctrine of the Trinity teaches a mystery that demands faith on the part of believers, it can be surprising that in the hands of certain medieval philosophers the Trinity becomes a logical puzzle, a puzzle how to reconcile the doctrine with certain tenets of Aristotelian philosophy. However, this development did not necessitate abandoning the requisite attitude of faith; it just meant that the faith of those who pursued this type of investigation was illuminated (or alternatively, encumbered) by philosophical theory.

I. BACKGROUND

A key component of the philosophical background to the logical puzzle of the Trinity can be found in the account of the various types of terms given in Aristotle's *Categories*. A privileged class of terms (let us say *per*

se terms) divides into substance-terms and accident-terms. Beyond that class there are denominatives. The terms so classified are not construed as totally unmediated by language and thought; rather, they are construed as already having a certain conceptual content. Thus the terms *man* and *the one approaching* are different terms, and different types of term, even though in some context they stand for the same being.

I take it that terms are picked out by abstract or concrete nominal expressions. But I make no assumptions about the ontological status of terms – whether they are mental or linguistic items, or whether they are objective entities. I do assume that various relations hold among terms, including relations of inherence and denomination. But when we say things like '*Colour* inheres in *body*' or '*Coloured* is denominated from *colour*', the truth of what we say is not dependent on the actual existence of colours, coloured things or bodies.

Different configurations of inherence and denomination, discernible in the text of the *Categories*, give us necessary conditions for being a substance-term, and also for being an accident-term.[1]

Three necessary conditions for a term *A* being a substance-term are:
(1) *A* does not inhere in any term as subject. This is a necessary condition of being a substance term, but it is not sufficient because it also applies to terms like *rational* that differentiate one species from another but do not themselves characterise a substance.[2]
(2) A second necessary condition is that *A* has accident-terms inhering in it.[3] Aristotle calls this feature the most characteristic mark of substances.
(3) A third necessary condition is not stated explicitly but may be conjectured as assumed in the text. It is noticeable that a differentia term such as *rational* is denominated from an abstract term, in this case the quality-term *rationality*, whereas we do not find any abstract terms in the *Categories* from which substance-terms are denominated. While *man* is a substance-term, no term *humanity* is mentioned. It seems then that a third necessary condition of substance-terms is that they are not denominated from any term.
(4) Finally, it is clear that a substance term does not denominate anything.
It seems then that we have three classes of terms as shown in Figure 1.

[1] Paul Thom, *The Logic of the Trinity: Augustine to Ockham* (New York: Fordham University Press, 2012), pp. 13ff.

[2] Aristotle, *Categories* 5, 2b21.

[3] Aristotle, *Categories* 5, 4a10.

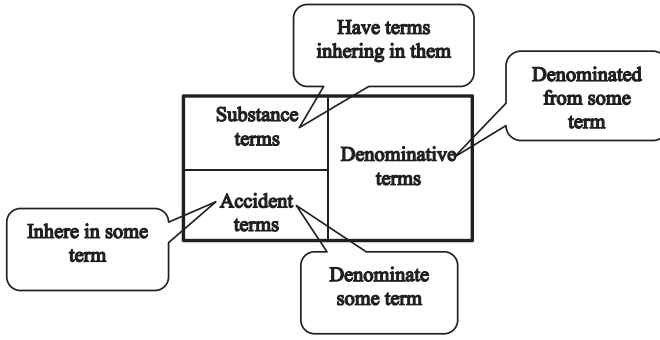

Figure 1. Types of term in the *Categories*

In distinguishing these three types of term, we have appealed to *four* relative states in which a term *A* can find itself: (1) *A* inheres in some term, (2) *A* is inhered in by some term, (3) *A* denominates some term, and (4) *A* is denominated by some term. Each of these states can be present or absent in a given term. So there are in principle not 3, but 16 types of term – each one of which is characterised by the presence or absence of each one of the four relative states. These types of term are shown in Figure 2, together with the locations of the types of term recognised in the *Categories*.

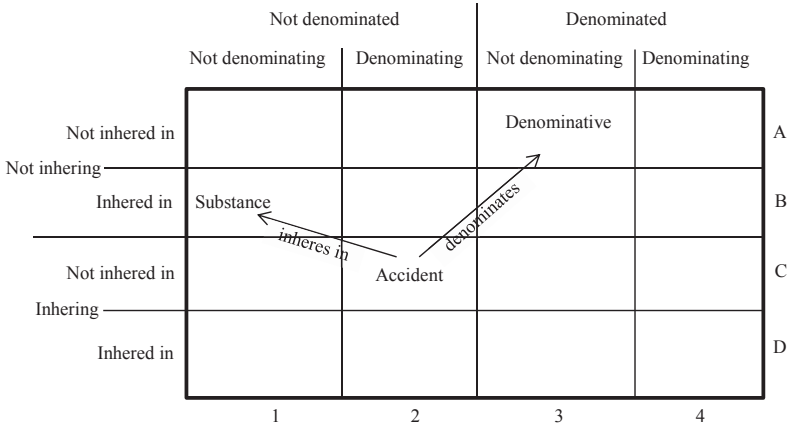

Figure 2. The ontology of the *Categories*

Other ontologies based on notions of denomination and inherence ought to be able similarly to locate the types of term they recognise.

The class of substance-terms in the *Categories* can be characterised by the following configuration of inherence and denomination: these terms are inhered in, they don't inhere, they are not denominated, and they don't denominate.

Accident-terms share one of these characteristics: they are not denominated. We do not find second-order terms like *Rationality-ness* in the *Categories*. But in respect of the other three characteristics, what substance-terms possess accident-terms lack. Thus accident-terms inhere in some term, they are not inhered in by any term, and they denominate.

Before we proceed it will be useful to make some observations about the ways in which terms figuring in the ontology of the *Categories* are inter-related by inherence and denomination.

(1) Every term enters into either a relation of inherence or a relation of denomination.
(2) What denominates something inheres in something.
(3) Denomination is irreflexive.
(4) What is denominated is not inhered in.

These observations can be spelt out as follows. Firstly, if we consider only substance-terms, accident-terms and denominatives, it is apparent that substance-terms are inhered in, accident-terms inhere, denominatives are denominated; so in each case the term in question stands towards something in a relation of inherence or a relation of denomination or the converses of these relations.

Secondly, confining consideration again to substance-terms, accident-terms and denominatives, we observe that only accident-terms denominate, and only accident-terms inhere; thus what denominates something inheres in something.

On observation 3, it might be thought that the irreflexivity of denomination is clear from the fact that when Aristotle introduces this relation he notes a linguistic difference between the denominating term and its denominative.[4] But on the other hand, it might be argued that this linguistic difference is accidental to the *ontological* relation between denominating and denominated, and that consequently it cannot be a sound basis for concluding that the relevant ontological denomination is irreflexive. There is, however, another, ontological, argument that can

[4] *Categories* 1, 1a12.

be appealed to. Aristotle conceives of the denominated term as being *derived* from the denominating term; and under this aspect it seems that the relation must be irreflexive – at least, this is so if a term cannot be derived from itself.

Regarding observation 4, it can be argued that (a) only denominatives are denominated (and denominatives are not substance-terms), and (b) only substance-terms are inhered in (to be inhered in is the mark of substance). Against this reasoning it might be argued that sometimes an accident-term inheres in another accident-term. For example, my pallor may change in its relational accidents, by spreading more widely over the surface of my face. If accident-terms can have accident-terms inhering in them, then it is not true that only substance-terms are inhered in. However, it remains true that only substance-terms or accident-terms are inhered in, and that denominatives are neither substance-terms nor accident-terms.

These observations correspond to various possible states of our table. Observation 1 would be reflected in a table that has no terms in cell 1A. Observation 2 would be reflected in a table where there are no terms in cells 2A, 2B, 4A or 4B. Observation 3 does not require that there will be no terms in column 4 – but that any term in that column (a term that both denominates and is denominated) must denominate and be denominated in respect of two different terms. Observation 4 requires that there will be no terms in cells 3A, 3C, 4A and 4C. Thus there is some overlap between what is excluded by observations 2-4.

II. THE TERM *GOD*

In applying this kind of framework to the divine realm, the constraints imposed by the faith (in a broad sense that includes the teachings of the Church Fathers and Councils) are that among terms applying to the Godhead there be only one substance-term, no accident-terms, and that the Godhead exhibits simplicity. More specifically, *God* must be a substance-term, nothing else applying to the Godhead can be a substance-term, *God* must not inhere as an accident in anything, and nothing can inhere in *God* as an accident. The notion of divine simplicity can be interpreted in more than one way.[5] On one interpretation, it excludes

[5] Paul Thom, 'Shades of Simplicity', in Anselm Ramelow (ed.), *God* (Munich: Philosophia Verlag, 2013), pp. 323-340.

any denomination within the Godhead. On another reading, it allows for denomination provided that this does not entail any dependency.[6] Other constraints include logical consistency and interpretive sense and comprehensiveness. In some developments the account will seek to give an ontological grounding for all the relevant predications. The aims of the present investigation are more modest: I will confine attention to logical coherence. I will examine three different ways of introducing *God* into our framework.

III. A *GOD* UNRELATED BY INHERENCE AND DENOMINATION

The *Categories* scheme as it stands is ill-suited to representing a theology in which there is a substance-term in which no accident-terms inhere. If accidents are defined as what inheres, such a theology requires that there be a substance-term in row A or row B. If the term *God* is not denominated from anything and nothing is denominated from *God* then this term must be in cell 1A. The term *God* thus does not stand to anything in a relation of inherence or a relation of denomination or in the converses of these relations. Such a term clearly satisfies the requirement of divine simplicity: at least it excludes any complexity that might be due to the relations of inherence and denomination.

The possibility of terms of type 1A is familiar to Aristotelians, though not because such terms form part of the *Categories* landscape. Negative terms like *Non-man* are of this type. To allow for such a term being a substance-term, the definition of substance-terms will have to be revised, dropping the requirement that a substance-term has something inhering in it.

The ontology of Aristotle's *Categories* gives us an initial set of links (consequences and incompatibilities) between being a substance-term and various configurations of inherence and denomination. The task of adapting this ontology to a set of theological requirements is similar to that of adapting a theory to accommodate recalcitrant data. One has to place an ordering on the initial set of links, which will determine which links can be abandoned ahead of others. The link with the weakest strength in a context requiring that there be a substance-term in which

[6] Paul Thom, *The Logic of the Trinity: Augustine to Ockham* (New York: Fordham University Press, 2012).

nothing inheres is the link between being a substance-term and being inhered in. If that link is broken, a substance-term can be redefined as one that (1) does not inhere, (2) is not denominated and (3) does not denominate. The first condition excludes rows C and D. The second condition excludes columns 3 and 4. The third condition excludes columns 2 and 4. The remaining cells are 1A and 1B. Created substance-terms belong in 1B, divine substance-terms in 1A.

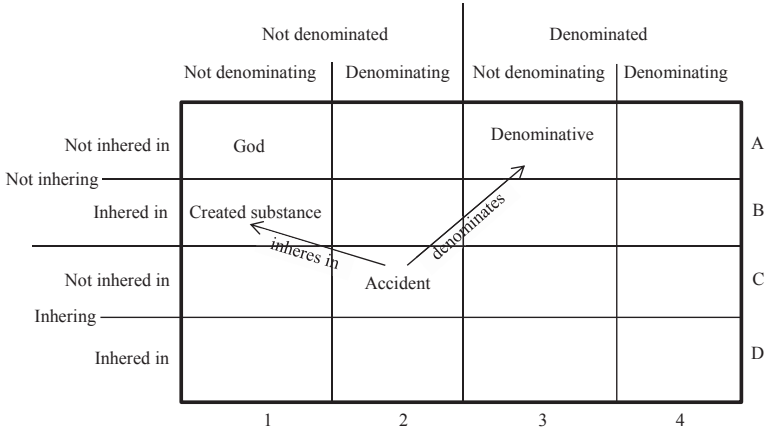

Figure 3. The *Categories* ontology plus divine substance

An accident-term can still be defined as one that inheres, is not inhered in, is not denominated, and denominates.

Redefining substance-terms in this way brings us into conflict with the first of our observations about the *Categories* ontology, viz. that every term enters into either the domain or the converse domain of either inherence or denomination. The term *God* now appears as an exception to this observation. But this should not worry a hypothetical theologian who wants to fit a term *God* into an ontology derived from the *Categories*. Aristotle's theory of categories gives an account of substance-terms that admit of temporally alternating accident-terms – an account in terms of inherence, denomination, and a couple of other relations. It is not surprising that all of the terms that are recognised in his account enter into these relations in one way or another. Once the account is extended so as to allow for other types of term, this restriction can no longer be expected to hold.

IV. A SELF-RELATED *GOD*

Let us now suppose that we want to make *God* a term which does stand in one or other of our relations, but only in a reflexive way. This account again satisfies the demand for divine simplicity, at least in the sense that it excludes any multiplication of substance-terms in the Godhead. Some thinkers, among them Bonaventure, have thought of *God* in this way, claiming that *God* is a self-denominating term:

> Since *quod est* as well as *quo est* are found among lower things, by reason of which we have both concrete and abstract signification (as when we say *man* and *humanity*), this is also our understanding in the divine realm, although we do not there understand these two to be different. Accordingly, we signify abstractly by the name *Deity* and concretely by the name *God*. And thereby we give Him a name by which we signify *quo est* (and this is the essence), as well as *quod est* (and this is the substance).[7]

The terms *God* and *divinity*, denominated and denominating, have the same signification. These terms now appear in cell 4A.

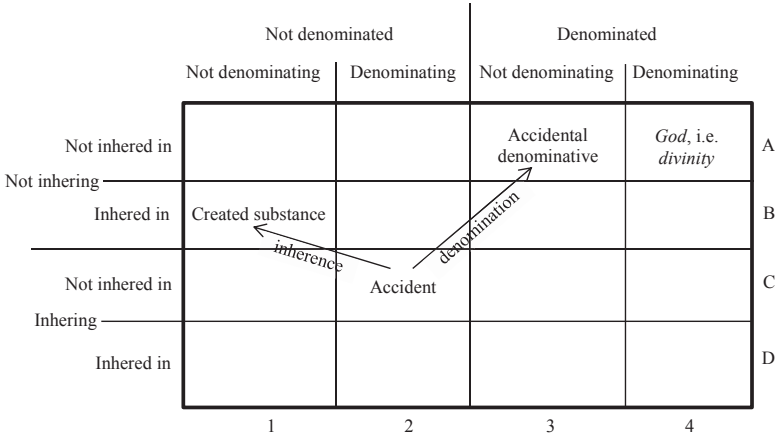

Figure 4. A self-related *God*

[7] Bonaventure, Sent. I d.23 a.1 q.3. ... cum in *communi* in inferioribus inveniatur *quod est et quo est*, ratione cuius significatur in concretione et in abstractione, ut dicatur *homo et humanitas*: sic in divinis intelligimus, quamvis non *intelligamus* in differentia illa duo. Ideo et in abstractione *significamus* per hoc nomen *deitas*, et in concretione per hoc nomen *Deus*. Et ideo imposuimus ei nomen, quo significaretur ipsum *quo est*, et hoc est essentia; et ipsum *quod est*, et hoc est substantia ....

In order for *God* to be counted as a substance-term in the same sense as non-divine substance-terms, the definition of a substance-term will have to be altered. This can be done by redefining a substance-term as satisfying the following three conditions. (1) It does not inhere, (2) it is inhered in iff it is not denominated, (3) it is denominated iff it denominates. The first condition rules out rows C and D. The second rules out 1A, 1C, 2A, 2C, 3B, 4B, 3D, 4D. The third condition rules out columns 2 and 3. The remainder is 1B and 4A. 1B is the profile of a created substance-term, 4A of a divine substance-term.

The definition, however, may be too wide because it includes everything falling into cell 4A, including terms that are denominated by something other than themselves. The difficulty can be met by replacing the third condition by the specification that the term is not denominated by anything other than itself.

A difficulty with this account arises from our earlier observation that there are no self-denominating terms in the *Categories*. Bonaventure is aware of this difficulty. In order to meet it he makes a distinction within the class of denominatives:

> To compare one thing to another as informing it or denominating it, is not thereby to posit a diversity or distinction between them; for, Deity is compared in this way to God. To compare things as a principle and that of which it is the principle, is a different sort [of comparison]: this *is* to import a distinction.[8]

In other words, denomination has to be irreflexive if it is understood as treating the denominating term as a principle from which the denominated term flows. But if we remove this element from the notion of denomination, there is no objection to a term's being denominated from itself. The definition of accidents is unaffected.

A line of objection could be raised against Bonaventure's idea of isolating an element in the notion of denomination which renders the relation irreflexive, and by abstraction generating from the relation a transform of it which lacks that element. First of all, in order to be sure that the abstracted relation is not itself irreflexive, one would need to be sure that one had removed *all* elements that would render it irreflexive.

[8] Bonaventure, *Sent*. 1.25.2: Quando aliquid comparatur ad aliud ut informans sive denominans, non ponitur propter hoc diversitas sive distinctio unius ad alterum; sic enim comparatur deitas ad Deum. Alio modo comparatur alterum sicut principium ad principiatum; et tunc de necessitate importatur distinctio.

A sound argument leading to the desired conclusion requires a universal premise. Bonaventure does not address this point; and it is not at all clear how he, or anyone, could do so. Secondly, even if by this kind of process one could arrive at a suitably abstracted description of a non-reflexive relation, that would not yet prove that the description was satisfiable. One would still need a consistency (satisfiability) proof for the description.

In any case, Bonaventure's specification of the content of the relation of denomination is too specific. Rather than being primarily a relation of what is principled to its principle, the relation of being denominated from what denominates it seems better described as that between something derivative and that from which it is derived. Given this, it's clear that the conclusion at which Bonaventure wishes to arrive can indeed be arrived at. Since mutual inter-derivability is a limiting case of derivability, and since denomination shorn of its linguistic marker is reducible to derivability, mutual denomination may be accepted as a limiting case of denomination. Derivability is a process of transformation in which each step is in accordance with a predetermined set of rules. If there is a single instance in which a rule is reversible, then the derivability relation is non-reflexive. That doesn't mean that all sub-relations of derivability are non-reflexive; but Bonaventure clearly thought that the specific kind of derivability that connects a denominated with a denominating term is non-reflexive.

Bonaventure and his contemporaries were happy to allow that a term for a created substance, such as *man*, had a corresponding abstract term *humanity* – even though such abstract terms do not figure in the *Categories*. A full account of his views would have to take that into account. However, my present objective is not to analyse Bonaventure's views, but to illustrate how our framework accounts for a term *God* as self-related by denomination.

V. A *GOD* RELATED TO OTHERS

Let us turn to a third way of construing the term *God* – as entering into relations of inherence or denomination with other terms. Gilbert of Poitiers, who died in 1154, is a famous exponent of such a view. Gilbert holds that *God* and *divinity* have different significations and are connected by a non-reflexive relation of denomination. On the face of it, such a view conflicts with divine simplicity. But Gilbert has a response to this charge, which we will see presently.

Gilbert radically overhauls the conceptual scheme of the *Categories*. Instead of speaking of the relation of denomination he uses the opposition between *quo est* and *quod est* (or alternatively, that between a subsistence and a subsistent):

> So he says BEING, i.e. the subsistence which is in a subsistent, IS DIFFERENT FROM WHAT-IS, i.e. the subsistent in which the subsistence is: for example, corporeality and body, humanity and man.[9]

Both types of term can stand for substances.

> For not only a subsistent but also a subsistence is called 'substance' in that both stand under accidents, though for different reasons.[10]

Gilbert also has a novel way of conceiving of the Aristotelian relation of inherence. He understands inherence to be the relative product of two relations, the first of which relates an accident-term to a subsistence, while the second relates that subsistence to a subsistent. He calls the former relation accompaniment, the second being the relation of denomination. Thus, instead of saying that *colour* inheres in *body*, Gilbert wants to say that *colour* accompanies *corporeality* which denominates *body*.
He explains the relation of accompaniment as follows:

> And we say that colour and line follow corporality, rather than corporality following colour and line. For they don't cause corporality; it causes them. It is the very being of a body; but they *accompany* it in a body. So it exists at first, after which it is a body (for it is indeed their substance). They are accidents first of corporality, and through it of body. For they stand under it – both corporality (which they accompany) and body (in which they inhere).[11]

---

[9] Gilbert of Poitiers, *Expositio in Boecium librum De Bonorum Ebdomade* §35. Ait ergo: DIUERSUM EST ESSE i.e. subsistentia, que est in subsistente, ET ID QUOD EST i.e. subsistens in quo est subsistentia: ut corporalitas et corpus, humanitas et homo.

[10] Gilbert of Poitiers, *Expositio in Boecium librum primum De Trinitate* 4 §99. Non enim subsistens tantum sed etiam subsistentia appellatur 'substantia' eo quod utraque accidentibus, diuersis tamen rationibus, substant.

[11] Gilbert of Poitiers, *Expositio in Boecium librum primum De Trinitate* 4 §26. Et dicimus quod non corporalitas colorem aut lineam sed color et linea corporalitatem secuntur. Non enim hec corporalitatis sed horum corporalitas causa est. Qua ratione illa corporis est esse: hec uero in eodem corpore illi adsunt. Ideo primum illa, deinde quod ea corpus est, uera ratione est horum substancia: hec uero primum corporalitatis et per eam corporis Taccidencia. His enim uere substat et corporalitas, cui assunt, et corpus cui insunt.

The language Gilbert uses here suggests that he views the relation of accompaniment as holding only when that which is accompanied provides a metaphysical substratum for that which accompanies.

As a result of his innovations, Gilbert is committed, even in conceptualising the created world, to terms of types 2B and 3B, neither of which is found in the *Categories*. 2B are not denominated but do denominate. 3B are denominated and do not denominate. 3B is how substance-terms appear in a language that admits abstract terms derived from them. 2B is how those abstract terms appear in such a language.

Gilbert treats *divinity* and *God* as being of types 2A and 3A. This treatment is perfectly analogous to his treatment of the abstracts of created substance-terms and the created substance-terms themselves. However, the created substance-term is inhered in while the divine one is not, and the abstract created substance-term is accompanied by something while the abstract divine one is.

> For that by which He is – the essence (which in Greek is called *ousia*) – cannot be non-simple. Nor can something other in it accompany the same essence, by which it is. For God would not be simple if His essence were established from several essences, or if forms accompanied the same in it, of which either God Himself truly was (or His essence were with reason said to be) the subjected matter.[12]

Even though *God* is not the same as *Divinity*, there is nothing other than *Divinity* by which *God* is, and *Divinity* is only because *God* is from it.[13] This is Gilbert's reinterpretation of what divine simplicity requires.

We already know of terms that satisfy the profile of 3A, namely accidental denominatives. As for 2A, a term of this type must not inhere in anything nor have anything inhering in it, not be denominated by anything but must denominate something. In other words, such a term does not enter into inherence relations in any way, and enters into denomination relations only by denominating not by being denominated. There doesn't appear to be anything ruling out such a state

[12] Gilbert of Poitiers, *Expositio in Boecium librum primum De Trinitate* 2.37: Neque enim ea, qua ipse est, essencia – que Grece usia dicitur – potest esse non simplex. Neque in eo eidem essencie adesse aliud aliquid potest quo ipse sit. Non enim Deus simplex esset si uel eius essencia constaret ex multis essenciis uel eidem adessent forme in illo quarum uel ipse Deus uere esset uel eius essencia ratione diceretur 'subiecta materia'.

[13] Gilbert of Poitiers, *Expositio in Boecium librum primum De Trinitate* I,2,89. Non enim est a diuinitate aliud quo Deus sit. Nec est unde diuinitas ipsa sit nisi quod ea Deus est.

of affairs: the fact that *A* denominates *B* doesn't require that something should denominate *A* (we know this from the case of accident-terms). Nor of itself does it require that *A* should inhere in anything (it seems that any abstract term denominates, but privative abstract terms, such as *blindness*, do not inhere). Nor does *A*'s denominating something entail that something should inhere in *A* (again, in the case of accident-terms we know that they denominate, but do not have anything inhering in them – at least, this is true of accident-terms of the highest order).

On Gilbert's reckoning *God* turns out to be a denominative just like created accidental denominates such as *brave* (but different from created substance-terms). At first sight this seems unacceptable. However, even within the slender resources to which we have limited ourselves, we can draw a distinction between two types of denominative. Terms like *brave* are denominated from accident-terms, the term *God* is denominated from the non-accidental term *Divinity*. Or, confining ourselves to what can be expressed solely through inherence and denomination: *brave* is denominated from a 2C-term not from a 3A-term; with *God* it is the other way round.

Another apparent difficulty with Gilbert's idea of *God* is that this term is denominated but is not inhered in – contrary, so it seems, to one of our observations about the *Categories*, viz. the observation that in the *Categories* what denominates something inheres in something. But actually, the objection is misconceived. To say that what denominates something inheres in something is not to say that what denominates *A* inheres in *A*. And in the *Categories* what is denominated is never what is inhered in. The denomination relation connects what is abstract with what is concrete. The inherence relation connects what is abstract with a substance-term. So, even in the *Categories* it is not the case that every denominated term is inhered in. So, from the point of view of the *Categories* there is nothing untoward about *God* being denominated but not inhered in.

A third difficulty concerns Gilbert's construal of *divinity*. Contrary to what we observe in the *Categories*, this term denominates but does not inhere. Its failure to inhere follows, in Gilbert's view, from the fact that it is not an accident-term. However, the coextension of denominating and inhering terms that we observe in the *Categories* does not appear to be have a principled basis. If we assume that what inheres must be an accident-term, we can argue that what inheres must denominate. For, an accident-term must denominate: all accident-terms are abstract, and

thus can be made concrete, in which case they denominate some term. So there is a reason why what inheres must denominate. But there doesn't seem to be any reason why, conversely, what denominates must inhere. Rather, any term – even a term beyond the limits of what is considered in the *Categories* (e.g. terms for negations, privations, compounds, Platonic Forms) – so long as it is abstract, must denominate. But terms falling outside the range of what is considered in the *Categories* cannot inhere, because they are not accident-terms.

Gilbert's innovations necessitate a revision of the definition of substance-terms. A substance-term satisfies the following two conditions (which are the first and third of the three conditions defining substance-terms in a Bonaventure-style scheme). (1) It does not inhere, (2) it is denominated iff it denominates. The first condition excludes rows C and D. The second condition excludes columns 1 and 4. What remains are the cells 2A, 3A, 2B, 3B. These are the profiles respectively of the abstract and concrete expressions for divine substance-terms, and the abstract and concrete expressions for created substance-terms.

VI. WHAT DOES THIS TYPE OF ANALYSIS SHOW?

The type of analysis used in this paper may have a use in reconstructing the conceptual schemes that structure certain strands of medieval theological thought. It may even help in understanding the changes that are evident when one of these schemes gets supplanted by another.

Acknowledgment. This paper was delivered to the conference *Analytic Theology: Faith, Knowledge, and the Trinity* held in Prague on 19-20 September 2013. My warm thanks go to the organisers of the conference for their hospitality, and to those who participated in the discussion of my paper for their helpful comments. The conference was generously funded by the John Templeton Foundation and this publication was made possible through the support of a grant from the John Templeton Foundation, grant #15571 ('Analytic Theology'). The opinions expressed are those of the author and do not necessarily reflect the views of the John Templeton Foundation.

BIBLIOGRAPHY

Aristotle, *Categoriae et liber de interpretatione* ed. L. Minio-Paluello (Oxford: Clarendon Press, 1949)

Bonaventure, *Commentaria in quatuor libros sententiarum Magistri Petri Lombardi* (The Fathers of the Collegii S. Bonaventura (eds.), Florence: Quaracchi, 1882)

Gilbert of Poitiers, *Expositio in Boecium librum De Bonorum Ebdomade*, in Nikolaus M. Häring (ed.), *The Commentaries on Boethius by Gilbert of Poitiers* (Toronto: Pontifical Institute of Mediaeval Studies, 1966), 183-232.

Gilbert of Poitiers, *Expositio in Boecium librum primum De Trinitate*, in Nikolaus M. Häring (ed.), *The Commentaries on Boethius by Gilbert of Poitiers* (Toronto: Pontifical Institute of Mediaeval Studies, 1966), pp. 62-160

Thom, Paul, 'Shades of Simplicity', in Anselm Ramelow (ed.), *God* (Munich: Philosophia Verlag, 2013), pp. 323-340

Thom, Paul, *The Logic of the Trinity: Augustine to Ockham* (New York: Fordham University Press, 2012)

HILDEGARD OF BINGEN: A FEMINIST ONTOLOGY

JANE DURAN

University of California at Santa Barbara

Abstract. Two major lines of argument support the notion that Hildegard of Bingen's metaphysics is peculiarly gynocentric. Contra the standard commentary on her work, the focus is not on the notion of *viriditas;* rather, the first line of argument presents a specific delineation of her ontology, demonstrating that it is a graded hierarchy of beings, many of which present feminine aspects of the divine, and all of which establish the metaphysical notion of interpenetrability. The second line of argument specifically contrasts her thought to that of Aquinas and Meister Eckhart, noting areas of similarity and difference. It is concluded that the visionary origins of Hildegard's work may have to some extent precluded our understanding of it, and that her work merits consideration not only philosophically and theologically but from the standpoint of its early presentation of a gynocentric worldview.

Hildegard of Bingen possesses an odd place in the current project of resurrecting the work of women philosophers. Unlike other women thinkers who have traditionally been accorded at least some small place in the pantheon – Hypatia, Anne Conway, Damaris Cudworth Masham – Hildegard has historically been sufficiently well-known so that the contention that she was lost does not stand up. On the other hand, precisely because Hildegard already has a place, at least in theological thinking,[1] there is enough commentary on her work that the feminist scholar who chooses Hildegard as a subject may find herself in the somewhat unusual position, vis-a-vis a woman philosopher, of going against the grain.

Linda Lopez McAlister, in her introduction to a recent work on women in philosophy, remarked that Hildegard was the only woman

[1] The main work is *Hildegardis Scivias*, Adelgundis Fuhrkotter, ed., *Corpus Christianorum Continuatio Medievalis*, Vol. LXIII (Turnhaut: Pontificat, l978). Due to the renascence of interest in mystical thinkers, Hildegard's work is now widely cited.

thinker categorizable as a philosopher of whose existence she had been aware as a student.[2] Matthew Fox has done extensive work on Hildegard, and she is widely perceived as standing in the same mystical tradition as Meister Eckhart.[3] But Hildegard's visions are informative for us, I shall argue, in a number of ways. Fox, Newman and others have argued that Hildegard brings a peculiarly female dimension to her thinking.[4] The concept of *viriditas*, for example, seems consonant with an ecologically-oriented wholeness, and strikes a chord with regard to current work by feminist theorists on deep ecology and ecofeminism.[5] The challenge, I believe, is to provide an account of Hildegard's ontology which is specific enough to achieve philosophical adequacy and simultaneously sufficiently developed from a feminist point of view to be able to sustain the assertion that portions of Hildegard's work are gynocentric.

Both Conway and Masham have received critiques which focus on their construction of ontologies or metaphysical viewpoints that might be deemed to be feminist.[6] Hildegard's work requires a similar explication. In this paper I shall be concerned to develop her ontology in such a way that delineation of its mystical origins helps us understand the nature of its gynocentrism. I shall first be concerned with some particulars of Hildegard's metaphysics.

I.

An area of difficulty with presentation of Hildegard's ontology is immediately apparent upon reading a work like the *Scivias*, for example,

---

[2] Linda Lopez McAlister, 'Some Remarks on Exploring the History of Women in Philosophy', in *Hypatia*, Vol. 4, No. 1 (Spring l989), 1.

[3] See, for example, Matthew Fox, *Illuminations of Hildegard of Bingen* (Santa Fe, NM: Bear & Co., l985). Fox specifically makes the comparison with Eckhart on pp. 16-l7; Fox and Bruce Hozeski (eds.) composed an English edition of *Scivias* through Bear & Co., l986.

[4] Helen John, 'Hildegard of Bingen', unpublished ms. (Trinity College, Washington, DC, l990).

[5] See *Hypatia*, Vol. 6, No. l (l99l): this special issue contains approximately a dozen articles on ecofeminism.

[6] See the relevant articles from *Hypatia*, op. cit.; Jane Duran, 'Anne Viscountess Conway: A Seventeenth Century Rationalist', pp. 64-79, and Lois Frankel, 'Damaris Cudworth Masham: A Seventeenth Century Rationalist Philosopher', pp. 80-90. With regard to Conway's work, a superbly edited version is available: Anne Conway, *Principles of the Most Ancient and Modern Philosophy*, Peter Loptson, ed. (The Hague: Martinus Nijhoff, l982).

even in its English excerpts and translations. The problem is that this work (and many others attributed to Hildegard) is the product of a series of visions, and the visionary influence presents us with assertions which, taken as a whole, are either inconsistent or incoherent.[7] Even a commentator like Fox, who wishes to employ Hildegard's visions for his new theology, presents us with brief excerpts from Scivias, the letters and De Operatione Dei which give seemingly contradictory accounts of the structure of the universe.[8] Nevertheless, some points of comparison seem clear, once the nature of the visions becomes apparent.

Hildegard presents us with a dualistic metaphysics, but with a metaphysical apparatus that allows for complex ontological relationships between various gradations of being. Just as Anne Conway posited a being midway between God and humankind (Conway refers to this being as 'Adam Kadmon', following cabbalistic tradition),[9] Hildegard seems to hold an ontological view which allows for the triune deity, the standard sorts of Christian spiritual entities such as angels, and then intermediate entities, such as the Virtues, which are above human and animalkind in the ontological hierarchy but which are still essentially spiritual entities. In various of Hildegard's visions these beings are personified as animals or entities with more human features but still metaphysically separate from the lower realm of existence. This same ontological separation seems also to come out of some of Hildegard's visions with regard to manifestations of the Trinity. Uhlein's translation of parts of *De Operatione* describes one vision as

> ... Love appearing in a human form, the Love of our Heavenly Father ... Love – in the power of the everlasting Godhead, full of exquisite beauty, marvellous in its mysterious gifts.[10]

In other words, Hildegard's visions allowed her to create epicycles on her ontology, as it were, which show a degree of union and interpenetration between the spiritual and material worlds which would be rare even for a visionary Christian. It is this aspect of Hildegard's

---

[7] An interesting contrast is that between the cosmic structure given in *Scivias* (Latin ed., op. cit.), p. 42 and *Welt und Mensch* (German translation of Hildegard's *De Operatione Dei*) (Salzburg, l965), p. 35.

[8] Fox makes this comparison, which is taken from the primary sources referred to in the previous fn., on pp. 35 and 39 of *Illuminations*.

[9] Duran, 'Conway', p. 66.

[10] Gabriele Uhein, *Meditations with Hildegard of Bingen* (Santa Fe, NM: Bear & Co., l982), pp. 40, 37.

work which has usually been labelled 'feminist' by the commentators, and it ties into the other aspect of her work that has received extensive attention, the notion of *viriditas*. This notion of a life-force – this time a manifestation of the spirit through that which is material – is, according to Fox, '... God's freshness that humans receive in their spiritual and physical [capacities]'.[11] Thus Hildegard sees creation as having the capacity to manifest physically that which is spiritual. Although this does not give her any unusual sort of monistic view (which we find, for example, in some of the philosophers of the early modern period), it apparently yields for her a view which allows the spiritual and material to intermingle, even if the nature of the intermingling is not clear.[12] The notion of a dualism which allows for penetration or admixture is a metaphysically sophisticated concept, preceding occasionalist and other Cartesian notions, and allowing for the gradations of spirituality that I have mentioned above. Uhlein, working from De Operatione, translates Hildegard as writing:

> O Holy Spirit, you are the mighty way in which everything that is in the heavens, on the earth, and under the earth, is penetrated with connectedness, penetrated with relatedness.[13]

The contention that viriditas or a similar manifestation of this commingling is peculiarly gynocentric is drawn out further by Newman, who in her commentary focuses on the extent to which some of the personifications of the Virtues mentioned earlier are also highly feminized figures. As Newman notes, this was characteristic of certain strands of medieval thought, but Hildegard perhaps utilized these personified entities in new ways. Illumination Six in Fox's collection of excerpts, largely from the Scivias, focuses on the vision that sees the universe as an organic part of the deity; the Virtues, Newman notes, are personae such that '... wherever [they] appear, we will find the Platonizing cosmology that captivates 12th century thinkers: the divine ideas, eternal in the mind of God and bodied forth in creatures; ...'[14]

Thus a cursory impression of Hildegard of Bingen's ontology is that it is a graded dualism which allows for some interaction between spirit and

[11] Fox, *Illuminations*, p. 32.

[12] Again, Conway seems to hold such a monism.

[13] Uhlein, *Meditations*, p. 4l.

[14] Barbara Newman, *Sister of Wisdom: St. Hildegard's Theology of the Feminine* (Berkeley, CA: University of California Press, l987), p. 44.

matter. The notions of connectedness and relatedness, which we would clearly deem gynocentric and which fly in the face of the more atomistic thinking of others from the same period,[15] seem even more powerful when we remember that Hildegard wrote a number of empirically-based pieces, and that at least one, her medical work, does not lay claim to being of visionary origin. In the next section a development of Hildegard's more overtly poetic and musical work will follow, coupled with an analysis of what can be gleaned from her metaphysics once these more literary pieces are examined.

II.

Hozeski's analysis of Hildegard's morality play, *Ordo Virtutum*, is of great help in a further refinement of her notions of commitment in terms of entities. The strong counter that could be made to any attempt to come to grips with her hierarchy is that the results of such an account, as mentioned above, are metaphorical at best and unstable – or unintended – at worst, but a look at *Virtutum* seems to bear out some of the niceties of her ontology.

Throughout *Scivias*, certain entities recur: the *Virtutum* repeats this theme. Although Hozeski's main concern is to establish Hildegard as chronologically prior to the canonically-verified liturgical play tradition, his analysis of the structure of the *personae* is revealing.

> The 'Ordo' of the Virtues begins with a list of 'Personae'. The first of the 'Personae' are the Patriarchs and Prophets, followed by the Souls and by Humility, a Queen. There are three Choruses. The first consists of six members: Knowledge of God, Charity, Hope, Chastity, Heavenly Love and Discretion. The second Chorus has five members: Faith, Contempt of the World, Discipline, Patience and Modesty. The third also has five: Fear of God, Obedience, Innocence, Mercy and Victory. The Devil is listed separately and finally in the 'Personae'.[16]

This listing of the Virtues would not be so striking were it not consonant with a good deal of the *Scivias*, and were the Virtues not placed on a par, metaphysically, with several other entities. Hozeski

[15] Fox, for one, specifically compares her to Aquinas in this regard.

[16] Bruce W. Hozeski, 'Hildegard of Bingen's *Ordo Virtutum*: the Earliest Discovered Liturgical Morality Play', in *American Benedictine Review*, Vol. 26, No. 3 (September l975), 251-259. (This citation p. 252)

notes, for example, that '... a ... Soul ... passes through ... the conquest of the Devil by the coming of Christ with His virtues ...';[17] several of the scenes have the Soul in conversation with the individual Virtues, the Devil and so forth. Although Hozeski mentions Hildegard's fondness for certain kinds of visual images, which are also found throughout the *Scivias*, her poetry and some musical works, the visual images are only very generally related to each other, and vary in metaphorical intensity. He notes, for example, that the finale of *Virtutum* has '... images of love, of darkness and brightness, death and growth ... with those of dryness and water, and fruitful flowers, plants and trees'.[18] These images are consistent with many of the vaguer results of the *Scivias*, as found in the illustrations (and as treated by Fox) – more direct constructions seem to figure in any activity involving notions of individual salvation or the activity of the soul.[19]

Barbara Grant, writing in *Signs*, seems to come to a similar conclusion about the importance of a sort of ontological intermediary (or set of intermediaries) for Hildegard. Noting the Gnostic origins of some of what Hildegard appeared to use as subject matter, both for the *Virtutum* and her *Symphonia*, Grant writes:

> Wisdom makes regular appearances in Hildegard's visions and speaks as the Wisdom of God. We have recently learned a great deal about the Sophia-Sapientia figure from the recovery in this century of Gnostic documents ... The consensus is that the female counter-part of the godhead in Gnosticism is the Wisdom of Judaism ... she both precedes and is instrumental in the creation of the world .... [S]he has existed from eternity, and ... she assisted and harmonized the creation of the various parts of the physical universe.[20]

Song 59, as translated by Grant, is specifically addressed to Wisdom and provides us with a precise visual interpretation of her – although this is far from unusual for a vision-driven work, the consistency of Hildegard's assertions on this score merits some degree of confidence.

---

[17] Ibid.

[18] Ibid., p. 259.

[19] Hozeski (p. 256) also cites Hardison as noting that the liturgical plays of the type created by Hildegard were constrained by a 'sacramental psychology'. (Hozeski citing O.B. Hardison, *Christian Rite and Christian Drama in the Middle Ages* (Baltimore: n.p., l965), p. 289).

[20] Barbara L. Grant, 'Five Liturgical Songs by Hildegard von Bingen (l098-ll79)', in *Signs*, Vol. 5, No. 3 (Spring l980), 557-567. (This citation pp. 56l-562.)

Given the nature of Hildegard's mysticism and the concomitant philosophical ramifications, the feminist theorist who wishes to pursue the importance of notions such as connectedness and penetration for Hildegard's work must make some comparisons with other metaphysical commitments of the time. Fox wants to make an easy comparison with Aquinas; although this can be done, it is not always to the point since it is clear that the origins of Aquinas's conceptualizations are vastly more empirical. The more obvious point of comparison is Meister Eckhart. Fox does himself make this comparison, and this contrast – along with some of the more standard Aristotelian material from Aquinas and others – provides a base from which we may expand on the significance of Hildegard's gynocentrism.

III.

The commentators on Hildegard seem to differ in precisely what it is that makes her work of philosophical importance. That is, if we accept the fact that Hildegard, as a visionary, is working from a different conceptual standpoint than many even of the lesser-known medieval thinkers, it must be inquired precisely what it is about Hildegard's work that makes it philosophical and comparable to other philosophical – rather than, say, poetic – work of the period.

Brunn and Epiney-Burgard answer this question implicitly by arguing that Hildegard's work is 'didactic'. They note that:

> In general, Hildegard's visions have been considered didactic rather than ecstatic; that is to say, through them, she transmitted her knowledge in allegorical form. Her visions are not hallucinations, even when received in wakefulness, but they allow her to penetrate 'with the eyes and ears of the inner man' into the realm of the 'spiritual senses', where what one sees and hears is received in a supranatural light, 'shadow of the Living Light', and sometimes, very exceptionally, she penetrates that light itself.[21]

In other words, one might want to make the case for the elucidation of Hildegard's metaphysics and even a nascent epistemology – largely on the basis of what the goal of the visionary presentation is. We may think of Hildegard herself as not only motivated toward but actually

[21] Emily Zun Brunn and Georgette Epiney-Burgard, *Women Mystics in Medieval Europe* (New York: Paragon House, l989), pp. 8-9.

enthusiastic about the presentation of her metaphysics; this is presumably what Brunn and Epiney-Burgard mean when they say that here visions are '... not hallucinations'.

Given the foregoing (a view which is, of course, consonant with Fox's, as well as Hozeski's), one might be surprised to find that the intersections between Hildegard's work and those of other medieval philosophers are rather more obvious than might initially be thought. Aquinas's pedantic and semi-Aristotelian style is in direct contrast, of course, but not all of the philosophical points fail to be congruent. Consider the following passage from Aquinas' 'On the Composition of Essence and Existence in Created Substances', from *De ente et essentia*:

> Hence there is no composition of matter and form of any kind in the soul or an intelligence. On this score, then, essence cannot be said to be in them as it is in corporeal substances. But composition of form and existence is there ... Now whatever pertains to anything is either caused by the principles of its nature, as man's ability to laugh, or it stems from some extrinsic principle, as the sun's luminosity in the atmosphere ... Therefore everything whose existence is something other than its nature must derive its existence from another.[22]

The preceding passage is exemplary because it contains less of the type of distinction frequently occurring in Aquinas and seldom occurring in Hildegard – that between primary and secondary substances, accidents and so forth. Rather, the language lends itself to the characterization of the deity which occurs throughout Hildegard's works, and which is cited here in the excision from Uhlein's translation of *De Operatione* and Newman's commentary on Hildegard's cosmology: the deity as the source of all things, and form as influencing both the individual soul and (on Hildegard's view), the 'divine ideas', or Virtues.

Hildegard's notion that an examination of creation paves the way for further knowledge of the deity is also found in, for example, the work of Peter Lombard: as Wippel and Wolter have it, an intriguing passage in Lombard's *Libri Sententiarum* reads:

> ... it has been shown here just how some likeness at least of the Trinity is found in creatures. Without interior revelation or a revelation of this

[22] Thomas Aquinas, originally from *De ente et essentia*, Roland-Gosselin, ed. (Paris: J. Vrin, l948); translated by A.B. Wolter and excerpted in *Medieval Philosophy: From St. Augustine to Nicholas of Cusa*, John F. Wippel and Allan B. Wolter, eds. (New York: Free Press, l969), pp. 330-332.

> doctrine, however, adequate knowledge of the Trinity neither was nor can be obtained from a contemplation of creatures. That is why ancient philosophers could only see the Trinitarian truth in a haze, as it were, and from afar .... Nevertheless, our belief in what we cannot see is helped by means of the things that were made.[23]

That there should be at least some strong connection between the metaphysics which can be gleaned from Hildegard's visions and other medieval ontologies should not be so surprising when one remembers that, historically, the eleventh century is regarded as a watershed in the rise of Western culture. Palmer, whose European history is more or less a benchmark for generations of students, remarks that '... such a time [a time of dynamic change] has been the last century ... Such a time, also, began in Europe in the eleventh century.'[24] He also reminds us of the facts of the early years of Hildegard's life: she was born around the time of the election of Gregory VII (Hildebrand), a 'dynamic and strong-willed' reformer. Brunn and Epiney-Burgard are at pains to emphasize Hildegard's interest in the autonomy of her order; surely some of the impetus for her zeal in trying to spread her cosmology and its corollaries came from the historical events which provided the context and backdrop for Churchly innovation.[25]

IV.

Perhaps the thinker with whom Hildegard bears the strongest comparison is Meister Eckhart. Fox makes this comparison explicitly;[26] for our purposes the more important point is that the Meister is regarded as standing squarely within the mystical tradition, and yet a number of philosophical positions are freely attributed to him and recur in extended commentary. The desire to characterize Meister Eckhart as a philosopher does not appear to stem solely from his use of the more Aristotelian and Thomistic terminology which, of course, was subsequent (in its extensive usage) to Hildegard's time – rather, Meister Eckhart is,

---

[23] Peter Lombard, *Libri IV Sententiarum* (Collegii S. Bonaventurae, l916), tom. I, 30-33, excerpted in
Wippel and Wolter, op. cit., and translated by Wolter, p. 208.

[24] R. R. Palmer, *A History of the Modern World* (New York: Knopf, l964), p. 22.

[25] Brunn and Epiney-Burgard, op. cit., pp. 3-4.

[26] Cf. fn. 3.

according to McGinn, a mystic without specifically being a visionary.[27] In other words, it appears that the virtue that Meister Eckhart possesses is that he engages in speculative ratiocination about some of the same problems that Hildegard pursued through ecstatic and visionary means. What is remarkable, however, is that again we find areas of consonance between the two thinkers. Meister Eckhart writes that the '... soul ... has a drop of understanding, a little spark, a little sprout ...', and that it '... has powers which work in the body'.[28] In a more extended commentary in Sermon 9 (one of his works in German, more accessible than several other of his commentaries), he explains:

> ... God is something that of necessity must be above being. Whatever has being, time or place does not touch God. He is above it. God is in all creatures, insofar as they have being, and yet he is above them. That same thing that he is in all creatures is exactly what he is above them. Whatever is one in many things must of necessity be above them. Some masters maintain that the soul is only in the heart. This is not so, and learned masters have gone astray here. The soul is complete and undivided at the same time in the foot and in the eye and in every part of the body.[29]

In his gloss on this part of the Sermon, Bernard McGinn notes the extent to which Eckhart seems to rely on what Aquinas would refer to as the via negationis, that is positing what God is not;[30] he also notes that Eckhart makes use of the fact that '... the conceptual distinction between substance and relation is legitimate in thinking about God as three-who-are-one, but is not in any way applicable to the hidden Godhead'.[31] In any case, for our purposes the noteworthy point is the obvious comparison between Hildegard and Meister Eckhart: Hildegard cites the 'mysterious gifts' of the Godhead (in Uhlein's translation of De Operatione, previously excised), and notes the Godhead's 'connectedness' to all living beings. Meister Eckhart notes that God is '... in all creatures, insofar as they have being ...' and goes on to make an analogy between the way in which God is 'in' things of the material world, and the way in which the soul is in the body. The soul, he asserts, is even in the foot.

---

[27] Bernard McGinn, *Teacher and Preacher* (New York: Paulist Press, l986), p. 14.
[28] Sermon 9 ('Quasi stella ...'), in McGinn, op. cit., p. 257.
[29] Ibid., p. 256.
[30] Ibid., p. 25.
[31] Ibid., p. 20.

Thus Hildegard seems to receive, visionarily, the same sort of knowledge – knowledge which, as Newman emphasizes, she wants to impart to others – that Eckhart receives largely through contemplation. It seems appropriate to assert that we would want to label both of these theorists philosophers.

V.

If we can find intersections between the interests and assertions of Hildegard and those of other, male, philosophers of the medieval period – as we have done above – it might be pertinent to ask in what her gynocentrism consists. Fox, Newman and Hozeski all seem to agree that there is a particularly strong gyncentred style to Hildegard's work, even if there is some slight disagreement about the manner in which it presents itself.[32] The concept of *viriditas* alone is not enough to enable us to make this assertion: it is structurally similar to many of the statements Meister Eckhart makes about the immanence of the divine nature. Nor can we posit the mystical or even visionary modes of Hildegard's epistemology as the primary factors here; clearly, these are elements of knowledge reception which are quite common among the medieval thinkers of a number of strands of the Christian tradition.

What strikes one preliminarily about Hildegard's thought is the nature of her ontology and the gradedness of its construction. Recent work on feminist epistemology in the analytic tradition has posited the 'dialectical' and 'connected' nature of female styles of coming-to-believe as paramount; androcentric styles have frequently been characterized as normative, and detached and distanced from the purported objects of knowledge. In the paradigmatic case, we can agree with Bordo, for example, that Descartes' style is peculiarly androcentric because the desire to pull oneself away from the objects of the senses and to categorize over them in such a way as to achieve freedom from error seems to recapitulate, according to psychoanalytic theory, certain elements of male personality formation.[33]

Hildegard's hierarchy of being, like Anne Conway's several hundred years later (although Conway, of course, employs much more standard

[32] Cf. the title of Newman's work on Hildegard, fn. 14.

[33] Susan Bordo, *The Flight to Objectivity* (Albany, NY: SUNY Press, l987). There is an extended discussion in this work of object relations theory and its relationship to an analysis of Descartes' work.

philosophical terminology, including the standard rationalist constructs of her time), demonstrates an intermingling of gradations throughout its structure. Although we are tempted to label Hildegard a standard dualist, insofar as Christian thought requires this, Hildegard, like Eckhart, precedes to some extent the Spinozan 'all things that are are in God' by establishing her ontology as permeated by the divine. In addition, the strong place of the Virtues – Knowledge of God, or the traditional Sophia figure of the Gnostic tradition being chief among them – ensures a sort of concern for interaction of the principles of masculine and feminine which is almost more characteristic of thinkers outside the Christian heritage.

Several of the commentators report Hildegard's own awareness of her status as a female, and this fact seems in itself significant, for example, to Newman who notes:

> [Her] texts ... interwoven as they are with references to her own simplicity, frailty, and femininity, insist on her authority with a defiance proportional to her fear that her books would indeed be concealed, altered, abridged, ridiculed, or ignored. In her Vita she tells how, when she was founding her new monastery at the Rupertsberg, many people asked 'why so many mysteries should be revealed to a foolish and uneducated woman, when there are many powerful and learned men', and some wondered whether she had been seduced by evil spirits.[34]

Somewhat heterodoxically, I would like to reiterate that the salient points here are neither Hildegard's status as female, nor her awareness of it. Rather, Hildegard's metaphysics is peculiarly gynocentric, even given the mystical epistemic base upon which it relies. Although she does not and cannot fashion the sorts of questions that would later be puzzling for dualists with regard to the interaction of two substances, she, like Conway at a later point, avoids these questions by positing interpenetration as primitive and by propounding a sort of metaphysical view of the union of the material and the divine which is, in its own way, a kind of precursor of interactive monism. Rupert of Deutz, a male mystic writer of the same period, shares with Hildegard the notion of visionary experience as a form of knowledge acquisition, but as Newman notes, 'His visions adorn and authenticate his writings, but they do not

[34] Barbara Newman, 'Hildegard of Bingen: Visions and Validation', in *Church History*, Vol. 54, No. 2 (June l985), l63-l75. This citation is from p. l7l.

determine the whole, and the inspiration he avows is of a far more general and less challenging sort than hers.'[35] Augustine, like Hildegard, compares the corporeal light and the spiritual light, but for Augustine, '... such light-speculation ... is envisaged [as] a mystical ascent involving brief moments of ecstasy rather than a continuous illumination such as Hildegard's.'[36]

Hildegard, like the later Conway and Masham, has developed a gynocentric philosophy based on a metaphysics and epistemology that reject the divorce and detachment so common to male theorists. Although the commentary on her work, much of it done within the framework of Church scholarship, has tended to emphasize her status as a visionary and a woman, my claim has been that neither of these facts fully explains the strength of Hildegard's work. Her ontological constructs, with the linkages between levels of being and the divine intercession of female or feminine figures, are themselves the most striking points of her theology. Hildegard, like Meister Eckhart, is a medieval mystic whose work speaks to us today. Unlike Meister Eckhart, she presents us with a gynocentric worldview that may serve as an antidote to our contemporary ills.

[35] Ibid., p. l73.
[36] Ibid., p. 168.

HUME, CAUSATION AND TWO ARGUMENTS CONCERNING GOD

JASON MEGILL

Carroll College

Abstract. In *Dialogues Concerning Natural Religion*, Hume (1779/1993) appeals to his account of causation (among other things) to undermine certain arguments for the existence of God. If 'anything can cause anything', as Hume claims, then the Principle of Causal Adequacy is false; and if the Principle of Causal Adequacy is false, then any argument for God's existence that relies on that principle fails. Of course, Hume's critique has been influential. But Hume's account of causation undermines the argument from evil at least as much as it undermines arguments for theism, or so I argue. I then suggest that Hume's account of causation can be used to formulate an alternative argument against classical theism.

INTRODUCTION

Hume's (1779/1993) ingenious attack on some of the traditional arguments of natural theology has influenced many. One important aspect of Hume's critique involves his account of causation. Some of the arguments in natural theology (e.g., some versions of the Cosmological argument) rely on certain causal principles; if Hume's account of causation is true, then these causal principles – along with the arguments that rely on them – must be rejected. This paper argues that if one accepts Hume's critique of certain arguments for theism, one must also reject the most famous argument for atheism, the argument from evil. That is, Hume's account of causation undermines the argument from evil just as

much as it undermines certain arguments for theism. This is good news for the theist. However, it also appears that Hume's account of causation can be used to formulate a novel argument against the existence of God.

In section I, I briefly discuss Hume's account of causation, including his famous claim that at least in principle, anything could be a cause of anything. In section II, I discuss the Principle of Causal Adequacy. This principle (i) was used in various arguments for theism, but (ii) is false if Hume is right about causation. In section III, I argue that Hume's account of causation undermines arguments from evil as much as they undermine arguments for theism. Given the considerable influence Hume's account of causation has had, theists might try responding to the problem of evil by appealing to Hume's account of causation. In section IV, I conclude by formulating a new Humean argument against theism.

I. 'ANY THING MAY PRODUCE ANY THING'

Hume's discussion of causation is well-known; my exegesis will be brief. Hume (1748/1975: 25) begins his discussion by drawing a distinction between two types of knowledge, (i) knowledge of 'relations of ideas' (Hume's term for analytic truths) and (ii) knowledge of 'matters of facts' (Hume's term for synthetic truths). Hume (1748/1975:25) claims that relations of ideas are 'discoverable by the mere operation of thought, without dependence on what is anywhere existent in the universe', while matters of fact 'are not ascertained in the same manner' (Hume 1748/1975). In other words, analytic truths are *a priori* while matters of fact are not, and so are *a posteriori*. Furthermore, Hume (1748/1975: 25) holds that 'the contrary of every matter of fact is conceivable'; given this, matters of fact are possibly false and so are not necessarily true. On the other hand, relations of ideas are not conceivably false and so are necessary. In short, Hume believes that (i) analytic truths are *a priori* and necessary while (ii) synthetic truths are *a posteriori* and contingent.

Hume (1748/1975) then analyzes the concept of causation. He discusses a number of criteria it appears that a causal relation must meet. For example, if *A* causes *B*, then (i) *A* and *B* should be contiguous, (ii) *A* should occur before *B*, and importantly, (iii) there should be a *necessary* connection between *A* and *B*. For Hume, there should be a necessary connection between cause and effect because causal inferences take us beyond what we are given in the senses (Noonan 2007). When making

a causal inference, we move 'from the observed to the unobserved' (Noonan 2007: 59), i.e., from the present to future; and the only way we could be justified in doing so is if there is a necessary connection between cause and effect. If there is no necessary connection between cause and effect, then we cannot be certain that the effect will occur even though the cause has. The problem, however, is that it appears that causal relations are not necessary, as Hume attempts to show with various arguments. For example, we can conceive of a cause occurring without its effect; this implies that causal relations cannot be necessary (assuming that conceivability entails possibility). Furthermore, Hume can find no evidence of a necessary connection between causes and effects in our experience: all we see is one event followed by another. Moreover, if we consider some object, some putative cause, in isolation, we can find no intrinsic property of the object that would tell us that the object could be a cause or what it could cause. There is no 'quality which binds the effect to cause and renders the one an infallible consequence of the other' (Hume 1748/1975: 63). That is, for Hume, if there is a necessary connection between cause and effect, we should be able to tell, *a priori*, what a given cause can cause (again, for Hume, necessity is associated with the *a priori*); since we cannot do so, causal relations cannot be necessary. But again, causal relations must be necessary if they are to warrant inferences about the future. It follows that our causal inferences are unjustified.[1]

The key claim for our purposes here is the following: there is no 'quality which binds the effect to cause and renders the one an infallible consequence of the other' (Hume 1748/1975: 63). There are no constraints on what can cause what; in principle, anything can cause anything.

II. CRITIQUE OF NATURAL THEOLOGY

The Causal Adequacy Principle claims that 'no cause can produce or give rise to perfections or excellences that it does not itself possess' (Russell 2005: Section 3). In the early modern period, this principle was used in a number of arguments for God's existence. The most famous example is Descartes's (1641/1984) trademark argument in the 3rd Meditation:

[1] This is, of course, an incomplete account of Hume's views on causation. For example, nothing has been said about the impossibility of giving a non-circular justification of induction. Nevertheless, this incomplete account is sufficient for our purposes here.

Descartes has the idea of God.Given that ideas must have a cause, Descartes's idea of God must have a cause. And given the Principle of Causal Adequacy, this cause must be adequate to cause the idea of God. Only God could be an adequate cause of the idea of God; so, God must exist. Others used the Causal Adequacy Principle in theistic arguments as well. Clarke (1704/1998), for example, argues that since any effect cannot have a 'perfection' that was not in the cause, there is no way that matter could give rise to thought. So, given that the universe has a first cause, this first cause cannot be material and unthinking (for otherwise thought would be unable to arise at all); that is, the first cause is an immaterial thinking being. Others who appealed to the Causal Adequacy Principle to establish the truth of theism include Cudworth (1678/1978) and even Locke.[2]

Clearly, Hume's account of causation and the Principle of Causal Adequacy are inconsistent. The Principle of Causal Adequacy places constraints on what can cause what: a cause must be sufficient or 'appropriate' in various ways to produce its effect. However, Hume's account of causation denies the existence of such constraints: anything can cause anything. If Hume's account of causation is correct, then the theistic arguments of Descartes, Cudworth, Clarke, Locke and others cannot be sound.

III. THE PROBLEM OF EVIL

The argument from evil is the most prominent argument for atheism; it is an argument against the existence of God, where God is understood as a being that is omnipotent, omniscient and omnibenevolent (the '3-O God'). Logical arguments from evil claim that the existence of the 3-O God is logically impossible; the combination of this being and some of the evil that exists in the actual world are inconsistent.[3] Evidential arguments from evil claim that the presence of at least some of the evil

[2] See Russell (2008, Chapter 3) for an excellent discussion of the Causal Adequacy Principle and the role it played in various theological arguments before and during Hume's time. The principle was so closely related to theism at that time that Hume's denial of it was cited by Hume's religious critics as evidence of his religious scepticism.

[3] For instance, one might make the argument: assume that the 3-O God exists. But this being would know when an innocent child is dying of leukaemia (it knows everything), it should want to prevent the death (it is all-good, after all) and it has the ability to do so (it is all-powerful). So no innocent children should die of leukaemia. Yet they do. The

in the actual world lowers the epistemic probability that the 3-O God exists.[4] Of course, theists have tried to respond to the problem of evil in various ways. Perhaps the actual world is the best possible world (Leibniz (1710/1951))? Perhaps there is no best possible world; God cannot be faulted for not actualizing something that cannot be actualized (Schlesinger (1964; 1977) and Forrest (1981))? Perhaps the evils in our world serve a purpose; e.g., they develop our souls (Hick 1966)? Perhaps God is not obligated to create the best possible world (Adams 1972)? Perhaps much of the evil in the actual world is a consequence of our free will (see Plantinga 1974; 1977)? Perhaps there is a reason why evil exists, but we simply do not know it (see Wykstra 1984)? And so on.

But, oddly enough, Hume could be an unlikely ally in the theist's effort to account for the existence of evil: Hume's account of causation and the *logical* problem of evil, at least, are *inconsistent*. Assume, as Hume's account of causation claims, that anything can cause anything. In logical notation, this can be written: (1) $\forall x \forall y \Diamond \text{Cause}(x, y)$.[5] But also suppose that it is contradictory, and so logically impossible, for the 3-O God to cause our universe: given some of the gratuitous evil in the actual world, an all-powerful, all-knowing and all-good being would not and could not have caused this world. That is, (2) $\neg \Diamond \text{Cause}(\text{God, our universe})$. But with substitution (or universal elimination) on line (1), one can infer: (3) $\Diamond \text{Cause}(\text{God, our universe})$. (2) and (3) contradict one another. So, not both (1) and (2) can be true. We must reject either (i) Hume's account of causation (or at least key parts of it) or (ii) the logical problem of evil. So, accepting Hume's critique of natural theology, or rather that part of the critique that depends on his analysis of causation, while simultaneously endorsing the logical problem of evil, is problematic.[6] And perhaps more importantly, the theist has another response to the logical problem of evil. We can agree with Hume that anything can cause anything. But then the 3-O God could cause our universe.

---

contradiction implies that our initial assumption is false: the 3-O God does not exist. For the logical argument from evil, see Mackie (1955) and McCloskey (1960).

[4] For the evidential argument from evil, see Rowe (1978; 1979; 1988; 1996) and Draper (1989).

[5] Though one might want to add the proviso that $x \neq y$, to avoid cases in which an entity causes itself.

[6] It is not difficult to find philosophers who think both (i) that Hume's critique of, say, the Cosmological argument is a serious one and (ii) that the problem of evil provides some evidence for atheism. Sobel (2004) is but one example.

One might wonder if *evidential* arguments from evil are also inconsistent with Hume's account of causation. They are. Again, evidential arguments claim that the existence of some of the evil in the actual world lowers the (epistemic) probability that the 3-O God exists. But recall that Hume claimed that anything can cause anything. There is nothing intrinsic to a cause that suggests what it could cause. So, it appears that not only can anything cause anything, but the probability that any *x* could cause a given *y* is the same. Just as we cannot look at a putative cause and determine, *a priori*, what it could cause, we cannot look at a putative cause and determine, *a priori*, what it is *likely* to cause either. Therefore, we should not be able to consider the concept of the 3-O God and determine – *a priori* – what it could likely and unlikely cause. The 3-O God could cause any universe – no matter how wonderful or horrific – with equal probability, if Hume is correct. Likewise, for Hume, we cannot determine, *a priori*, what could cause, or even likely could cause, a given effect. But then we cannot determine, *a priori*, what caused a given effect (for if we could, then there must be some connection between the nature of a cause and what it could cause in principle, and there is no such connection). So, we cannot examine the actual world, and infer that – given the evil that exists – it is *less likely* that it was caused by the 3-O God.

IV. A HUMEAN ARGUMENT AGAINST CLASSICAL THEISM

Even though Hume's account of causation undermines the argument from evil, it appears it can be used to formulate a different argument against classical theism (or against the existence of the 3-O God). While the claim that anything can cause anything makes it possible that the God of classical theism caused our universe, it also makes it highly improbable that the God of classical theism caused our universe.

Consider the following argument: first, suppose – as is suggested by Hume's account of causation – that anything can cause anything with equal probability. Then, grant to the theist that the universe has a cause. And consider a set *S* of, say, 100 properties. We stipulate that *S* contains three properties: omnipotence, omniscience, and omnibenevolence. As for the other 97 properties in *S* ... one can pick just about any properties one likes, so long as they are not logically inconsistent with any of the 3-O properties. One can form subsets of *S* by selecting properties in *S*. For

example, one subset will contain all three of the 3-O properties and no other properties; some subsets will contain the 3-O properties along with other properties in *S* (so, e.g., a set will contain the 3-O properties plus, say, property 37, whatever it is); some subsets might contain none of the 3-O properties; and so on. One can think of each subset as corresponding to a being; for instance, a subset that contains the 3-O properties might correspond to the 3-O God. But if it is correct that anything can cause anything, then any subset of properties in *S* is just as likely to correspond to a being that could cause the universe as any other. So, one might ask, 'what is the probability that the cause of the universe has all three of the 3-O properties?' Of course, the probability that the being will be omnipotent is 1/100. The probability that this omnipotent being will also be omniscient is 1/99. And the probability that this omnipotent and omniscient being will also be omnibenevolent is 1/98. So the probability that the being will have all three of these divine attributes – and so the probability that the cause of the universe is a being that can be called the God of classical theism – is 1/100 * 1/99 * 1/98. That is, 0.00000103. It is extremely improbable that the God of classical theism caused the universe, at least given Hume's account of causation.

The theist might respond that the calculation is too crude: it has been argued, for example, that an omnipotent being will also be omniscient.[7] If so, then given that the being is omnipotent, the probability that it will also be omniscient is 1. This would indeed raise the probability that the 3-O God caused our universe, but it would still only be 1/100 * 1/98, and so would still be very low; it would be 0.0001, to be precise. Perhaps the theist can appeal to the doctrine of divine simplicity? If divine simplicity is true, and so all of God's properties are indivisible, and so are unified in some sense, then perhaps if a being has one of the divine attributes, it has all of them? This would greatly increase the odds in favour of the theist ... but they would still only be 3/100, and so 1/33. These are still rather poor odds. And there do not seem to be too many other possible responses: the theist might argue that an omnipotent being could give itself other properties. So an omnipotent being could make itself all-knowing and all-good. But even here, the odds are only 1/100. The theist might argue that there are not even 100 properties ... if not, then *S* will have to be smaller, so the odds that the God of classical theism caused the universe will rise? But surely there are at least 100 properties. Of course, the theist

[7] See, for instance, Megill and Mitchell (2009).

could reject Hume's account of causation; indeed, this appears to be the most promising strategy.

In sum, Hume's account of causation has been taken by many to undermine some of the traditional arguments for God's existence. It does. But what some have failed to notice is that it also undermines the argument from evil. Hume's account of causation entails that the 3-O God could have caused our universe. However, Hume's account of causation can be used to formulate a different argument against classical theism. If the above arguments are sound, there are different strategies the theist and atheist might adopt. The theist might reject Hume's account of causation and search for other responses to the problem of evil; this might also make it possible to appeal to some of the traditional arguments for theism that Hume's account of causation undermines. Or the theist might accept Hume's account of causation, thereby solving the problem of evil; but the theist would then need to find some response to the alternative argument against theism just discussed. On the other hand, the atheist might endorse Hume's account of causation, thereby undermining many traditional arguments for God's existence and making it possible to appeal to the alternative argument against theism just discussed; they would have to abandon the argument from evil though. Or the atheist might reject Hume's account of causation and keep the argument from evil.

BIBLIOGRAPHY

Adams, R.M. 1972. 'Must God Create the Best?', *Philosophical Review*, 81(3): 317-332

Clarke, S. 1704/1998. *A Demonstration of the Being and Attributes of God*. Edited by E. Vailati (Cambridge: Cambridge University Press)

Cudworth, R. 1678/1978. *True Intellectual System of the Universe*, 2 Vols. (London, 1678; reprinted by New York & London: Garland)

Descartes, R. 1641/1984. *Meditations on First Philosophy* in *The Philosophical Writings of Descartes*, *Volume 2*. Translated by John Cottingham, Robert Stoothoff, and Dugald Murdoch (Cambridge: Cambridge University Press), pp. 1-62

Draper, P. 1989. 'Pain and Pleasure: An Evidential Problem for Theists', *Noûs*, 23: 331-350

Forrest, P. 1981. 'The Problem of Evil: Two Neglected Defenses', *Sophia*, 20: 49-54

Forrest, P. 1996. 'A Speculative Understanding of Evil', in *God Without the Supernatural* (Ithaca: Cornell University Press)

Hick, J. 1966. *Evil and the God of Love* (New York: Harper and Row)
Hume, D. 1748/1975. *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*, 3rd Edition, edited by L.A. Selby-Bigge and P. H. Nidditch. (Oxford: Oxford University Press)
Hume, D. 1779/1993. *Dialogues Concerning Natural Religion* in *Dialogues and Natural History of Religion*. Edited by J. A. C. Gaskin (Oxford & New York: Oxford University Press), pp. 29-130
Leibniz, G. 1710/1951. *Theodicy* (New Haven: Yale University Press)
Mackie, J.L. 1955. 'Evil and Omnipotence', *Mind*, 64: 200-212
McCloskey, H.J. 1960. 'God and Evil', *Philosophical Quarterly*, 10: 97-114
Megill, J and Mitchell, J. 2009. 'A Modest Modal Ontological Argument', *Ratio*, 22.3: 338-349
Noonan, H. 2007. *Hume* (Oxford: Oneworld Publications)
Plantinga, A. 1974. *The Nature of Necessary* (Oxford: Oxford University Press)
Plantinga, A. 1977. *God, Freedom, and Evil* (Grand Rapids, MI: Eerdmans)
Rowe, W. L. 1978. *Philosophy of Religion: An Introduction* (Encino, CA: Dickenson Publishing Company)
Rowe, W. L. 1979. 'The Problem of Evil and Some Varieties of Atheism', *American Philosophical Quarterly*, 16: 335-41
Rowe, W. L. 1988. 'Evil and Theodicy', *Philosophical Topics*, 16: 119-32
Rowe, W. L. 1996. 'The Evidential Argument from Evil: A Second Look', in *The Evidential Argument from Evil*. Edited by D. Howard-Snyder (Bloomington: Indiana University Press)
Schlesinger, G. 1964. 'The Problem of Evil and the Problem of Suffering', *American Philosophical Quarterly*, 1: 244-7
Schlesinger, G. 1977. *Religion and Scientific Method* (Boston: D. Reidel)
Sobel, J. H. 2004. *Logic and Theism: Arguments For and Against Beliefs in God* (New York: Cambridge University Press)

KANT'S NEGLECTED OBJECTION TO THE ONTOLOGICAL ARGUMENT

MICHAEL R. SLATER

Georgetown University

Abstract. This paper argues that Kant's most famous objection to the ontological argument – that existence is not a real predicate – is not, in fact, his most effective objection, and that his 'neglected objection' to the argument deserves to be better known. It shows that Kant clearly anticipates William Rowe's later objection that the argument begs the question, and discusses why Kant himself seems to have overlooked the force of this criticism in his attempt to demolish the traditional proofs for God's existence.

Many philosophers have thought that Kant's most significant and original criticism of the ontological argument is his objection that existence is not a real predicate (A599/B627), and there is little doubt that this is the most famous and influential of his criticisms of the traditional arguments for God's existence.[1] Yet as William Rowe argued almost forty years ago, this objection does not clearly provide a conclusive refutation of the ontological argument, as it seems to rest upon some mistaken or incomplete claims about the nature of predication.[2] Indeed, Rowe showed

[1] All references to the *Critique of Pure Reason* (1781/1787) are to the Cambridge Edition, translated and edited by Paul Guyer and Allen Wood (Cambridge: Cambridge University Press, 2008), and follow the pagination of the standard German *Akademie* edition of Kant's works.

[2] William Rowe, 'The Ontological Argument', in Joel Feinberg (ed.), *Reason and Responsibility: Readings in Some Basic Problems of Philosophy*, third edition (Encino and Belmont, CA: Dickenson Publishing Company, 1975), pp. 8-17. Reprinted in Steven M. Cahn (ed.), *Ten Essential Texts in the Philosophy of Religion: Classics and Contemporary Issues* (New York and Oxford: Oxford University Press, 2005), pp. 102-16. All page references are to the Cahn volume.

that one can grant the proponent of the argument the assumption that existence is a real predicate, as well as the definition of 'God' as an existing, wholly perfect being, and still show him that 'it will not follow from that definition that there actually exists something to which his concept of God applies'.[3] Focusing his critique on Anselm's version of the argument, Rowe objected to the ontological argument on two basic grounds: first, that it does not follow from Anselm's concept of God that some existing thing must exemplify that concept, any more than it follows from the concept of a 'magican' (an existing magician) that some existing thing must exemplify that concept; and second, that the argument begs the question regarding the existence of God.[4] Rowe's objections to Anselm's and other non-modal versions of the argument are successful, I think, but the latter of the two is hardly an original objection.[5] What I want to show in this paper is that Kant in fact anticipates this objection in *The Critique of Pure Reason*, though he does not focus specifically, as Rowe does, on Anselm's version of the argument.[6] If I am right, this neglected objection of Kant's deserves recognition on the part of philosophers of religion, not only insofar as it highlights another important and easily overlooked feature of his critique of the ontological argument, but also insofar as it highlights the cumulative nature of his critique.

A number of contemporary philosophers now think that Kant's critique of the traditional arguments for the existence of God in the First *Critique* was not only incomplete but also confused in some of its basic assumptions. In addition to the point about predication raised above, Kant's critique also seems to be incomplete insofar as it supposes that there are – and can only be – three types of theistic arguments (A591/B619), which overlooks the possibility of other types of arguments for the existence of God such as the moral argument. And it is confused, some critics argue, insofar as it supposes that the teleological argument (or as Kant terms it, 'the physico-theological proof') is logically dependent

[3] Ibid., p. 112.

[4] Ibid., pp. 111-16.

[5] A modal version of the ontological argument such as Alvin Plantinga's does not fall prey to these criticisms, in my view, though I will not argue for this claim here. For Plantinga's defence of this version of the argument, see 'The Ontological Argument' in James F. Sennett (ed.), *The Analytic Theist: An Alvin Plantinga Reader* (Grand Rapids and Cambridge: Eerdmans Publishing Co., 1998), pp. 50-71.

[6] Kant seems to have had Descartes' version of the ontological argument in the Fifth Meditation primarily in mind, though he clearly thought that his objections to the argument applied equally to other versions as well.

on the cosmological argument, and that the latter in turn depends for its success on the ontological argument and its attempt to demonstrate *a priori* the existence of a necessary being (A630/B658), with the consequence that a destruction of the ontological argument necessarily entails the destruct of the other types of arguments.[7] Although I accept these criticisms, I will not undertake to defend them here, both for reasons of space and because my aim is to focus on what Kant gets right about the failings of traditional ontological arguments, as opposed to what he gets wrong.[8]

Before discussing what I have termed Kant's neglected objection to the ontological argument, though, it is perhaps worth noting that while a sympathetic critic such as Rowe disagrees with his claim that existence is not (or cannot be) a real predicate, along with the objection that Kant develops from this claim, he nevertheless shares with Kant – and is very likely indebted to him for – the basic assumption that 'from the

[7] Perhaps the most forceful and influential of these critics is Alvin Plantinga, who views this supposition as one of Kant's 'fashionable confusions about the ontological argument'. See 'Reason and Belief in God' in Alvin Plantinga and Nicholas Wolterstorff (eds.), *Faith and Rationality* (Notre Dame, IN: University of Notre Dame Press), pp. 16-93; for a more extensive critique of Kant's objections to the ontological argument, see Plantinga's 'The Ontological Argument', ibid. But even generally sympathetic interpreters such as Allen Wood have expressed serious doubts about the adequacy of Kant's views on the traditional theistic arguments, including the problematic assumption mentioned above. See Allen W. Wood, 'Rational Theology, Moral Faith, and Religion' in Paul Guyer (ed.), *The Cambridge Companion to Kant* (New York: Cambridge University Press, 1992), pp. 394-416.

For a fundamentally different interpretation of Kant's views on the traditional theistic arguments see Paul Guyer, who claims – in apparent contrast to what Kant himself states at A630/B658 – that 'for Kant the ontological argument is a more involuted line of thought' than the cosmological argument, one that leads 'from the existence of anything contingent to something necessary (the cosmological argument as the first step of the ontological argument), to the idea of a being comprehending all reality, and then as it were back down to the idea of the existence of this being as a necessary existence – following from its own concept and not being dependent upon and therefore possibly precluded by the existence of anything else'. See Guyer, *Kant* (London and New York: Routledge, 2006), p. 147. Guyer does not provide evidence for this surprising interpretation, however, and even if he is right this does not seem to affect the larger objection made by some critics that Kant's views on this matter are insufficiently developed.

[8] In putting the point this way, I simply mean to suggest that some of Kant's criticisms of medieval and modern versions of the ontological argument are successful. What I do not think Kant shows, however, is that his criticisms represent a successful refutation of *any* version of the ontological argument. Indeed, his assumption that this is essentially a single argument, rather than a family of closely related arguments, is itself problematic.

logical analysis of a certain idea or concept we can never determine that there exists in reality anything answering to that idea or concept'.[9] This assumption is clearly on display in the First *Critique*, where we find Kant arguing that given the principle that all existential propositions are synthetic as opposed to analytic (A598/B626), it follows that 'if I think of a being as the highest reality (without defect), the question still remains whether it exists or not. For although nothing at all is missing in my concept of the possible real content of a thing in general, something is still missing in the relation of my entire state of thinking, namely that the cognition of this object should also be possible a posteriori' (A600/B628).[10] As Kant summarizes this objection to the ontological argument, 'thus whatever and however much our concept of an object may contain, we have to go out beyond it in order to provide it with existence' (A601/B629).[11] This conclusion looks virtually identical to Rowe's first objection to the ontological argument, with the major difference being that Rowe arrives at this conclusion by granting the proponent of the ontological argument the assumption that existence is (or can be) a real predicate, which Kant of course does not. This constitutes an advantage of Rowe's version of the objection, I think, but the basic objection itself is already anticipated in Kant's critique of the argument.

Rowe's second objection, however, that the ontological argument begs the question, adds little if anything to Kant's prior objection in the First *Critique*. Although Kant does not specifically state that proponents of the argument are guilty of committing this informal fallacy, it is nevertheless clearly implied by his remarks. In response to the would-be proponent, he writes: 'You have already committed a contradiction when you have brought the concept of existence, under whatever disguised name, into the concept of a thing which you would think merely in terms of its possibility.' (A598/B626)[12] What is highly interesting (and in some ways perplexing) about Kant's analysis of this defect in the argument, though, is that having identified this fallacious piece of reasoning he immediately proceeds to discuss how the proponent of the argument has committed a *tautology*, presumably on the assumption that that is the more serious problem facing the argument. The passage continues as follows:

[9] William Rowe, 'The Ontological Argument', p. 108.
[10] Kant, *Critique of Pure Reason*, p. 566.
[11] Ibid., p. 568.
[12] Ibid., p. 566.

> If one allows you to do that, then you have won the illusion of a victory, but in fact you have said nothing; for you have committed a mere tautology. I ask you: is the proposition, **This or that thing** (which I have conceded to you as possible, whatever it may be) **exists** – is this proposition, I say, an analytic or a synthetic proposition? If it is the former, then with existence you add nothing to your thought of the thing; but then either the thought that is in you must be the thing itself, or else you have presupposed an existence as belonging to possibility, and then inferred that existence on this pretext from its inner possibility, which is nothing but a miserable tautology. The word 'reality', which sounds different from 'existence' in the concept of the predicate, does not settle it. For if you call all positing (leaving indeterminate what you posit) 'reality', then you have already posited the thing with all its predicates in the concept of the subject and assumed it to be actual, and you only repeat that in the predicate. If you concede, on the contrary, as in all fairness you must, that every existential proposition is synthetic, then how would you assert that the predicate of existence may not be cancelled without contradiction? – since this privilege pertains only in the analytic propositions, as resting on its very character (A598/B626).[13]

Why does Kant construct his objection in this way? One possibility is that he assumes that forcing the proponent of the ontological argument into a dilemma – namely, either concede that the proposition 'God (defined as existing) exists' is an analytic proposition, in which case it commits a tautology, or concede that it is a synthetic proposition, in which case the predicate of existence can be denied without contradiction – is a better argumentative strategy (because a more effective criticism) than simply pointing out that she has begged the very question at issue. It is not entirely clear, however, that Kant fully appreciates the force of the objection he has just made, perhaps because he is so focused on laying out the dilemma mentioned above. Another possibility, then, is that *Kant constructs his objection in the way he does because he conceives the smuggled assumption of God's actual existence in the premises of the ontological argument in terms of a logical self-contradiction, rather than what it straightforwardly is: an illicit assumption that commits an informal as opposed to a formal fallacy*. Kant apparently thinks it necessary at this point to force the proponent of the ontological argument into a dilemma, but it seems to me that he has just made a fatal objection

[13] Ibid., p. 566.

to the argument, and that no further argument is needed – including not only the aforementioned dilemma, but also the famous objection that 'existence is not a real predicate' which he proceeds to develop at A599/B627. If this interpretation is right, then what I have termed Kant's 'neglected objection' turns out to be doubly neglected, not only by those with an interest in Kant's critique of the traditional arguments for God's existence, but also partly by Kant himself.

TRADITIONAL ISLAMIC EXCLUSIVISM – A CRITIQUE

IMRAN AIJAZ

University of Michigan-Dearborn

Abstract. In this paper, I give an account and critique of what I call 'Traditional Islamic Exclusivism' – a specific Islamic interpretation of religious exclusivism. This Islamic version of religious exclusivism rests on exclusivist attitudes towards truth, epistemic justification and salvation. After giving an account of Traditional Islamic Exclusivism by explaining its theological roots in the Qur'an and *ahadith* (reports of sayings attributed to the Prophet Muhammad), I proceed to critique it. I do so by arguing that Islamic epistemic exclusivism, which forms the main core of Traditional Islamic Exclusivism, is implausible. This criticism subsequently opens up further lines of criticism and discussion of both salvific and alethic exclusivism in an Islamic context. I conclude with some remarks about the implications and significance of my criticisms of Traditional Islamic Exclusivism.

I. INTRODUCTION

Contemporary Philosophy of Religion has seen a lot of discussion surrounding *religious exclusivism*, mainly due to the fact that philosophers of religion have, over the last few decades, started to take seriously the rich diversity in beliefs about religion. As John Hick explains, for a long time the labels 'Philosophy of Religion' and 'Philosophy of The Christian Religion' were treated as synonymous.[1] It is only during the last couple of decades that philosophers of religion in the West have, in Hick's words, 'increasingly felt obliged to take note of the fact that Christianity is only one of the great world faiths and that monotheism is only one of the

[1] John Hick, 'Religious Pluralism', in Charles Taliaferro et al., *A Companion to The Philosophy of Religion*, Second Edition (Oxford: Blackwell Publishing, 2010), pp. 710-17 (p. 710).

major types of religion'.[2] According to many philosophers of religion such as Hick, incorporating facts about the world's religious diversity into an assessment of the plausibility of religious claims is important. It also calls into question the reasonableness of maintaining an exclusivist attitude in religious matters.

But what *is* religious exclusivism? Even just a cursory survey of the relevant literature reveals that there is no consensus on how to understand it and that it is construed in a variety of different (and sometimes confusing) ways. David K. Clark, for example, understands it as follows: 'Only one true religion leads to God. Attaining the spiritual goal requires a believer to find and follow the one true faith, for other religious paths will *not* lead to the spiritual goal.'[3] Discussing religious exclusivism specifically within an Islamic context, Zain Ali defines what he calls 'Muslim Exclusivism' as follows: 'Islamic theism is overall epistemically superior [*sic*] and that all other incompatible beliefs are false.'[4] In just these two definitions alone, we see that religious exclusivism can be understood with reference to three *distinct* matters, even if they are conceptually connected: truth, epistemic justification and salvation.

In this paper, I will discuss religious exclusivism in an Islamic context. In particular, I will focus on exclusivism with respect to truth, epistemic justification and salvation in the Islamic religion – three tiers that form the basis of what I will call 'Traditional Islamic Exclusivism'. I will give an account of Traditional Islamic Exclusivism by basing it on the traditional Islamic understanding of religious truth, epistemic justification and salvation, after which I will proceed to critique it. I will then conclude with some remarks about the implications and significance of my assessment of Traditional Islamic Exclusivism.

II. TRADITIONAL ISLAMIC EXCLUSIVISM EXPLAINED

I will give the label *Traditional Islamic Exclusivism* to a view that, as I shall construe it, is comprised of a conjunction of three distinct positions situated within an Islamic context and based on traditional

[2] Ibid.

[3] David K. Clark, 'Religious Pluralism and Christian Exclusivism', in Francis J. Beckwith et al., *To Everyone An Answer*, (Downers Grove, IL: InterVarsity Press, 2009), pp. 291-307 (p. 293).

[4] Zain Ali, *Faith, Philosophy and The Reflective Muslim* (Basingstoke: Palgrave Macmillan, 2013), p. 138.

Islamic understanding: (1) *alethic exclusivism*, (2) *epistemic exclusivism* and (3) *salvific exclusivism*. Let me explain these in turn.

2.1 Alethic Exclusivism

First, I understand alethic exclusivism as follows. A person is an alethic exclusivist with respect to a proposition *p* if and only if, in believing that *p* is true, she takes to be false any beliefs that are incompatible with *p*. Thus understood, alethic exclusivism sounds obviously true and uncontroversial; indeed, it seems to be one of the fundamental principles that guide our reasoning. We believe in the falsity of propositions like 'Abraham Lincoln is alive' or 'George W. Bush is the President of The United States' precisely because we believe that the propositions 'Abraham Lincoln is dead' and 'Barack Obama is the President of The United States' are true. Those who subscribe to alethic exclusivism within an Islamic context will maintain the falsity of those beliefs that are incompatible with Muslim belief. For example, the Muslim alethic exclusivist will believe in the falsity of the propositions 'Zeus exists' or 'Jesus is the son of God' because their falsity is entailed by Muslim belief.

In addition to being a basic principle underlying rationality, alethic exclusivism also follows from a natural interpretation of the primary source of authority in Islam, the Qur'an. In reading the Islamic Scripture, one will constantly find denials of religious propositions that are incompatible with those that are affirmed as part of Islamic belief. In affirming the existence of Allah as the only God, for instance, the Qur'an denies the existence of any other deity (2:163). Or, to consider another example, in affirming the strict Islamic understanding of God's (indivisible) oneness (112:1), the Qur'an rejects the Christian doctrine of the Trinity (4:171). Many more examples of this sort can be found in the text.

2.2 Epistemic Exclusivism

Next, epistemic exclusivism. It is epistemic exclusivism that, as I see it, provides the principle basis and spirit of religious exclusivism in Islam. As with alethic exclusivism, epistemic exclusivism can be understood more generally without reference to religion. Here is how I will construe it. A person is an epistemic exclusivist with respect to proposition *p* if and only if, in believing that epistemic justification exists for *p*, she rejects the existence of epistemic justification for any propositions incompatible

with *p*. Why would a person be attracted to epistemic exclusivism, thus understood? There are a few possible answers. One that is pretty common is when a person believes that the epistemic justification for *p* is so strong or overwhelming that the existence of epistemic justification for propositions incompatible with *p* is highly unlikely.[5] As an example, consider Richard Dawkins' views about evolution and Young-Earth creationism, two incompatible perspectives about the origins of life and species on our planet. In his book *The Greatest Show on Earth*,[6] Dawkins states that '[t]he evidence for evolution grows by the day, and has never been stronger ... the "theory" of evolution is actually a fact – as incontrovertible a fact as any in science'.[7] He compares those who would deny evolution to people who believe that the Roman Empire never existed or that the Holocaust never happened.[8] Indeed, for Dawkins, the evidence for evolution is so strong that, as he stated controversially in 1989, 'It is absolutely safe to say that if you meet somebody who claims not to believe in evolution, that person is ignorant, stupid or insane (or wicked, but I'd rather not consider that).'[9] Furthermore, given his belief in the preponderance of evidence supporting evolution, Dawkins has always maintained that any 'evidence' to the contrary, particularly in support of Young-Earth creationism, is virtually non-existent. For instance, he states more recently: 'Just as I wouldn't expect a gynaecologist to have a debate with somebody who believes in the Stork-theory of reproduction, I won't do debates with Young Earth creationists.'[10]

The kind of epistemic exclusivism that I have described here is evident in traditional Islamic belief. Consider the two core beliefs

---

[5] It may be helpful to compare Gilbert Harman's explanation of how one can dismiss alleged evidence for not-*p* as 'misleading' in cases where one actually *knows p*: 'If I know that *h* is true, I know that any evidence against h is evidence against something that is true; so I know that such evidence is misleading.' See Gilbert Harman, *Thought* (Princeton, NJ: Princeton University Press, 1973), p. 148.

[6] Richard Dawkins, *The Greatest Show on Earth: The Evidence for Evolution*, (New York: Free Press, 2010).

[7] Ibid., p. vii.

[8] Ibid., pp. 3-4.

[9] Richard Dawkins, 'Ignorance is No Crime', *The Richard Dawkins Foundation*, available at: <http://old.richarddawkins.net/articles/114> (last accessed 1/30/14).

[10] Stoyan Zaimov, 'Richard Dawkins Explains Why He Doesn't Debate Young-Earth Creationists', *The Christian Post*, available at: <http://www.christianpost.com/news/richard-dawkins-explains-why-he-doesnt-debate-young-earth-creationists-107196/> (last accessed 1/30/14).

of the Islamic religion that are expressed in the *Shahada*, the Muslim declaration of faith[11] – belief in the existence and oneness of God and belief in Muhammad as God's Prophet. Both the Qur'an and the *ahadith* are clear that there is very strong epistemic justification for both of these beliefs and very little, if any, for those beliefs incompatible with them. In what follows, I will itemize (in no particular order) the different types of epistemic justification for belief in the existence and oneness of God as well as belief in Muhammad as God's Prophet, according to traditional Islamic understanding:

(1) Belief in God as Our Original, Natural Disposition (Fitrah): In a well-known *hadith*, the Prophet Muhammad is reported to have said: 'Every child is born with a true faith of Islam [*fitrah*] ... but his parents convert him to Judaism, Christianity or Magianism, as an animal delivers a perfect baby animal. Do you find it mutilated?'[12] The reference to *fitrah* is a reference to what, according to Islamic tradition, is our original disposition to recognize and worship only God (i.e. to be Muslims). The Qur'an also affirms this understanding of the term: 'So direct your face toward the religion, inclining to truth. [Adhere] to the *fitrah* of Allah upon which He has created [all] people. There is no altering of Allah's creation.' (30:30) For this reason, many individuals who come to Islam after leaving their non-Islamic faith or worldview refer to themselves as *re*verts and not *con*verts;[13] the idea here is that such individuals are *returning to* their original disposition to be Muslims. One way that the concept of *fitrah* can be used to support Islamic epistemic exclusivism is as follows: While belief in the existence and oneness of Allah is our natural disposition, beliefs that are incompatible with this cornerstone of Islamic doctrine can be dismissed as epistemically unjustified, as they are held merely on the basis of *taqlid* (imitation).[14]

---

[11] Translated into English, the *Shahada* reads as follows: 'There is no god but God, Muhammad is the Messenger of God.'

[12] Sahih Bukhari, 2:441, available at: <http://www.sacred-texts.com/isl/bukhari/bh2/bh2_442.htm> (last accessed 1/30/14).

[13] See, for example, '10 embrace Islam in Riyadh', *Arab News*, Tuesday, 4 February, 2014, available at: <http://www.arabnews.com/news/520291> (last accessed 2/5/2014). A search for the word 'revert' on this website's search engine brings up dozens of stories of people 'reverting' to Islam.

[14] The term is typically used by Muslim apologists in a pejorative sense and as a synonym for 'blind belief/faith'. See Al-Ghazali's discussion of *taqlid* as a belief forming mechanism to be contrasted with *fitrah* in his 'Deliverance from Error', in W. Montgomery Watt, *The Faith and Practice of Al-Ghazali* (London: G. Allen and Unwin, 1967), p. 21.

(2) The 'Islamic Anthropological Argument': A further component of epistemic justification for the Muslim view that Islam is our primordial religion comes by way of what I shall term The 'Islamic Anthropological Argument'. In essence, the argument maintains that every nation on earth was sent a messenger calling people to worship Allah alone and to shun all false deities (e.g. see Qur'an 10:47, 16:36 and 35:24). Some of these messengers are mentioned by name in the Qur'an while others are not, a point acknowledged explicitly in the text itself (40:78). The Islamic Anthropological Argument serves to augment the Muslim understanding of *fitrah* by pointing out that the messengers of God were sent to people of all nations to bring them *back* to the straight path of Islam (Qur'an 2:213). Islamic tradition maintains that our natural disposition is to worship God alone, even if it is subsequently corrupted. The pagan Arabs of Muhammad's time, for instance, acknowledged that God owns all that is in the heavens and the earth; that He is the Lord of the seven heavens; and, that in His Hand is dominion over all things. Despite believing all of this, they dismissed the (Islamic) belief in bodily resurrection as false. In noting all this, the Qur'an labels them as deluded and liars (23:81-89).

(3) Qur'anic Arguments for The Existence and Oneness of God: The Qur'an presents two main arguments for the existence of God as further sources of epistemic justification for Islamic belief. The Spanish Muslim philosopher, Ibn Rushd (Averroes), provides a helpful discussion of these two arguments in his *The Exposition of The Methods of Proof*.[15] According to Ibn Rushd, the two main Qur'anic arguments for the existence of God are best understood as arguments from providence and invention. Based on his discussion,[16] we can present these arguments in standard form as follows:

The Argument from Providence

(1) All existing things (found in the world) are suited to man's existence.
(2) Things that are suitable for man's existence are necessarily so due to an Agent both willing and intending such suitability.
    a. The suitability in question cannot be due to chance.

[15] See the translation of this work in Ibrahim Y. Najjar, (trans.), *Faith and Reason in Islam* (Oxford: Oneworld, 2001).

[16] Ibid., pp. 33-34.

(3) (Therefore) There exists an Agent (i.e. God) who providentially ordered the world for man.

The Argument from Invention

(1) All existing things (animals, plants, etc.) are invented.
(2) For everything invented, there is an inventor.
(3) (Therefore) There exists an Inventor (i.e. God) of every existing thing in the world.

After giving an account of The Argument from Providence and The Argument from Invention, Ibn Rushd maintains that these two arguments 'are the arguments favoured by religion'.[17] Qur'anic verses drawing our attention to evidence that points to the existence of God refer to (1) The Argument from Providence (e.g. 78:6-16), (2) The Argument from Invention (e.g. 86:5, 88:17, 22:73) or (3) a combination of The Argument from Providence and The Argument from Invention (e.g. 2:20-23).[18]

The Qur'an also presents some arguments for the oneness of God and against polytheism. The seeds of these arguments are found in verses like: 'Say: If there had been (other) gods with Him, as they say, behold, they would certainly have sought out a way to the Lord of the Throne!' (17:42); and, 'If there were therein gods beside Allah, then verily both (the heavens and the earth) would have been disordered. Glorified be Allah, the Lord of the Throne, from all that they ascribe (unto Him).' (21:22) Verse 23:91 explicitly repudiates the Christian doctrine of the divinity of Jesus as well as the doctrine of the Trinity: 'Allah has not taken any son, nor has there ever been with Him any deity. [If there had been], then each deity would have taken what it created, and some of them would have sought to overcome others. Exalted is Allah above what they describe [concerning Him].' Based on such verses, Muslim thinkers like 'Abd al-Jabbar and al-Ghazali formulated sophisticated arguments for the conclusion that there can be only one divine being.[19]

(4) The 'Inimitability' Argument for The Qur'an's Divine Origin and The Prophet of Muhammad: Islamic tradition has typically maintained

[17] Ibid., p. 35.

[18] Ibid.

[19] See, for example, 'Abd al-Jabbar's 'Book of the Five Fundamentals', in Richard C. Martin et. al, *Defenders of Reason in Islam: Mu'tazilism from Medieval School to Modern Symbol* (Oxford: Oneworld Publications, 1997), pp. 90-115 (p. 96); and, 'Al-Ghazali's Tract on Dogmatic Theology', edited, translated, annotated and introduced by A. L. Tibawi, *Islamic Quarterly*, IX (1965), 65-122 (p. 104).

that the primary miracle of the Prophet Muhammad is the Qur'an.[20] The miracle of the Qur'an is said to prove that Muhammad is God's Prophet (and *ipso facto* that there is a God, since Muhammad is God's Prophet only if God exists). Andrew Rippin summarizes the argument from the Qur'an's miraculousness, which originated during the early days of Islam, as follows:

> It would appear that, early on, Muslims had to defend their nascent religion against Christian theological attack in the area of the Fertile Crescent, especially Iraq. The following argument was constructed: miracles prove the status of Prophethood and the Qur'ān is Muhammad's miracle; therefore, Muhammad was truly a prophet and Islam is a true, revealed religion. All participants in the debate appear to have agreed on the first premise. What Muslims had to prove, and Christians disprove, was the validity [*sic*] of the second, for the conclusion, the truth of Islam, stood or fell on its credibility.[21]

As Rippin explains, the controversial premise here is whether the Qur'an is indeed a 'miracle'. The classical Islamic argument that emerged to support this premise is based on the alleged 'inimitability' of the Muslim Scripture. Rippin explains this argument thus:

> Over time, the argument became one concerned to prove the 'inimitability' of the Qur'ān, an argument which, its proponents were quick to point out, had a basis in the Qur'ān itself... [T]he production of a text 'like' the

[20] It is controversial within the Islamic tradition whether the Prophet Muhammad performed any miracles outside of bringing the revelation of the Qur'an. There are numerous *ahadith* that document Muhammad performing miracles such as splitting the moon, causing water to spring from his hands or multiplying a small quantity of food that fed over one thousand people (see Sahih Al-Bukhari, Volume 5, Book 58, No. 208; Volume 1, Book 4, No. 170 and Volume 5, Book 59, No. 428). The Qur'an, however, seems to strongly suggest that Muhammad performed no miracles apart from conveying the revelation of the Qur'an. In verses 17:90-93, for instance, we read that the unbelievers asked the Prophet to perform miracles, such as causing a spring to gush from the ground or have fragments of heaven fall upon them. In response, Muhammad says that here is merely a messenger. Verses 25:7-8 also note that the unbelievers complained that Muhammad did not come across as extraordinary or a performer of miracles. Based on such Qur'anic data, a number of Muslim thinkers, such as Muhammad Husayn Haykal, have maintained that the revelation of the Qur'an was the only miracle that one can legitimately ascribe to the Prophet Muhammad. See Haykal's *The Life of Muhammad* (Kuala Lumpur: Academe Art & Printing Services, 2008), pp. lxxxvi-lxxxvii.

[21] Andrew Rippin, *Muslims: Their Religious Beliefs and Practices*, Third Edition (London: Routledge, 2005), p. 38.

> Qur'ān is encouraged but known to be impossible: 'Produce a sūra like it [i.e., the Qur'ān], and call on whom you can, besides God, if you speak truthfully' (Qur'ān 10/38); 'Well then bring ten chapters the like of it, forged!' (Qur'ān 11/13). God has given the Qur'ān to Muhammad and because of its divine origin, no text 'like' it can, in fact, be produced. The inimitability of the text proves its divine authorship and thus its status as a miracle, confirming Muhammad's role and the veracity of Islam.[22]

According to Issa J. Boullata, general Muslim consensus maintains that the inimitability of the Qur'anic text refers mainly to its *stylistic* supremacy, which is held to be all the more remarkable given that the Qur'anic revelation was proclaimed by Muhammad, an illiterate man.[23] Muslim tradition maintains that this challenge to produce something similar to the Qur'an has never been successfully met since it was first raised. This alleged fact is offered as strong evidence for the divine origin of the Qur'an and the Prophethood of Muhammad.

In reading the Qur'an and *ahadith*, one gathers the impression that the four sources of epistemic justification I have itemized above see Islamic belief in the existence and oneness of God as well as belief in Muhammad as God's Prophet as being accompanied by very strong epistemic justification. Indeed, at times, the Author of the Qur'an uses an *incredulous* tone in noting rejection of Islamic belief, as in the following verse: 'How can ye reject the faith in Allah? – seeing that ye were without life, and He gave you life; then will He cause you to die, and will again bring you to life; and again to Him will ye return.' (2:28) Or again, 'Can there be any doubt concerning Allah, the Creator of the heavens and the earth?' (14:10) It is not surprising then that, for the average Muslim, Islam 'coincides in his mind with the irresistible logic of things', as Frithjof Schuon puts it.[24] A corollary of this aspect of the Qur'anic *Weltanschauung* is that the only alternative to having Islamic faith is deliberate and perverse rejection of it; that is to say, having Islamic faith or deliberately and perversely rejecting it are not only two mutually exclusive states of affairs, but they are also, according to the Qur'an, jointly exhaustive. In the Islamic Scripture, there is no room for

---

[22] Ibid.

[23] Issa J. Boullata, 'The Rhetorical Interpretation of the Qur'an: I'jaz and Related Topics' in *Approaches to The History of The Interpretations of The Qur'an*, Andrew Rippin, (ed.), (Oxford: Clarendon Press, 1988), pp. 139-157 (p. 142).

[24] Frithjof Schuon, *Stations of Wisdom* (Bloomington, IN: World Wisdom Books, 1961), p. 64, note 1.

non-belief that arises from, say, ignorance or (sincere) incredulity.[25] Based on all of this, it seems that epistemic exclusivism, as I have construed it, is indeed a significant component of Traditional Islamic Exclusivism.

2.3 Salvific Exclusivism

The final component of what I have termed Traditional Islamic Exclusivism is salvific exclusivism, which I will interpret as follows. A person is a salvific exclusivist with respect to a religion *r* if and only if, in believing that acceptance of *r* (or acceptance of its essential doctrines) leads to salvation, she rejects the view that religions other than *r* can lead to salvation.[26] A Muslim salvific exclusivist is one who maintains that acceptance of Islam is the only means to salvation; no other religion can offer such salvation. Achievement or loss of salvation, as understood in traditional Islam, is strongly associated with, if not reducible to, a person's entry into paradise or punishment in hell; and, whether a person is 'saved' or not is principally determined by whether he or she responded appropriately to Islamic belief (see e.g. Qur'an 2:82, 3:133, 5:119, 2:24, 4:55-56).

Perhaps the clearest statement of salvific exclusivism in the Qur'an is in verse 3:85: 'And whoever desires a religion other than Islam, it shall not be accepted from him, and in the hereafter he shall be one of the losers.' Now, the Qur'an does contain some verses stating that followers of other religions, such as Judaism and Christianity, will achieve salvation (e.g. 2:62; 5:69). It is difficult to assess what bearing such verses have on the traditional Islamic understanding of salvation, however, for a few reasons. First, these verses, which are few in number, contradict the ethos of the Qur'an as well as the *ahadith*, which generally subscribes to salvific exclusivism. Second, there are claims in both the Qur'an and *ahadith* explicitly stating that Jews and Christians will be condemned to hell for their beliefs. Consider, as an example, verses 5:72-73: 'They do blaspheme who say: "Allah is Christ the son of Mary" ... Whoever joins other gods with Allah, – Allah will forbid him the garden, and the Fire will be his abode ... They surely disbelieve who say: Lo! Allah is the third

[25] For more discussion on this, see Imran Aijaz, 'Some Ruminations about Inculpable Non-belief', *Religious Studies*, Vol. 9, Issue 3 (2013), 399-419, esp. 405-407.

[26] This understanding of salvific exclusivism can easily be modified to refer to specific denominations of religions, world 'philosophies', etc. But such a modification is not required for my purposes in this paper.

of three; when there is no Allah save the One Allah. If they desist not from so saying a painful doom will fall on those of them who disbelieve.' (see also 98:6) And, in one particular *hadith*, the Prophet Muhammad is reported to have said: 'By Him in Whose hand is the life of Muhammad, he who amongst the community of Jews or Christians hears about me, but does not affirm his belief in that with which I have been sent and dies in this state (of disbelief), he shall be but one of the denizens of Hell-Fire.'[27] Third, a number of Islamic scholars have maintained that verse 3:85, which affirms that salvation comes through adherence to Islam alone, abrogates those verses that mention the possibility of salvation for non-Muslims, such as Jews and Christians.[28]

A helpful summary account of the traditional understanding of salvific exclusivism in Islam is given by Sheikh Muhammed Salih Al-Munajjid, who writes:

> Allaah has clearly created everything, sent His blessings on His creation, and sent Messengers and Revealed Books to tell them that whoever believes in and worships Him alone, not associating any partner with Him, will enter Paradise, and whoever denies Him, or worships something else besides Him, or takes other gods instead of Him, or says that He has a wife or son, or that the angels are His daughters, or follows a law other than that which He revealed to judge between people in truth, or turns away from His religion, will be doomed in the Hereafter to the eternal punishment of Hell. This is exact justice, and this will be the fate deserved by the one who did not give his Creator His due.[29]

The thought here seems to be that, given the overwhelming evidence and justification for Islamic belief, religious beliefs and practices contrary to Islamic doctrine must constitute deliberate and perverse denial of the truth. And this, in turns, merits eternal punishment in hell – a consequence of God's Justice.

So, here is how the three exclusivist strands in Traditional Islamic Exclusivism – alethic, epistemic and salvific – come together. According

---

[27] Sahih Muslim, Book 1, No. 284, available at: <http://www.hadithcollection.com/sahihmuslim/129-Sahih%20Muslim%20Book%2001.%20Faith/8500-sahih-muslim-book-001-hadith-number-0284.html> (last accessed 3/14/14).

[28] See, for instance Al-Tabari's *The Commentary on The Qur'an*, Volume 1, J. Cooper, (ed.), (Oxford: Oxford University Press, 1989), p. 364.

[29] Sheikh Muhammed Salih Al-Munajjid, 'The Muslim belief concerning the destiny of non-Muslim monotheists in the Hereafter', Islam QA, available at: <http://islamqa.info/en/434> (last accessed 3/15/14).

to traditional Islamic understanding, there is very good evidence and justification for Islamic belief. Such evidence is available to everyone and points to the truth of God's existence and oneness as well as the Prophethood of Muhammad. Consequently, people *should* acknowledge the truth of Islamic belief and reject as false those beliefs that are incompatible with it. Those who fail to do this – all non-Muslims – do so because of their obstinate rejection of it (*kufr*). Such perverse rejection of Islamic belief merits eternal punishment in hell, as an instance of God exercising His Justice.

III. TRADITIONAL ISLAMIC EXCLUSIVISM CRITIQUED

In this section, the central thesis around which I will form my assessment of, and arguments against, Traditional Islamic Exclusivism is that it is highly implausible. I will begin by arguing that Islamic epistemic exclusivism is not tenable. My case against it will consequently open up a line of criticism against Islamic salvific exclusivism. Finally, I will assess what implications my views regarding epistemic and salvific exclusivism have for alethic exclusivism in an Islamic context.

3.1. Problems with Islamic Epistemic Exclusivism

As we have seen, the Islamic epistemic exclusivist affirms two key claims. First, there is (strong) epistemic justification for Islamic belief – belief in the existence and oneness of God, alongside belief in Muhammad as God's Prophet. Second, there is little, if any, epistemic justification for beliefs that are incompatible with Islamic belief. I will consider these two claims in turn, starting with an assessment of the four alleged sources of epistemic justification for Islamic belief.

To begin with, the claim that belief in the existence and oneness of God is our original, natural disposition (*fitrah*) seems, on the face of it, patently false. Perhaps the most obvious objection to it is that it appears to be incompatible with a number of facts about our global religious landscape.[30] Consider, the fact that there are plenty of people in the world who do not believe that there is a God or that there is only one God. As of 2010, there were 500 million Buddhists in the world, people

[30] See 'The Global Religious Landscape', Pew Research Center Forum on Religion & Public Life, available at: <http://www.pewforum.org/2012/12/18/global-religious-landscape-exec/> (last accessed 3/26/14).

who do not believe in the existence of any being like Allah.[31] Consider also the fact that Christianity and Hinduism, religions whose members typically reject the Islamic claim that God is strictly one, collectively had a total of 3.2 billion followers.[32] Indeed, the specific Islamic belief in the existence and oneness of God seems to be held by only a *minority* of the world's population.[33] Based on these data, one can construct the following argument: (1) If belief in the existence and oneness of God is our natural disposition, we would not see such diversity in beliefs about religion. Given that (2) we do, however, see such diversity, one can conclude that (3) belief in the existence and oneness of God is not our natural disposition. How might one reply to this argument? Well, given the validity of its *modus tollens* form, and the obvious truth of the second premise, the only recourse left for someone who wants to rationally resist accepting the argument's conclusion is to reject the first premise.

One way to do this is by arguing that the Islamic understanding of God's existence and oneness *can* be found in all of the various world religions and philosophies, albeit in ways that are subtle. For instance, it may be argued that, underneath layers of theological embellishments and aberrations in their respective religions, Jews, Christians, Sikhs and members of other theistic worldviews do, in fact, believe in the existence and oneness of God. Such a move is, however, problematic, for at least two reasons. First, although the beliefs of some, such as Jews and Christians, can perhaps be accommodated into the Islamic understanding of God's existence and oneness (despite the evidence cited above to the contrary), it will take quite a stretch of imagination to do this for the beliefs of others, such as Buddhists, Jains and adherents of non-theistic religions and philosophies. It is hard to see, for instance, how the Buddhist religion can be described as essentially Muslim, given the Buddhism's explicit denial of theism. Connected to this is the hermeneutical problem of accounting for those portions of the Islamic texts that explicitly repudiate other religions and philosophies because of their denial of God's existence and oneness.[34]

Another way to attempt to reconcile the traditional Islamic understanding of *fitrah* with some of the facts about religious diversity

[31] Ibid.

[32] Ibid.

[33] In the study conducted by the Pew Research Center, the population of Muslims in the world as of 2010 was 1.6 billion, or 23% of the world's population.

[34] See section 2.3 above.

cited above is to try and discredit those beliefs incompatible with Islamic belief by maintaining that they arise (merely) due to social conditioning or sin (or both). The idea here is that although everyone is born in a state of *fitrah*, social conditioning or sin prevents people from responding appropriately to their innate nature; and, it is this responding inappropriately that regrettably results in the kind of religious diversity that we see today. Here is how Sheikh 'Abd al-Rahmaan al-Barraak explains this:

> What it means when it is said that a child is born as a Muslim is that he is born inherently ready, when he reaches the age of discretion, if he is given the choice between Islam and its opposite, to prefer Islam over its opposite and to choose Islam as his religion, so long as there is nothing to prevent him from doing so, such as his whims and desires or tribalism. Following his desires makes him prefer falsehood so that he may attain some share of leadership or wealth, and tribalism or racial pride makes him follow his forefathers or elders, even if they are not following true guidance.[35]

This response essentially maintains that an informed, reasonable rejection of Islamic belief is not possible. Those engaging in religious practices that are incompatible with Islam are motivated primarily by vices such as whims and desires, tribalism, racial pride, etc. It is difficult to avoid seeing this sort of reply as merely dogmatic, especially when it seems that it cannot be supplemented by an argument that does not beg the question.[36] Not only that, but it relies on claims that are simply not plausible. Surely, there *are* plenty of cases involving informed, reasonable rejection of Islamic belief. Let us just consider the theistic component of Islamic belief. In 2009, David Bourget and David Chalmers conducted a survey of 1,972 philosophers in 99 leading departments of philosophy.[37] The found that 72.8% of the philosophers surveyed either accepted or

[35] Islam QA, 'What is the fault of children born in a kaafir environment?', available at: <http://islamqa.info/en/11783> (last accessed 3/27/14).

[36] Here, Karl Barth's famous interview comes to mind, where he declares that Hinduism is a form of unbelief and that Hindus are in rebellion against God. When asked how he knew this, since he had never previously met a Hindu, Barth replied that he knew this 'a priori'. As cited in James Lee Fredericks, *Faith Among Faiths: Christian Theology and Non-Christian Religions* (Mahwah, New Jersey: Paulist Press, 1999), p. 21.

[37] David Bourget and David Chalmers, 'What do philosophers believe?', *Philosophical Studies* (December, 2013), pp. 1-36.

leaned towards atheism.[38] Are we to think, then, that most philosophers today who are atheistic are simply irrational in their rejection of theism, or that their dismissal of theism is motivated simply by whims and desires, tribalism or racial pride? Are there *no* reasonable philosophers, those like the late J.L. Mackie, for instance, who predicate their rejection of God's existence on a careful consideration of the evidence?[39] Answering these questions in the negative not only seems implausible[40] but comes across as the epitome of close-mindedness. Furthermore, there is no independent evidence suggesting that the beliefs of all non-Muslims who do not accept Islamic belief can be explained by appealing to social conditioning or sin. Indeed, there is evidence to the contrary. Consider once again the family of philosophers, many of whom are trained specifically to *avoid* basing their convictions on beliefs that are popular, traditional, politically correct, emotionally appealing, etc. Indeed, in its history, argument forms like appeal to popularity, appeal to tradition, appeal to emotion, etc., have long been recognized as fallacious in logic – a core component in the study of philosophy. To be sure, this does not mean that sociological factors have no influence on the work done by philosophers, but it does suggest that a blanket dismissal of what philosophers believe (e.g. that God does not exist), by appealing to social conditioning, is not plausible.

But what of sin as a viable explanation of the beliefs of non-Muslims, beliefs that are incompatible with Islamic belief? If the reference to sin is taken to refer primarily to moral vices such as pride (a cardinal sin in Islam; see e.g. Qur'an 2:34) that prevent people from recognizing their being in a state of *fitrah*, then we are lead to what I call 'The Gandhi Problem'. This problem essentially concerns reconciling (1) the view that

[38] Ibid. 14.6% accepted or leaned towards theism, while 12.6% accepted or leaned towards the option of 'other'.

[39] Mackie's *The Miracle of Theism* (Oxford: Oxford University Press, 1983) is widely regarded among philosophers as one of the best works in philosophy of religion from an atheistic perspective. In his review of the book, Robert Adams (himself a theist) calls Mackie's work 'a book that will surely rank for some time among the best of philosophical "apologies" for atheism. The standard of exposition and argument is high, and Mackie's fair-mindedness is exemplary. He critically, but in many ways sympathetically, examines an impressive array of theistic arguments', Robert Merrihew Adams, Review of 'The Miracle of Theism, *The Philosophical Review*, Vol. 95, No. 2 (April, 1986), 309-316 (p. 309).

[40] Especially in light of some of the criticisms of the Islamic arguments for the existence of God mentioned below.

non-Muslims who hold beliefs incompatible with Islamic belief do so because of their sin, despite the fact that (2) many such non-Muslims are morally upright and virtuous individuals, people such as Mahatma Gandhi (a Hindu) or Mother Teresa (a Catholic). The more general version of this problem states that non-Muslims who do believe in things contrary to Islamic belief do not appear to be stubbornly resisting the Islamic religion our of pride, selfishness, etc. Rather, they simply do not hold Islamic belief.[41]

So, until further argument can be given to the contrary, the claim that belief in the existence and oneness of God is our original, natural disposition seems untrue. Facts about religious diversity suggest that (Islamic) theism is not epistemically privileged in the way that the concept of *fitrah* would have us believe.

My treatment of The Islamic Anthropological Argument will be brief, simply because, outside of the traditional Islamic account of it, there is no external historical evidence supporting the idea that every nation on earth was sent a messenger calling people to (return to) Islamic monotheism – the primordial core of all world religions and philosophies. The traditional Islamic view is very similar to the hypothesis of 'primitive monotheism' (*Urmonotheismus*), presented and defended by thinkers like Andrew Lang and Wilhelm Schmidt over a century ago.[42] In essence, this hypothesis maintains that all world religions and philosophies that are not monotheistic degraded from a monotheistic *Urreligion*, which consisted of belief in a deity who is eternal, the creator, omniscient and beneficent.[43] One problem with this hypothesis is that we do not have the means to fully assess its historical credibility.[44] Another problem with it is that there are alternative and better hypotheses that account for the origin and development of religion, such as the evolutionary theory. Not only is this particular theory supported by strong evidence, but it also contradicts the notion of an *Urmonotheismus* by placing monotheism at

[41] For further discussion of this particular point, see Aijaz, 'Some Ruminations about Inculpable Non-belief', pp. 12-19.

[42] Mircea Eliade, *The Quest: History and Meaning in Religion*, (Chicago, IL: University of Chicago Press, 1984), pp.23-24.

[43] Ibid., p. 24.

[44] Ibid., p. 25. For a summary account of some of the problems with Lang and Schmidt's work, see E.E. Evans-Pritchard, *Theories of Primitive Religion* (Oxford: Clarendon Press, 1965), pp. 103-105.

the *end* of a long line of religious evolution.[45] These are just two problems facing anyone who wants to defend the Islamic Anthropological Argument along the lines of the hypothesis of an *Urmonotheismus*.

Turning now to the Qur'anic arguments for the existence and oneness of God, these arguments come across as superficial and unconvincing, being susceptible to criticisms that are familiar in the Philosophy of Religion. The Argument from Providence fails because it relies on premises that are at best questionable and its conclusion does not carry much weight in providing evidence for Islamic monotheism. Consider the first premise of this argument, which states that all existing things (found in the world) are suited to our existence. This is clearly false, as a moment's reflection on instances of apparent dysteleology and evil in the world will suggest, e.g. suffering and death caused by diseases, harsh environmental conditions, natural disasters, etc. Such reflection also calls into question the assumption of anthropocentrism that underlies the argument, an assumption that is especially questionable in light of evolutionary biology, according to which human beings have emerged much *later* as a species, being preceded by bacteria, cockroaches and dinosaurs. Be that as it may, even if we were to accept the argument as sound, it would prove, at best, the existence of some sort of intelligent agency behind the creation of the universe. 'But beyond that position', as Hume famously notes, the defender of such an argument 'cannot ascertain one single circumstance, and is left afterwards to fix every point of his theology, by the utmost license of fancy and hypothesis.'[46] The Argument from Invention so obviously begs the question that there is no need to proffer anything further by way of criticism to show that it fails. The first premise of the argument states that all existing things (animals, plants, etc.) are invented. But in order to accept this premise, one already needs to be committed to the truth of the argument's conclusion, which states that there is an Inventor (i.e. God) of every existing thing in the world. The sort of argument for God's oneness that is found in the Qur'an is principally motivated by the thought that the existence of more than one divine being would (or could) result in one divine being frustrating the will of the other(s); and this, in turn, would (or could) result in chaos

[45] Lucius Boraks, *Religions of The East* (Kansas City, MO: Rowman & Littlefield, 1988), p. 9.

[46] David Hume, *Dialogues Concerning Natural Religion* [1779], ed. Martin Bell (London: Penguin, 1990), p. 79.

and disorder in the universe. But, as Richard Swinburne has argued, it not at all clear that the existence of multiple deities necessarily means that they will fail to 'work things out', so to speak, in a way that the activity of one deity does not interfere with what the other(s) do.[47] The Qur'anic arguments for the existence and oneness of God are, therefore, problematic and consequently fail to supply epistemic justification for the belief that (only one) God exists.

It may be objected here that I have discussed only simple or superficial formulations of these arguments and that there are more sophisticated variants out there. For instance, the spirit of The Argument from Providence, as distilled from the Qur'an and formulated by Ibn Rushd, can be found in the more 'updated' Fine-Tuning Argument that rests on recent discoveries about the fundamental physical structure of our universe.[48] Although this is true, it should be noted that the simplistic nature of the Qur'anic arguments for the existence and oneness of God is something that is displayed to its reader from a straightforward reading and interpretation of the Islamic text, as opposed to being, say, a gross oversimplification of more sophisticated arguments contained therein. What is more, merely shifting the focus to more sophisticated versions of the Qur'anic arguments leaves untouched and therefore intact my contention that the original, simple formulations are unconvincing, as I have argued. The import of this last point is significant since a great many traditional Muslims today *do* think that the original, simple formulations of the arguments for God's existence and oneness found in the Qur'an are successful. As W.M. Watt writes, commenting more generally on the traditionalist Muslim's attitude towards faith and reason:

> Traditionalist Muslims today like to claim that 'Islam is a religion based on reason'; but if asked to elaborate this point, they can only produce the sort of philosophical reasoning that was in vogue in the twelfth century ... [T]hey know of no philosophy since Averroes, and are completely unaware of the new challenges to religious belief produced by men like Hume and Feuerbach, not to mention our twentieth-century [now twenty-first century] philosophers.[49]

---

[47] Richard Swinburne, *The Coherence of Theism*, Revised Edition (Oxford: Clarendon Press, 1993), p. 233.

[48] See, for instance, Robin Collins' 'God, Design, and Fine-Tuning', available at: <http://home.messiah.edu/~rcollins/Fine-tuning/Revised%20Version%20of%20Fine-tuning%20for%20anthology.doc> (last accessed 4/1/2014).

Even contemporary defenders of traditionally simple theistic proofs acknowledge (even if implicitly sometimes) that such proofs require more updated and sophisticated defences. Consider as just one example the well-known Kalam Cosmological Argument,[50] formulated by medieval Muslim thinkers like Al-Ghazali. In his *Jerusalem Epistle*, Al-Ghazali states the argument as follows: 'It is self-evident to human reason that there must be a cause for the origination of anything originated. Since the universe is originated it follows that there was a cause for its origination.'[51] William Lane Craig, the foremost contemporary proponent of this argument, concedes that his defence of the argument's premises

> took [him] into extended discussions of such recondite and profound subjects as Cantorian set theory, transfinite arithmetic, the ontological status of sets, the nature of time as tensed or tenseless, Zeno's Paradoxes, Kant's First Antinomy, contemporary Big Bang cosmology (including critiques of alternative or non-standard cosmological theories such as the Steady State model, the Oscillating model, the Vacuum Fluctuation model, and Quantum Gravity models), thermodynamics and physical eschatology, and so on and so forth.[52]

Craig explains that, although the overall logic of the argument is extremely simple, establishing the truth of the premises can be 'a long and complex affair'.[53]

Simplistic arguments for the existence and oneness of God of the sort that are found in the Qur'an are just not convincing, especially in our present day after thinkers like Hume, Kant, Darwin, etc., have criticized the plausibility of theistic arguments and explanations. But what about the more 'sophisticated' theistic proofs in contemporary Philosophy of

---

[49] W.M. Watt, *Islamic Fundamentalism and Modernity* (London: Routledge, 1988), p. 5.

[50] The label is William Lane Craig's. See his 'The Existence of God and The Beginning of The Universe', available at: <http://www.reasonablefaith.org/the-existence-of-god-and-the-beginning-of-the-universe> (last accessed 4/1/2014).

[51] Al-Ghazali, 'The Jerusalem Epistle', a dual language edition with English translation and commentary published as: 'Al-Ghazali's Tract on Dogmatic Theology', edited, translated, annotated and introduced by A. L. Tibawi, *Islamic Quarterly*, IX (1965), 65-122 (p. 98).

[52] William Lane Craig, 'A Swift and Simple Refutation of The *Kalam* Cosmological Argument?', available at: <http://www.reasonablefaith.org/a-swift-and-simple-refutation-of-the-kalam-cosmological-argument> (last accessed 4/1/2014).

[53] Ibid.

Religion? It seems fair to say that whether these proofs are successful is controversial, not only in the Philosophy of Religion, where there is no consensus regarding their success (or failure), but also amongst members of the general public. As Robert McKim observes,

> It is obvious that for every Richard Swinburne who adds up what he thinks to be the relevant evidence and gets a result that supports theism, there is a J.L. Mackie who gets an entirely different result, and, in general, for every theist to whom the facts of her experience appear to confirm that God exists there are apparently equally well qualified nontheists, including members of nontheistic religions, agnostics, and atheists, to whom the facts of their experience have no such significance.[54]

In any case, the Qur'anic arguments for the existence and oneness of God are, in their simple, original forms, unsuccessful. They do not, therefore, provide epistemic justification for Islamic belief.

The 'Inimitability' Argument for the Qur'an's divine origin is, much like The Islamic Anthropological Argument, one that lacks any external, corroborating evidence. The conclusion of this argument is that the Qur'an is a miracle and the key supporting premise is that the text of the Qur'an – conveyed to use by Muhammad, an illiterate man – is inimitable. One problem with this argument is that the concept of a 'miracle' used in the conclusion is not the classical one we find in Hume, for whom a miracle is 'a transgression of a law of nature by a particular volition of the Deity, or by the interposition of some invisible agent'.[55] As Ibn Rushd notes in his *Incoherence of The Incoherence*, the miracle of the Qur'an does not involve 'an interruption in the course of nature ... like the changing of a rod into a serpent, but ... is established by way of perception and consideration for every man who has been or will be till the day of resurrection'.[56] If the idea of violating the laws of nature is built into the definition of a miracle, then it is understandable why a miracle can be seen as providing some evidence for the existence of a God. For, is seems that only something like a God could explain the violation of a law of nature that is confirmed to hold universally. But if an alleged

[54] Robert McKim, *Religious Ambiguity and Religious Diversity* (Oxford: Oxford University Press, 2001) p. 24.

[55] David Hume, 'Of Miracles', *An Enquiry Concerning Human Understanding*, available at: <http://www.bartleby.com/37/3/14.html> (last accessed 4/6/2014).

[56] Ibn Rushd, Simon van den Bergh, (trans.), *Averroes' Tahafut Al-Tahafut* (London: Trustees of the E.J.W. Gibb Memorial, 1978), p. 315.

'miracle' does not violate any laws of nature, then it becomes difficult to ascertain its evidential value. This is basically Swinburne's objection to the Islamic 'miracle' of the Qur'an. 'We have no reason to suspect that that illiterate creative genius cannot guess at truths normally accessible only to the literate,' as Swinburne says, 'or create a new religious style or movement.'[57] This point becomes all the more forceful if we consider the premise of the argument from the Qur'an's inimitability. The claim that the text of the Qur'an is stylistically inimitable is an aesthetic judgment and it is controversial whether such judgments are objective and can be evaluated according to objective criteria.[58] Moreover, it is clearly controversial what *kind* of aesthetic merit in a work counts as a mark of divine inspiration. No traditional Muslim would, for instance, regard the works of Shakespeare, Shelley or Keats as divinely inspired. But why not? What *aesthetic* difference is there between, say, Shelley's *Ozymandias* and one of the short chapters towards the end of the Qur'an, such that the former fails to qualify as divine inspiration while the latter does? It is difficult to see how one can answer this question without sliding into subjective and arbitrary aesthetic judgments. The fact that, outside of the Muslim community, there is no consensus among Arabic-speaking people regarding the stylistic inimitability of the Qur'an further supports this point. I am therefore in agreement with Rippin's description of the argument from the Qur'an's inimitability as a 'dogmatic one, essential to the proof of the status of the text, but one which operates (like many other religious arguments) within the presuppositions of Islam alone'.[59]

In sum, then, none of the four alleged sources of epistemic justification for Islamic belief succeed in showing that the Muslim belief in God's existence and oneness, along with belief in the Prophethood of Muhammad, is epistemically justified. This component of Islamic epistemic exclusivism is therefore false. The Islamic denial of similar epistemic justification for beliefs that are incompatible with Islamic beliefs is not a significant concern, given my purposes in this paper. I will, however, note the following point. If epistemic justification for Islamic belief is *lacking*, then people who are non-Muslim cannot

[57] Richard Swinburne, *Revelation: From Metaphor to Analogy*, Second Edition (Oxford: Oxford University Press, 2007), p. 128.

[58] See Oliver Leaman, *Islamic Aesthetics: An Introduction* (Notre Dame, IN: University of Notre Dame Press, 2004), p. 142.

[59] Rippin, *Muslims: Their Religious Beliefs and Practices*, pp. 39-40.

be regarded as perversely rejecting what they know (or, at least, are epistemically justified in believing) to be true. This observation dovetails with facts in our experience that are evident. There are plenty of people around us who cannot be regarded as obstinately rejecting Islamic belief. More specifically, there are people of *integrity*, people who are, that is, wise, careful and judicious thinkers, intelligent, clever, honest, reflective, serious, etc., who hold beliefs that are incompatible with Islamic belief.[60] The idea that all non-Muslims can be regarded as *kafir*, walking around with their heads held high and stubbornly rejecting Islamic belief, which they know to be true, is simply a theological fantasy.

3.2. Problems with Islamic Salvific Exclusivism

If the criticisms of Islamic epistemic exclusivism in the previous section are correct, then we have arrived at two important conclusions. First, Islamic belief – that is, belief in the existence and oneness of God as well as Muhammad's Prophethood – lacks epistemic justification. Second, and consequently, the Qur'anic stance that non-Muslims are *kafirs*, i.e. perversely rejecting Islamic belief, is implausible. As noted in the preceding section, the concept of the *kafir* is, in the Qur'an, conceptually tied to all cases where a person fails to respond appropriately to Islamic belief. That is, a person who does not practice Islamic belief fails to do so out of a stubborn refusal to acknowledge what he *knows* (or at least has strong epistemic justification for believing) to be true. It is clear from reading the Qur'an that possessing strong epistemic justification is a *necessary* condition for being in a state of *kufr*. If, however, Islamic belief *lacks* epistemic justification (and therefore cannot constitute a case of knowledge either),[61] then it follows that no non-Muslim is in a state of *kufr*. If a non-Muslim cannot be labelled as a *kafir*, then this has serious ramifications for salvific exclusivism in Islam.

As noted previously, according to the traditional Islamic understanding of salvific exclusivism, failure to achieve salvation is treated as equivalent to condemnation in hell. The main reason that the Qur'an cites for people's eternal punishment in hell is that they refused to acknowledge the truth of Islamic belief. Each inhabitant of hell is a *kafir*.

[60] The description of 'people of integrity' here comes from Robert McKim, *Religious Ambiguity and Religious Diversity*, p. 129.

[61] According to most epistemologists, epistemic justification is at least a *necessary* condition for knowledge.

Now, this particular aspect of Islamic salvific exclusivism opens itself up to criticism in at least two different ways.

First, one can criticize whether the traditional understanding of hell in Islam is compatible with God's Compassion and Mercy, which is stressed throughout the Qur'an. Taken at face value, the Qur'anic descriptions of the punishment that awaits those who reject God are, quite frankly, sadistic and horrifying, to say the least. Let me just consider a few verses here. First and foremost, the main description of hell in the Qur'an is a fire whose fuel is men and stones (2:24). God will ensure that the blazing fire of hell never subsides (17:97). It is eternal and inescapable (5:37). As soon as the fire of hell finishes roasting the skin of its inhabitants, God will create new skins for them to continue the cycle of punishment (4:56). People in hell will have their foreheads, flanks and backs branded. They will be given purulent water to drink, which they will be able to sip but not swallow (14:16-17). They may also be given boiling water that will tear their bowels (47:15). The thirsty who ask for drink will be showered with water that is like molten lead, scalding their faces (18:29). Many more such verses can be cited, all of which effectively portray hell as a gruesome torture chamber and the God that created it as a God of (excessive!) retribution. Now, even if each inhabitant of hell is a *kafir*, the Qur'an's description of God's punishment awaiting such a person simply does not seem reconcilable with His Compassion and Mercy. Indeed, even if this point is considered with respect to God's *Justice* alone, where a person did not acknowledge His Creator appropriately and give Him 'His due' as Sheikh Al-Munajjid suggests, it is hard to see how this sort of punishment is *fair*. At most, the *kafir* refuses to acknowledge God and sin for a hundred years or so, an infinitesimal flicker in the timeline of creation from the perspective of an eternal deity. In response, God retaliates disproportionately by condemning the *kafir* to horrific forms of torture, *eternally*. In thinking about all this, one has ample reason for taking seriously the idea that perhaps a literal interpretation of verses that describe hell in the Qur'an is wrong. Perhaps, as Muhammad Ali suggests, the doctrine of hell should understood as *remedial* (by taking mention of the fire of hell as a metaphor for purification) rather than *retributive*.[62]

[62] Muhammad Ali, *The Religion of Islam* (Cairo: National Publication & Printing House, 1967), p. 307.

Rather than pursue this line of criticism, which certainly has force, another simple argument against Islamic salvific exclusivism is to point out that, whatever the Qur'an may say about the fate of the *kafir* in hell and however plausible or not this may be, it is simply inapplicable to the situation of non-Muslims in our world. Why? Because a non-Muslim, as noted, cannot be regarded as a *kafir*. It remains an open question, then, whether non-Muslims who hold beliefs that are incompatible with Islamic belief may be saved. Surely, given the Islamic understanding of God as Gracious, Merciful, Forgiving, etc., such a possibility has to be taken seriously. Rejecting it gives rise to the sort of point that Hick makes several times in his work:

> [T]the basic criticism of both Christian and Muslim [salvific] exclusivism is that it denies by implication that God, the sole creator of the world and of all humanity, is loving, gracious and merciful, and that His love and mercy extend to all humankind. If God is the creator of the entire human race, is it credible that God would set up a system by which hundreds of millions of men, women and children, the majority of the human race, are destined through no fault of their own to eternal torment in hell?[63]

Hick's basic criticism here can be further strengthened by noting that, at least as far as Islamic salvific exclusivism is concerned, non-Muslims are not guilty of *kufr*.

3.3. Islamic Alethic Exclusivism

It should be clear that my criticisms of both Islamic epistemic and salvific exclusivism have no bearing on *alethic* exclusivism in Islam. As explained earlier, exclusivism with respect to truth is both a basic principle of rationality and grounded in a natural interpretation of the Qur'an (see section 2.1). A person can, however, be an alethic exclusivist in an Islamic context while rejecting the traditional Islamic understanding of both epistemic and salvific exclusivism (of course, *which* religious propositions the Muslim alethic exclusivist affirms and denies may change in light of the criticisms proffered in this paper!). Equivocating on the different forms of religious exclusivism can lead to confusion. A clear example of such equivocation and confusion can be found in Gavin D'Costa's criticism of Hick's pluralism. D'Costa argues

[63] John Hick, 'Religious Pluralism and Islam', available at: <http://www.johnhick.org.uk/article11.html> (last accessed 4/6/2014).

that religious pluralism must always be a form of exclusivism and that nothing called pluralism really exists. This is because even the religious pluralist holds some form of truth criteria and anything that falls short of that criteria is not true. In this regard, religious pluralism and exclusivism operate within the same logical structure.[64] But this sort of objection fails to take into consideration the fact that both religious exclusivism and pluralism can be understood with reference to a number of other issues such as epistemic justification and salvation, alongside the question of truth.

IV. CONCLUSION

Religious exclusivism, especially the version that I have termed Traditional Islamic Exclusivism, has a number of philosophical, theological and practical implications. Thinking carefully about the practical implications is especially important today, where the Islamic religion is facing increasing scrutiny because of the religiously inspired, but morally questionable, actions of many Muslims. To consider just one recent example, in 2012 Afghan senators wanted to remove the word 'friendship' in a pact with France based on their belief that Muslims cannot be friends with infidels.[65] Several other examples can be cited where morally questionable or wrong actions by Muslims were inspired by their commitment to Traditional Islamic Exclusivism. The vast majority of Muslims in the world, of course, live their lives peacefully. What they may not recognize, however, is that their actions are not congruent with the traditional Islamic understanding of religious exclusivism. For this, they are often shunned by Imams from the pulpits in the mosques who argue that, in light of Islamic theology, Muslim practice needs to change (e.g. no mingling with non-Muslims, taking them as friends, etc.). But perhaps it is not the religious *practice* of most Muslims that needs to change, but, rather, their traditional understanding of Islamic *theology*, including religious exclusivism.

---

[64] Gavin D'Costa, 'The Impossibility of A Pluralist View of Religions', *Religious Studies*, Vol. 32, Is. 2 (1996), 223-32 (pp. 225f.).

[65] The Telegraph, 'Afghan bid to "de-friend" France', October 3 2012, available at: <http://www.telegraph.co.uk/news/worldnews/asia/afghanistan/9584584/Afghan-bid-to-de-friend-France.html> (last accessed 4/6/2014).