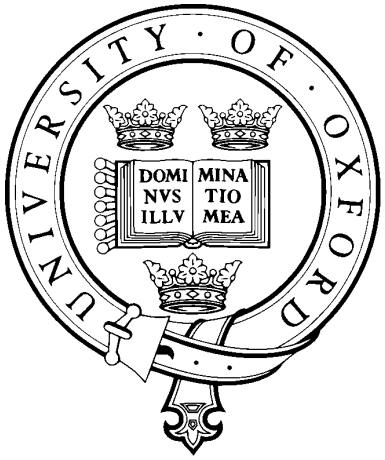
Reformed Epistemology and Naturalistic Explanations of Religious Belief

An inquiry into the epistemological implications of the cognitive science of religion



Max Baker-Hytch
St Cross College
Faculty of Philosophy
University of Oxford

Supervisor
Prof. Brian Leftow

Examiners
Dr Tim Mawson
Professor Duncan Pritchard (Edinburgh)

Thesis for the degree of
Doctor of Philosophy in Philosophy
Trinity Term 2014

To my parents, who have been there for me from day one, and to my wife, Michelle, who has the patience of a saint.

Thesis Abstract

Reformed Epistemology and Naturalistic Explanations of Religious Belief:
An inquiry into the epistemological implications of the cognitive science of religion

Max Baker-Hytch
Faculty of Philosophy
St Cross College

Thesis submitted for the degree of D.Phil in Philosophy
Trinity term 2014
Final word count: 74,857

Reformed Epistemology is an influential view in contemporary philosophy of religion, according to which theistic beliefs that are the product of our native, non-inferential cognitive faculties often constitute knowledge if God exists. My aim in this thesis is to ascertain whether Reformed Epistemology is viable in light of contemporary scientific explanations of the mechanisms of religious belief-formation, especially the Cognitive Science of Religion (CSR). I argue for a qualified "yes." To begin with, I attempt to carefully reconstruct and scrutinise some currently popular "debunking arguments" from CSR's findings, which aim to show that non-inferential religious beliefs are not knowledge, even if true, given the causal origins that CSR ascribes to them. I try to show that in various ways these arguments fail. Subsequently, I attempt to find a better such argument. The strongest debunking argument, I contend, is one that focuses upon the diverse and mutually inconsistent outputs of the religious belief-producing mechanisms described by CSR. However, I go on to argue that even supposing that this argument succeeds in showing that religious beliefs that are partly the product of contingent cultural influences are not knowledge even if true, there remains a body of what I term "core propositions"—propositions concerning the existence of some kind of personal, supernatural creator and moral lawgiver, in which humans are naturally disposed to believe regardless of their particular cultural setting—that can be known if God exists. Finally, I try to show that merely having this core supernaturalistic knowledge would provide someone with the cognitive contact with God that is sufficient for having a personal relationship with God (if God exists), even if only *de re* relationship. I argue, moreover, that God would have positively good reasons for creating a world in which human beliefs about life's most important matters, including religious matters, are significantly dependent upon testimony and hence subject to the ebbs and flows of cultural tides.

Acknowledgements

Firstly, I am deeply grateful to my thesis supervisor, Brian Leftow, for his mentoring over the past four years, and in particular, for his laser-like precision and attention to detail in engaging with my work, for setting an admirable example of rigour, clarity, and intellectual generosity that I have sought (though no doubt not yet succeeded) to emulate, and ultimately, for helping me to grow as a philosopher.

I am also indebted to other friends and fellow philosophers for conversations which have shaped and enriched my thinking about the topics addressed in this thesis and for providing incisive comments on my work. Among those to whom I am especially grateful in this regard are Chris Kyle, Wes Skolits, Tim Mawson, Rik Peels, Trent Dougherty, Kelly Clark, Vince Vitale, and Tim O'Connor. I am grateful, moreover, to members of the New Insights and Directions in Religious Epistemology project for allowing me to participate in the project's weekly reading group and to contribute to the project's termly workshops, and for conversations which have helped me to get to grips with a number of crucial concepts in contemporary epistemology.

I would like to thank the Faculty of Philosophy for giving me the opportunity to write and present a series of lectures based upon this thesis, which gave me the opportunity to test run some of the material in front of an audience of sharp-witted undergraduates. I would also like to thank my funding body, the Arts and Humanities Research Council, whose generous funding via the Block Grant Partnership covered all of my fees as well as a termly stipend.

Finally, I am deeply thankful to those without whose friendship the past four years would have been so much duller and without whose support and encouragement this thesis would never have been finished. I am thinking in particular of Richard Park, Christine Park, Britton Brooks, Michelle Brooks, Luke Martin, James Crocker, Wes Skolits, Chris Kyle, Geoff Dargan, Christian Hofreiter, Ard Louis, Mary Louis, and most of all, my wife, Michelle Baker-Hytch.

Contents

CHAPTER TWO

CHAPTER THREE

Introduction

Reformed Epistemology and the Cognitive Science of Religion

> Epistemology without contact with science becomes an empty scheme. Science without epistemology is—insofar as it is thinkable at all—primitive and muddled.
>
> Albert Einstein

§0.0 Etiology and Epistemology

Philosophical lore has it that for vast stretches of the history of Western philosophy it was held with near unanimity that knowledge is equivalent to justified true belief: belief that is both true and is held on the basis of good epistemic reasons. Then, in 1963, a little-known Midwestern philosopher named Edmund Gettier published a three-page paper that devastated this traditional tripartite account of knowledge with a couple of simple counterexamples, in which the believing subject has a belief that is both true and internally well-justified but which is merely accidentally true and thus, intuitively, is not an instance of knowledge. Epistemology would never be the same again. Whether or not this potted history with which many of us were introduced to epistemology is fully accurate, what

seems hard to deny is that Gettier's contribution brought right to the forefront the way in which factors external to the mind of the believing subject—factors other than just the truth of the belief—are relevant to whether a belief amounts to knowledge. Developments in epistemology in light of Gettier can be seen to bear upon the philosophy of religion in at least two crucial ways.

On the one hand, certain post-Gettier developments in epistemology have made life easier for theistic philosophers of religion who wish to argue that theistic beliefs often have what it takes to count as knowledge. A number of epistemologists in the wake of Gettier's revolution have dropped internal justification altogether from their analyses of knowledge and replaced it with a condition (or conditions) that may be met just in virtue of factors external to the mind of the believing subject.[1] Such epistemologists are known as epistemic externalists.[2] If epistemic externalists are correct in thinking that knowledge doesn't require internalist justification, then knowledge is actually considerably easier to come by

[1] Prominent epistemic externalists, such as Greco (2010), Pritchard (2005), Plantinga (1993a), and Goldman (1967), all put forward their putative conditions as replacements for the internal justification condition on knowledge. According to Greco, "Over the years, externalism about epistemic justification has gained ground, largely because internalist theories seem to entail unacceptable skeptical consequences. More and more, the choice seems to be between (a) externalism and knowledge and (b) internalism and skepticism" (2010: 6).

[2] Notably, epistemic externalism gets glossed in at least a couple of different ways. One way of putting it is that externalists hold that knowledge merely requires true belief that bears some appropriate objective connection to its truth-maker, regardless of whether the subject is aware of that connection. Another way of putting it is that externalism is the view that epistemic justification supervenes solely upon features that are external to the consciousness of the subject, and that knowledge is belief which is both true and justified in this non-traditional sense.

than may have been traditionally assumed.[3] The epistemic credentials of theistic belief have often been attacked on the grounds that there is a paucity of evidence for God's existence, or that at any rate, most theistic believers don't base their theistic beliefs on good evidence even if such evidence is available. But if all that is required for knowledge is that one's belief is true and has the right sort of objection connection to the truth—regardless of whether one is aware of that connection, then an unreflective theistic believer may in fact be able to know that God exists if He in fact does. As we shall see shortly, Alvin Plantinga, the foremost defender of Reformed Epistemology, has skilfully exploited the rise of epistemic externalism with its relaxed requirements for knowledge. Making use of a certain externalist theory of knowledge, Plantinga (2000) has argued that theistic beliefs that are not based upon evidence may, nevertheless, often have what it takes to count as knowledge if God exists.

On the other hand, the way in which the etiology of belief occupies centre stage in post-Gettier epistemology makes for a different set of challenges for theistic philosophers. Even if it is true that theistic beliefs needn't be held on the basis of natural theological inferences in order to count as knowledge, they still need to have the right sort of causal pedigree, in addition to being true. Critics of religion have long speculated about the causes that underlie the seemingly near-universal human religious impulse, and the favourite candidates appear *prima facie* not to be

[3] Robert Pasnau (2013) argues that for much of the history of Western philosophy, epistemology has been concerned with the attainment of a much higher epistemic ideal—one that involves, *inter alia*, the grasping of certain sorts of reasons for one's beliefs.

the sorts of causal sources that would be very likely to produce knowledge. David Hume (1956 [1757]) conjectured that a human tendency to anthropomorphize nature was the culprit. According to Ludwig Feuerbach (1989 [1841]), religion is the result of the way in which we humans feel compelled to project onto the heavens a being who embodies the negation of our weaknesses and limitations. Karl Marx (1964 [1843]: 41-42) proposed that a "perverted world consciousness" is to blame; that the alienation of the masses from the fruits of their labour leads them to console themselves with a spiritual "opium"—an illusory hope of afterlife compensation. And Sigmund Freud (1973 [1927]), whose own genealogy of belief in God seems to have had perhaps the most enduring influence, speculated that a wish-fulfilment mechanism leads people to reach out for a heavenly father figure once they become aware that their earthly fathers cannot protect them against all the world's terrors and against the ultimate threat, death. Whilst these classic genealogies of religious belief have largely been based upon armchair speculation rather than careful empirical investigation, what cannot be avoided in light of Gettier's contribution to epistemology is the way in which discoveries about the etiology of religious belief are in principle highly relevant to its epistemic status. It is the aim of this thesis, then, to take up the task of examining the implications that the latest accounts of the mechanisms of religious belief-formation have for the epistemic status of theistic beliefs. In particular, it is my aim here to attempt a rigorous investigation of the epistemological implications of discoveries about the etiology of theistic beliefs that are not arrived at via sophisticated natural theologi-

cal argumentation, but rather, are arrived at by simply trusting the deliverances of one's native cognitive faculties.

§0.1 Reformed Epistemology

As the name suggests, Reformed Epistemology is a view about the epistemology of religious belief that takes its cues from certain strands of Reformed theological thought.[4] At the heart of Reformed Epistemology is the thought that natural theological argumentation is not needed in order to obtain knowledge of God; that belief in God that is simply the product of one's native, non-inferential cognitive faculties may have what it takes to count as knowledge. In 1983, Plantinga and Nicholas Wolterstorff co-edited a ground-breaking collection of essays under the title *Faith and Rationality*, which in various ways sought to defend this thought.

[4] The name "Reformed Epistemology" may be slightly misleading insofar as it suggests commitment to certain highly controversial soteriological claims, such as the predestinarian theses embodied in the Canons of Dort, when in fact it is neutral on such matters, as far as I can see. Indeed, many contemporary proponents of Reformed Epistemology are libertarians about free will and reject such doctrines as unconditional election, total depravity, and limited atonement. The strand of Reformed thought from which Reformed Epistemology takes its name is the notion that God has endowed human beings with a natural awareness of divinity (*sensus divinitatis*) in virtue of which non-inferential knowledge about God can be had. On this matter, John Calvin, one of the preeminent theologians of the reformation, wrote that "There is within the human mind, and indeed by natural instinct, an awareness of divinity [*sensus divinitatis*]. This we take to be beyond controversy. To prevent anyone from taking refuge in the pretense of ignorance, God himself has implanted in all men a certain understanding of his divine majesty. Ever renewing its memory, he repeatedly sheds fresh drops" (1960 [1536]: 44).

These early defences of Reformed Epistemology tended to focus on showing how theistic belief could satisfy an internalist condition on knowledge despite being held in a non-inferential manner, or put another way, on showing how theistic belief could be *properly basic* with respect to internalist justification.

Plantinga (1983b) took his chief interlocutor in this quest to be the evidentialist objector to theism, who contends both that theistic belief must be based upon good reasons for thinking God exists, if it is to be justified, and that there are in fact no such reasons available. Against this, Plantinga complained that "[P]resumably *some* propositions can properly be believed and accepted without evidence. Well, why not belief in God? Why is it not entirely acceptable, desirable, right, proper, and rational to accept belief in God without any argument or evidence whatever?" (1983b: 39). Plantinga went on to identify what he took to be the crucial assumption underlying the evidentialist demand: namely, the classical foundationalist view of the structure of knowledge, which has been a dominant epistemological paradigm in Western philosophy for the past two millennia. According to classical foundationalism (or, at least, Plantinga's rendering of it), an ideally rational individual's noetic structure consists of a foundation of non-inferentially justified beliefs—properly basic beliefs—upon which is built a super-structure of derived beliefs. Beliefs that can be properly basic include only those beliefs whose truth is either evident to the senses (or incorrigible, on medieval and modern versions) or self-evident. All other sorts of beliefs—theistic beliefs included—must be arrived at via chains of valid deductive or inductive inference that

ultimately terminate in the foundations.[5]

Plantinga levelled two crucial objections against this picture. Firstly, it appears to have fairly serious sceptical consequences. All of us actually believe a much wider variety of propositions in a non-inferential manner than merely those that are self-evident or evident to the senses (or incorrigible): for instance, propositions about the reality of the external world, about other minds, about the past, and so on. And so classical foundationalism appears to have the highly counterintuitive consequence that "enormous quantities of what we all in fact believe are irrational" (1983b: 53). Secondly, classical foundationalism is self-referentially incoherent: the classical foundationalist thesis itself is neither self-evident nor evident to the senses (nor incorrigible), and it is very far from obvious how it could be derived from propositions that are of the aforementioned sorts.[6] For these reasons Plantinga rejected the classical foundationalist criterion for proper basicality and proposed an alternative approach for arriving at such criteria. On this approach, we are to frame hypotheses as to the necessary and sufficient conditions for proper basicality and test them against a sample set of candidate beliefs that seem intuitively to be properly basic, and if necessary, revise the sample set in light of the

[5] What makes this *classical* foundationalism is that it restricts the class of properly basic beliefs in the aforementioned fashion. Foundationalism *simpliciter* is just the thesis that *some* beliefs can be properly basic, and stands in opposition to coherentism, according to which there can be no properly basic beliefs.

[6] For one thing, given that classical foundationalism is a normative thesis, it doesn't seem that it could be derived from any propositions that are evident to the senses.

hypothesis and under the pressure of argument.[7] The result is a kind of epistemic pluralism (though not subjectivism): interlocutors who begin with differing sample sets of properly basic beliefs will arrive at differing criteria for proper basicality, and neither will be able to offer the other a rationally compelling reason for thinking hers is the correct one.[8] As Plantinga writes, "Must my criteria for proper basicality conform to theirs [i.e., to the non-theist's]? Surely not. The Christian community is responsible to its set of examples, not to theirs" (1983b: 77). Importantly, according to the weak foundationalism proposed by Plantinga, whilst properly basic beliefs are (by definition) not based upon evidence, they are not without grounds of a certain sort; grounds having to do with the conditions under which they are formed. Plantinga proposed that even non-inferential theistic beliefs may in fact be grounded in this sense and thus may be properly basic in

[7] Roderick Chisholm (1973) distinguishes two broad ways of doing epistemology: methodism and particularism. The methodist approach begins by drawing up a criterion for justification or knowledge and then proceeding to evaluate cases of belief with the criterion in hand. Particularism moves in the opposite direction: the particularist approach begins with our pre-theoretic judgments about whether given cases of belief are instances of justified belief or knowledge, and on that basis attempts to formulate a criterion for justification or knowledge. Plantinga's approach is closer to particularism, as Chisholm characterises it, though Plantinga's caveat that the sample set should sometimes be revised in light of a hypothesis suggests that it is better to think of his approach as striving for some kind of reflective equilibrium.

[8] Nicholas Wolterstorff (2009) terms the situation that has resulted from the down fall of classical foundationalism *dialogical pluralism*. He writes that "I think it is now widely accepted that the charge of irrationality against theistic beliefs will have to be made on an individual ad hoc basis; if one were to defend the charge by appealing to some general epistemological theory, usually it will be at least as rational for the believer to retain her conviction concerning the acceptability of her theistic belief and reject the theory, as to accept the theory and give up that conviction" (2009: 161).

virtue of being formed under the appropriate sorts of conditions—conditions, for example, involving the experience of guilt, gratitude, danger, a sense of God's presence, a perception of beauty and complexity, and so on. These, Plantinga contends, are the sorts of conditions under which God (if He exists) would intend for human beings to form theistic beliefs. Here, then, we start to see the first hints of the proper functionalist epistemology that comes to characterize Plantinga's later work: a properly basic belief is grounded in virtue of being properly formed; formed in the conditions in which it was *meant* to be formed.

That proper functionalist epistemology comes to its full flower in *Warrant and Proper Function* (1993a). Warrant is that property that turns true belief into knowledge. A warranted belief, then, is one that amounts to knowledge if it is true. According to the theory of warrant that Plantinga sets forth in *Warrant and Proper Function*, a belief is warranted just in case it is subject to no defeaters and is the causal product of a cognitive faculty that is functioning in accordance with a truth-aimed design plan and in the environment for which it was designed (1993a: 46-47). Knowledge consists in true belief that meets these conditions. This proper functionalist account of knowledge becomes crucial to Plantinga's later, defining work in religious epistemology, *Warranted Christian Belief* (2000). In *Warranted Christian Belief*, Plantinga begins by drawing a crucial distinction between two distinct families of objection to Christian theism: according to *de facto* objections, Christian theism is false or probably so; according to *de jure* objections, whether or not Christian theism is true, Christian belief is in some way epistemically subpar.

Much philosophical attention has been given to *de facto* objections—refuting *de facto* objections is, after all, part of the longstanding project of natural theology—but Plantinga's interest is in the significantly less studied *de jure* family of objections. The first part of the book is devoted to trying to get to grips with this sort of objection. After a lengthy survey of various varieties of *de jure* objection, Plantinga concludes that by far the most promising version is the "Freud-and-Marx complaint"—the "F&M complaint" for short (2000: Ch. 5). Epitomised in Freud's claim that religious belief arises from wish-fulfilment and Marx's charge that it is the result of a perverted consciousness of the world, the F&M complaint, according to Plantinga, is that theistic belief *lacks warrant*; lacks what it takes to count as knowledge even if theism is true. And the reason is that on the Freudian and Marxist stories about the causal origins of theistic belief, theistic belief is the product of belief-forming faculties that are either malfunctioning (Marx) or are aimed at a goal other than truth (Freud). So, the F&M complaint, on Plantinga's rendering of it, is that even if theism is true, theistic belief lacks what it takes to count as knowledge because it is not the product of a cognitive faculty that is both truth-aimed and functioning properly. Having set up his target, Plantinga devotes the second part of the book to dismantling it. That is, he seeks to show that theistic belief, and also specifically Christian theistic belief, very likely *does* have what it takes to count as knowledge if it is true. The reason, in short, is that if God exists, he would very likely have created human beings with a cognitive faculty whose aim is the production of true, non-inferential beliefs about God:

> The basic idea…is that there is a kind of faculty or a cognitive mechanism, what Calvin calls a *sensus divinitatis* or sense of divinity, which in a wide variety of circumstances produces in us beliefs about God. These circumstances, we might say, trigger the disposition to form the beliefs in question; they form the occasion on which those beliefs arise. Under these circumstances, we develop or form theistic beliefs—or, rather, these beliefs are formed in us; in the typical case we don't consciously choose to have those beliefs. Instead, we find ourselves with them, just as we find ourselves with perceptual and memory beliefs. (2000: 172-73).

The key idea, then, is that if theism is true, it is very epistemically likely that human beings possess a *sensus divinitatis* and that non-inferential beliefs about God are frequently the product of this faculty, which, in producing non-inferential theistic beliefs, would be working properly in accordance with a divine design plan that is successfully aimed at the truth. That is, it is very epistemically likely given theism that many non-inferential theistic beliefs amount to knowledge. According to Plantinga, then, whether or not theistic beliefs have warrant depends crucially upon whether or not theism is true. So, the *de jure* objection turns out not to be independent of the *de facto* objection. Plantinga's overarching argument might be captured in the following syllogism:

(P1) IF *S*'s belief that *p* is produced by a truth-aimed cognitive faculty that is functioning properly in the environment for which it was designed, THEN *S*'s belief that *p* is (*prima facie*) warranted.

(P2) IF theism is true, THEN very probably, many beliefs about God that are held in the basic way are the product of a truth-aimed cognitive faculty that is functioning properly in the environment for which it was designed.

(C) THEREFORE, IF theism is true, THEN very probably, many beliefs about God that are held in the basic way are (*prima facie*) warranted.

Impressive as Plantinga's case for (C) is, there seem to me to be at least two serious shortcomings with it, and an important part of the contribution made by this thesis is its attempt to rectify these shortcomings.

The first shortcoming I see is this. Plantinga's argument for (C) is thoroughly dependent upon the truth of his proper functionalist account of knowledge. But that account is not widely accepted among contemporary epistemologists. Obviously, truth in philosophical matters is not settled by popular consensus, but insofar as Plantinga's case for (C) is anchored to one particular, not very popular account of knowledge, it is rendered considerably less dialectically effective than it might otherwise have been, just as the dialectical effectiveness of an argument for some interesting claim in metaphysics would be diminished were it made dependent upon the truth of a highly controversial thesis such as universalism about composition. This dialectical worry deepens when we consider that Plantinga's proper functionalist account of knowledge is likely to be fairly objectionable to non-theistic philosophers, who are among the audience for which Plantinga's argument for (C) is presumably intended. It turns out that on Plantinga's account

of knowledge, humans cannot have knowledge unless there exists a God (or someone very much like God) who has consciously established a design plan for human cognition.[9] Whether or not Plantinga is right to think that this is so, it would be better for Reformed Epistemology, all else being equal, if arguments for (C) could be found that don't depend upon so controversial an account of knowledge.[10]

In the two decades that have elapsed since the publication of Plantinga's *Warrant and Proper Function*, two approaches to analysing knowledge have come to the fore, namely, modal reliabilist and virtue-theoretic approaches. A modal reliability condition for knowledge is fundamentally aimed at excluding the sort of epistemically problematic luck that is manifested in Gettier cases. Early steps in the direction of locating a modal reliability condition for knowledge had already been

[9] Plantinga writes that "there is…no denying that intention and conscious design crucially enter into what it is that determines the function(s) of an artifact and whether a given way of functioning is or is not proper functioning" (1993a: 196).

[10] In any case, there do seem to be some powerful reasons for thinking that Plantinga's account of knowledge is inadequate. For one thing, as Linda Zagzebski (1994: 67-69) has shown, that account is not Gettier-proof. Zagzebski presents a case in which Mary, using her 20/20 vision under ideal lighting conditions, sees that someone who looks just like her husband is sitting in the chair, and on that basis correctly believes that her husband is in the room, though in fact the man in the chair is her husband's twin brother and her husband is obscured from view in the far corner of the room. Even though Mary's true belief that her husband is in the room is intuitively not an item of knowledge, it does appear to satisfy Plantinga's conditions for knowledge: it is very plausibly the product of a properly functioning faculty (eyesight) which is truth-aimed and which is operating both in the broad environment for which it was designed (planet earth) and in the specific conditions under which it was designed to operate (close proximity and good lighting). One can quite easily construct other cases that exemplify these features. So, Plantinga's conditions for knowledge appear not to be sufficient.

made several decades ago, for instance, in the conclusive reasons account offered by Fred Dretske (1971), according to which one knows that *p* only if one wouldn't have had one's actual reasons for believing *p* if *p* had been false, as well as in the relevant alternatives account of perceptual knowledge presented by Alvin Goldman (1976), which required that the subject would have avoided false belief in all of the alternative possibilities that are relevant. Subsequently, Robert Nozick's (1981) tracking account of knowledge made use of a counterfactual sensitivity condition, according to which someone's belief that *p* counts as knowledge only if she wouldn't have believed that *p* if *p* had been false. In more recent years, however, the safety condition has come to prominence—a condition which, put very roughly, requires that the subject couldn't easily have been mistaken about the target proposition. An early version of safety was proposed by Ernest Sosa (1999) in the context of a discussion of the problem of radical scepticism. Sosa argued that the intuition that sceptical scenarios cannot be known to be false rests upon the thought that counterfactual sensitivity is required for knowledge, but he contended that in fact only safety, and not sensitivity, is required for knowledge, and that provided that one indeed isn't the victim of a radically sceptical scenario then one may have a safe belief that one isn't. A safety condition features prominently in the knowledge-first epistemology of Timothy Williamson (2000), though Williamson doesn't attempt to offer a reductive analysis of knowledge that features such a condition. Duncan Pritchard (2005, 2009, 2012), on the other hand, has developed a safety condition that does feature in a reductive analysis of

knowledge. Recently, Pritchard (2012) has argued that a successful reductive account of knowledge must in fact answer to two independent "master intuitions" about knowledge: namely, the intuition that knowledge cannot be lucky and the intuition that knowledge arises from cognitive ability. The safety condition takes care of the former, he contends, but a separate, virtue-theoretic condition is required in order to take into account the latter intuition, and in particular, in order to explain why a subject lacks knowledge in a case in which she has a belief that is safe yet whose accuracy is not appropriately connected with her cognitive abilities.[11] The virtue-theoretic condition that he includes in his account of knowledge is close to that advocated by Sosa (2009, 2007) and John Greco (2010).[12] The virtue-theoretic account of knowledge advocated by Greco and Sosa holds that knowledge is a cognitive achievement: a cognitive success that is due to

[11] Pritchard (2009: 40-41; 2012: 260) discusses one such case involving a man named "Temp," who forms a belief about the temperature of the room by looking at a thermometer that he justifiably believes to be working properly, but which is in fact broken and fluctuating randomly. Also unbeknownst to Temp, an unseen agent is constantly adjusting the temperature of the room to ensure that it always matches Temp's belief. It appears, then, that Temp's true belief is safe; that there is no nearby world in which Temp forms a false belief about the temperature by way of the same belief-forming method. And yet it appears very doubtful that Temp really knows the temperature.

[12] Notably, though, Pritchard (2012) employs a weak virtue-theoretic condition in his hybrid anti-luck virtue account of knowledge, requiring only that the subject's cognitive success is in significant part due to her cognitive abilities, rather than that her cognitive success is *primarily* due to her abilities. Pritchard contends that the stronger kind of requirement makes trouble for testimonial knowledge, in that in most cases in which someone acquires a true belief via testimony, her cognitive abilities at best play only a fairly minor role in accounting for her cognitive success. See Pritchard, Millar, and Haddock (2010: §2.6); Kallestrup and Pritchard (2012a).

one's cognitive abilities. Sosa (2007: Ch. 2) likens the acquisition of knowledge to an archer's shot whose accuracy manifests the archer's abilities. On this influential account, a subject knows that *p* only if the accuracy of her belief that *p* is explained by or manifests her stable and reliable cognitive abilities.

I said that Reformed Epistemology would be better off if a good case for (C) could be made that didn't depend upon the truth of Plantinga's proper functionalist account of knowledge. Given the prominence of modal reliabilist and virtue-theoretic accounts in contemporary epistemology, Reformed Epistemology would be considerably better off if it could be shown that if theism is true then theistic beliefs are likely to amount to knowledge by the lights of these sorts of accounts. And that is part of what this thesis will strive to do.

A second shortcoming with Plantinga's case for (C) is that it doesn't take account of the best scientific accounts of the causal mechanisms of religious belief-formation that are currently on offer. The *de jure* objections of Marx and Freud to which Plantinga responds are built upon theories of the origin of religious belief that are considerably out of date and weren't based upon serious empirical investigation in the first place. But there is a branch of contemporary cognitive science that is devoted to studying the cognitive mechanisms underlying religious belief-formation, and religious sceptics have begun to pose *de jure* objections to theistic belief that invoke current cognitive scientific theories about the character of those mechanisms. It is one thing to show that if God exists then non-inferential theistic beliefs probably amount to knowledge, but it is another thing to establish the

stronger conditional claim that if God exists *and* the mechanisms by which non-inferential religious beliefs arise are roughly as current cognitive science alleges them to be, then non-inferential theistic beliefs arising from those mechanisms probably amount to knowledge. Plantinga has not established that the latter conditional is true. But in this thesis I aim to respond to *de jure* objections based upon the latest scientific theories of the origins of religious belief and to establish that latter conditional—or at any rate, a qualified version thereof (more on the qualification shortly).

§0.2 The Cognitive Science of Religion

Contemporary cognitive science takes as its starting point the thesis that the human mind is not a blank slate, but rather, is significantly constrained in its information processing capabilities by certain inherent biases and tendencies. Cognitive science of religion (henceforth CSR) is thus an attempt to understand the cognitive biases and tendencies that are universal to all humans and which play a causal role in the production of religious beliefs. CSR is interested in those features of human cognition that are what Justin Barrett (2004, 2011) terms *maturationally natural*; that is, features that emerge in all developmentally normal human beings without the need for any special training or practice. A key theoretical postulate of contemporary cognitive science is that the operation of the human mind involves two parallel processing systems: the intuitive system and the reflec-

tive system. According to Daniel Kahneman (2003, 2013), one of the foremost proponents of this "dual-processing" view of cognition, the intuitive system can be characterized as quick, effortless, emotion-laden, and automatic, compared with the reflective system, which is slow, deliberate, and involves significant mental effort. Concerning the deliverances of the intuitive system—namely, intuitive (or nonreflective) beliefs—Barrett writes that

> Nonreflective beliefs map closely onto what might be called tacit or intuitive knowledge. They are products of the intuitive system, and are likewise cognitively natural in McCauley's sense of having a high degree of automaticity and requiring little conscious effort to produce. These nonreflective beliefs are representations that we have whether or not we know we have them. They are nonreflective in that they do not require conscious, deliberate, reflective resources to form them. We might never be aware of many of our nonreflective beliefs even though they guide information processing, speech, and other actions… Nonreflective beliefs include such ideas as "Snakes are scary," "Rainbows have bands of color," "People have minds," "I am," "A definite article does not precede a proper name," "Unsupported objects fall," "The sun moves relative to the earth," and "Dogs have puppies." When cognitive scientists talk about what infants or chimpanzees think or know, they generally are referring to nonreflective beliefs. (2011: 48-49).

The operations of the intuitive system and the reflective system are thought to interact in several important ways, including that the deliverances of intuitive

cognition constrain what an individual is liable to endorse upon reflection, that intuitive beliefs are liable to serve as defaults for reflective beliefs in the absence of consciously accessible reasons for rejecting them, and that the intuitive system may be trained through heavy repetition, resulting in what Robert McCauley (2011) calls *practiced naturalness*, as happens when someone learns to drive or play the piano.[13]

The intuitive system, then, comprises various cognitive modules that generate intuitive beliefs. An example of such a cognitive module is the Theory of Mind (ToM) module, which generates intuitive beliefs about the presence of other minds and about the beliefs, desires, and intentions of those putative other minds. CSR postulates that a cluster of such modules working together in concert inclines human beings to form religious beliefs. The current CSR literature seems to indicate that belief of the following propositions is part of what Harvey Whitehouse (2008: 20) has termed the "cross-culturally recurrent religious repertoire":

(a) There is a personal first cause of the universe.

(b) There exists at least one disembodied superhuman agent who possesses beliefs, intentions, and desires, and who:

(c) possesses super-properties such as immortality, epistemic infallibity (i.e. never having false beliefs), and unlimited informational access;

[13] For more on this, see Barrett (2011: Ch. 3).

(d) is interested in and aware of the moral conduct of human beings, including their secret conduct, and is liable to issue rewards and punishments based upon such conduct; and

(e) is responsible for designing apparently purposeful features of the universe.

The cognitive mechanisms by which intuitive beliefs in these propositions are generated are theorized to be as follows. Intuitive belief in (a) appears to be the result of intuitive expectations concerning causation, in conjunction with a preference for postulating persons as causes. Johan De Smedt and Helen De Cruz write that

> [T]he evidence suggests that humans are cognitively predisposed to think that every contingent object or event must have a cause. Thus, the assumption that God is the cause of the universe seems plausible. God is perceived as a good explanation for the universe, because human intuitive notions of causality favour personal over impersonal causes. (2011: 116).

The disposition to believe (b) is thought to be the product of a number of cognitive modules working in concert, especially the Hypersensitive Agency Detection Device (HADD) and the Theory of Mind (ToM) module. HADD tends to generate the intuitive belief that an agent or agents are responsible for patterns or traces of events that resist straightforward mechanistic explanation. According to Kelly Clark, this is so "especially when those traces appear to be (a) teleological or

purposeful, and (b) inexplicable by biological or mechanical means" (2010: 511).[14] The Theory of Mind (ToM) module, responsible for ascribing beliefs, intentions and desires to agents, is liable to generate intuitive beliefs about the mental states of the postulated agent or agents. Added to this, agent concepts which deviate in small but striking ways from our intuitive ontology of agents are particularly liable to be remembered and transmitted, *especially* when those deviations create significant potential for drawing explanatory inferences.[15] According to Barrett, a deity is probably not conceptualised in terms of a human being who breaches various intuitive expectations associated with the category HUMAN, but rather, in terms of the category MIND—a category the activation of which does not by itself automatically generate the intuitive expectation of having a body. Hence, a disembodied mind is in fact not the major deviation from our intuitive ontology that one might think it is. What's more, the property of being everywhere may well be intuitively easy to represent if "conceptualized as having a single location that is well-distributed (such as an enormous cloud or unbounded substance)" (Barrett 2008: 328). Where the concept of a deity *does* deviate from our intuitive ontology is in breaching the expectation that a mind has limited informational access concerning the world and concerning other people's mental states. A mind that can see and hear anything and can read the minds of others is counterintuitive, but not hugely

[14] See also Murray and Goldberg (2010: 186-87). Notably, an evolutionary explanation for the existence of HADD is offered by some CSR researchers, for instance, Scott Atran and Ara Norenzayan, who write that "supernatural agents are readily conjured up because natural selection has trip-wired cognitive schema for agency detection in the face of uncertainty" (2004: 720).

[15] See Boyer (2001); Barrett (2008); Barrett and Lanman (2008).

so (particularly if that mind is represented as occupying every point in space), and the fact that such an entity seems to offer the potential to explain a vast array of otherwise puzzling patterns of events makes it extraordinarily intuitively appealing.[16] Hence intuitive belief in (c).

As for (d), a minded agent with infallible and unlimited informational access to the mental states of others is highly intuitively appealing also for the way it connects with intuitive beliefs about objectively moral obligations.[17] Summarizing Pascal Boyer's theorizing on this topic, Aku Visala writes that "If humans tend to be moral realists and supernatural agents have full access to strategic information, then supernatural agents automatically see the rightness or wrongness of some action" (2011: 72). One strand of current theorizing in CSR postulates that there would likely have been a significant evolutionary advantage for human groups who believed in the existence of morally interested deities who punished cheaters and defectors from the group. Summarizing this hypothesis, Kelly Clark and Justin Winslett write that

> As group size increases, the ability to detect cheaters, who are a threat to the stability of the group, decreases. Because of the threat of such norm-violations to the group, detection of cheaters and their punishment is paramount but very costly… Belief in a supernatural punisher would transfer the "costs" of detection and pun-

[16] For more on this, see Visala (2011: 56-57). And for more about studies conducted on children, which suggest that children find divine properties intuitively easy to conceptualise, see Barrett and Richert (2003).

[17] For an account of the evolutionary origins of intuitive moral realism, see Joyce (2006).

> ishment to (a) nonhuman agent(s), thereby reducing the human costs… In order to prevent cheaters, such agents must have access to strategic information—they must be superknowing about defections or cooperations that are hidden from merely human agents. And they must be believed to efficaciously exercise a kind of moral providence—to reward virtue and to punish vice. So supernatural punishment theory leads to the expectation of high moralizing gods with strategic information who exercise moral providence. (2011: 932).

The mechanisms by which intuitive belief in (e) arises also seem to involve HADD. (e), it should be noted, is not quite as clearly part of the set of pan-cultural, maturationally natural intuitive beliefs as are the previous three. Studies do seem to indicate that children exhibit a marked preference for reasoning about artefacts and organisms in terms of the purpose for which they exist, and there is some evidence that purpose tends to be thought of as being assigned by one or more supernatural agents.[18] Remarking on this, Deborah Kelemen suggests that "Perhaps children's generalized attributions of purpose are, essentially, side effects of a socially intelligent mind that is naturally inclined to privilege intentional explanation and is, therefore, oriented toward explanations characterizing nature as an intentionally designed artefact" (2004: 296). What is particularly striking is that such intuitive teleology doesn't appear to subside when a person reaches adulthood, but rather, experiments indicate that even scientifically educated adults tend to opt for teleological explanations of natural phenomena when subjected to

[18] See Kelemen (2004); Kelemen and DiYanni (2005).

the sorts of time-pressured conditions that prevent careful reflection and thus require participants to call upon the deliverances of intuitive cognition.[19]

Lastly, I must insert the following disclaimer: CSR is a relatively young field, and so, any philosophical conclusions drawn with reference to current CSR are necessarily provisional in nature, and may indeed be undermined as the field continues to develop.

§0.3 Plan of the Thesis

This thesis is divided into four chapters. Broadly speaking, the first half of the thesis attempts to get to grips with *de jure* objections arising from the findings of CSR, and develops what I take to be the strongest such argument. The second half of the thesis argues that even given the success of that argument, there remains a certain body of core religious content that may be known if theism is true, and goes onto draw out the significance of this claim for the problem of divine hiddenness.

Chapter One: Epistemological Debunking Arguments. The first chapter of the thesis attempts to carefully reconstruct and scrutinise some currently popular debunking arguments from CSR's findings—*de jure* objections which aim to undermine the epistemic status of religious beliefs in light of the causal origins that CSR ascribes to them. I outline why such arguments should aim for the following conclusion:

[19] See Kelemen and Rosset (2009).

(D) Whether or not God exists, beliefs about God produced by humans' native cognitive faculties do not constitute knowledge.

Arguments in favour of (D) generally employ a premise asserting a reliability or anti-luck condition for knowledge along with the premise that religious beliefs produced by the non-inferential belief-forming mechanisms described by CSR fail to meet such a condition. While of course theistic beliefs are not knowledge if theism is false, the defender of (D) must also secure the far more interesting claim that even given the truth of theism, theistic beliefs are not knowledge. This chapter explores the difficulties facing the advocate of (D) in establishing the latter claim. In particular, I try to show how it is difficult to run successful arguments for that claim given various properties that theists have traditionally ascribed to God, such as metaphysical necessity and causal efficacy. I conclude that the argument for (D) which involves a counterfactual sensitivity condition for knowledge fails as an independent argument against theism, since it requires certain controversial assumptions about the relative likelihoods of a world containing humans given theism and given atheism—assumptions that cannot legitimately be made in the context of a debunking argument which grants the truth of theism. I argue that a debunking argument that employs a causal condition for knowledge succeeds only if causal determinism is true, since it remains possible and perhaps even probable, given causal indeterminism, that God is causally related to theistic beliefs in the required way. And lastly, I examine a debunking argument that employs an evolutionary-success-in-virtue-of-truth condition for knowledge. I first attempt to

clarify that argument and modify its premises so as to avoid some obviously implausible consequences, but I eventually conclude that a key premise of the argument nonetheless has far wider sceptical implications than its advocates acknowledge.

Chapter Two: A Debunking Argument from Diverse Deliverances. Having tried to show that the most popular debunking arguments currently on offer are flawed in serious ways, I attempt to find a better argument for (D). I suggest, in fact, that the best such argument is one which in its basic form was around long before the advent of CSR, although the theories of CSR may now add some extra weight to one of its premises. CSR postulates that the activity of the various bio-psychological mechanisms involved in the generation of religious beliefs significantly underdetermines the eventual content of those beliefs. CSR thus reinforces what some sceptics have long held: that contingent cultural influences are responsible for determining the specific contours of a person's religious beliefs. The argument in question, then, begins with the claim that had a typical religious believer been subject to substantially different cultural influences, then she would have held substantially different religious beliefs. To begin with, though, I try to show that this argument cannot plausibly be spelled out in terms of the failure of religious beliefs to meet the safety condition for knowledge. A safety condition that is sufficiently strict for the purposes of the argument will also rule out a number of what are intuitively clear cases of knowledge, or so I argue. I suggest that the argument is better spelled out in terms of the failure of religious beliefs to satisfy a

process type reliability condition on knowledge, a condition that requires that the salient process type that produced one's belief doesn't give rise to a high proportion of mutually inconsistent beliefs. If process types are individuated too narrowly, though, such that the specification of the salient type involves particular religious texts or testimony chains, then the argument won't succeed: such a type won't be responsible for enough mutually inconsistent beliefs to render the type unreliable. Rather, process types will have to be individuated at a rather higher level of generality, without making mention of particular texts and testimony chains, so that sufficiently many mutually inconsistent religious belief tokens fall under the salient type. I suggest that the argument can succeed if framed in this way, though I go on to note some potentially unattractive consequences that this wide type approach has vis-à-vis moral and scientific beliefs.

Chapter Three: Knowledge of Core Content. I proceed in this third chapter to argue that even supposing that the debunking argument set out in Chapter 2 succeeds in showing that religious beliefs that are partly the product of contingent cultural influences are not knowledge even if true, there remains a body of what I term "core propositions"—that is, propositions in which humans are naturally disposed to believe regardless of their particular cultural setting—that may be known if theism is true. This core content includes that there is a personal explanation for the cosmos, that the cosmos is providentially ordered, and that there is a moral law with a transcendent personal source. These core propositions are true if the God of Abrahamic monotheism exists, and moreover, knowledge of such propositions will amount

to knowledge *de re* of God if he exists. Much of the chapter is taken up with an attempt to demonstrate that if theism is true, then beliefs of core propositions can satisfy all of the following putative conditions for knowledge: (i) a modal reliability condition, namely, safety; (ii) a virtue-theoretic "cognitive achievement" condition; (iii) a process reliability condition. I consider two distinct ways in which core propositions might be known. Firstly I suggest that core propositions might be known entirely non-inferentially, in roughly the same manner as we arrive at knowledge about other minds or the reality of the external world. Secondly, drawing upon some recent literature on knowledge from falsehood, I also suggest that an adherent of a religious tradition that is significantly mistaken may be able to arrive at knowledge of core propositions via simple deductive inferences from certain false non-core propositions that she has accepted from her tradition.

Chapter Four: A Defence against the Demographic Divine Hiddenness Challenge. Even supposing that knowledge of core religious propositions is available if God exists, the picture I have painted raises some troubling questions. Why would a perfectly loving God be content to permit a situation in which the pushes and pulls of cultural tides play such a significant role in determining the specific contours of an individual's religious beliefs? Why, indeed, would God create a world that is religiously ambiguous to the extent that different human cultures can end up with such diverse ideas about the nature of divine reality? This is the challenge that this final chapter will take up. My aim here is to present a kind of defence for theism in the face of this challenge, which I term the *demographic divine hiddenness challenge*.

That is, the aim of the chapter is to try to make a case for thinking that despite the initial appearances to the contrary, it shouldn't in fact be surprising given God's existence that we would observe the sort of world we do—a world in which human beliefs about many of life's most important matters, including religious matters, are influenced to a significant degree by the norms and consensuses of the communities in which we find ourselves. For one thing, I suggest that the way in which humans are heavily dependent upon testimony and are liable to be influenced by their communities especially on matters that are not readily susceptible to empirical investigation—factors that make an uneven distribution of theistic belief highly likely—actually make possible some great goods that God would not otherwise be able to realize. What's more, I argue that knowing core propositions entails having a cognitive grasp of God that is sufficient for relating personally to God. Hence, even despite an uneven geographic distribution of theistic belief, personal relationship with God—even if only *de re* personal relationship—is very widely available.

Chapter One

Epistemological Debunking Arguments

> According to the general postulate of psychology, there is not a single one of our states of mind, high or low, healthy or morbid that has not some organic process as its condition. Scientific theories are organically conditioned just as much as religious emotions are; and if we only knew the facts intimately enough, we should doubtless see "the liver" determining the dicta of the sturdy atheist as decisively as it does those of the Methodist under conviction anxious about his soul.
>
> William James

§1.0 Introduction

In the Introduction to this thesis we surveyed two significant intellectual projects, one philosophical and one scientific: Reformed Epistemology and the cognitive science of religion (CSR). Reformed Epistemology's central thesis, I suggested, is that belief in God that is simply the product of one's native, non-inferential cognitive faculties likely has what it takes to count as knowledge if God exists. CSR theorists, we saw, have gathered substantial empirical evidence for the existence of a set of cognitive biases, or modules, that predispose developmentally

mature human beings towards religious beliefs. But a number of writers allege that if CSR is roughly correct, the epistemic status of religious beliefs is undermined or "debunked." These writers argue from the findings of CSR, along with premises concerning the necessary conditions for knowledge or justification, that non-inferential theistic beliefs that have been produced by the mechanisms postulated by CSR are not knowledge. The aim of this chapter is to attempt rigorous reconstructions of such arguments and to subject them to thorough scrutiny.

§1.1 Preliminaries: The Nature of a Debunking Argument

An epistemological debunking argument is an argument to a conditional conclusion: that even if someone's belief that *p* is a true belief, she does not *know* that *p*. To use Alvin Plantinga's term, an epistemological debunking argument is a kind of *de jure* objection. By contrast a *de facto* objection, as Plantinga terms it, alleges that theistic belief is in fact false, or probably so.[1] Herman Philipse points out that there have been two basic *de jure* objections in the history of Western philosophy. On the one hand, there is the *evidentialist objection*, "which presupposed that in order to be rational or justified, theistic beliefs have to be argued for on the basis of sufficient evidence, which was assumed to be impossible" (2012: 60). And on the other hand, there is the *irrationalist explanation of religion*, according to which "religious beliefs result from, and are sustained by, psychological or sociological

[1] See Plantinga (2000: vii-xi).

mechanisms that are either not aimed at truth or not well-functioning" (2012: 60). Epistemological debunking arguments from the findings of CSR are *de jure* objections of the latter sort. That sort of *de jure* objection seeks to show that regardless of whether theistic belief is true, something about its causal genesis prevents it from constituting knowledge. Such an objection clearly needs to be able to succeed on the assumption that theism is true, for it is trivial that theistic belief is not knowledge if it is false.[2] And so, this *de jure* objection needs to be logically independent of any *de facto* objection—that is, independent of any argument to the falsity or probable falsity of theism.[3] Notably, this is precisely parallel to the manner in which Sharon Street and Richard Joyce formulate their debunking arguments against moral realism from the findings of evolutionary psychology of morality. They argue that even if there exist mind-independent moral facts, we know nothing about those facts (including whether or not they exist).[4, 5]

But why think this is the best way to interpret the challenge to theistic belief

---

[2] Debunking arguments are usually presented as seeking to establish that $(p \vee \sim p) \supset \sim K_S(p)$, ("whether or not p is true, S doesn't know p"), but this obviously entails both $\sim p \supset \sim K_S(p)$ and $p \supset \sim K_S(p)$, and so both must be established. Since it is trivial given the factivity of knowledge that $\sim p \supset \sim K_S(p)$, it is the latter, $p \supset \sim K_S(p)$, which is of interest throughout this chapter.

[3] See Plantinga (2000: ix).

[4] See Street (2006); Joyce (2006).

[5] Guys Kahane notes that "It is often overlooked…that they [i.e. debunking arguments against metaethical objectivism] presuppose the truth of metaethical objectivism." He goes on: "It's important not to overestimate what debunking arguments can achieve. All they can show is that someone is not justified in believing that p; although they can show the falsity of second-order beliefs about justification, on their own they do nothing to show that p itself is false" (2011: 103, 08).

posed by CSR explanations? Is there not another way to frame that challenge: namely, as the claim that a purely naturalistic explanation of theistic belief is *simpler* than a theistic explanation of theistic belief, and thus more likely to be true? In this vein, Philipse writes that "Since such a secular explanation of the occurring theistic beliefs would be sufficient, there is no need to postulate God, or gods, or any other supernatural entity to account for the existence of theistic beliefs" (2012: 61). Consequently, the argument goes, theism is less likely than atheism to be true and therefore belief in God is an improper response to the available evidence; it is unjustified.

The flaw with this way of construing the problem seems to me to be fatal, however. Suppose one is comparing two rival explanations for a particular episode of jewellery theft in London, one hypothesis involving opportunist thieves and the other involving an entire criminal network masterminded by James Moriarty. Suppose that while both hypotheses are sufficient to explain the happening, the Moriarty hypothesis is less simple as an explanation for this incident. (Notably, there is considerable disagreement about which features of a hypothesis are relevant in determining the relative simplicity of that hypothesis—fewness of types, fewness of tokens, fewness of unobservable theoretical properties, and fewness of adjustable parameters are among the prominent candidates.[6] I don't need to take a stance on that issue here, though. Rather, for present purposes we can simply stipulate that due to whatever consideration is relevant, the Moriarty

[6] See Swinburne (1997: 24-26).

hypothesis is less simple as an explanation for this particular bit of evidence, *viz*., the jewellery theft.) However, one could not reasonably conclude just on that basis—i.e., on the basis of evaluating the hypotheses with respect to one particular incident of jewellery theft—that the Moriarty hypothesis is *all-things-considered* less simple than its rivals. For suppose it were also the case that at the same time that theft occurred, numerous other thefts of exactly the same sort of jewellery piece took place all across the country. When the evidence under consideration is widened to include all those other thefts, the Moriarty hypothesis might well turn out to be the simplest when compared with its rivals, which invoke countless unconnected criminal gangs. Here it is helpful to employ a distinction made by Richard Swinburne between a theory of a narrow region of reality and an overall world theory. As Swinburne puts it,

> In holding simpler theories to be more probable than complex theories, the inquirer is holding it to be more probable that the world as a whole is simple than that it is complex. Hence we should postulate, on grounds of simplicity, as more likely to be true, that theory of a narrow region which makes our overall world theory the simplest for the total data. That theory may not be the simplest theory of the narrow aspect of the world, considered on its own. (1997: 42).

So, let us grant Philipse's claim that the secular explanation is simpler (in whatever sense is relevant) than the theistic one as regards a particular narrow aspect of reality, *viz*., the existence and persistence of religious belief. Still, we need further

argumentation in order to arrive at the conclusion that the secular explanation is a simpler one than theism when the evidence under consideration is widened to include *all* the evidence for which those hypotheses could account. But if Philipse's claim about simplicity is to constitute an independent argument for the probable falsity of theism, then it can't, without begging the question, already assume that theism is less simple as an *overall world theory* than the secular explanation. Rather, the greater simplicity of atheism compared with theism needs to be argued for at the level of the world taken as a whole. So I contend that Philipse's suggestion is not a viable way of construing the challenge to theism posed by CSR's findings.[7]

---

[7] In fact, even granting Philipse that atheism *is* all-things-considered simpler than theism, it is not clear that even that would be sufficient on its own to show that atheism is all-things considered more *probable* than theism. Let us suppose, as Philipse seems to think, that atheism and theism both render the existence and persistence of religious belief equally probable. Still, whether atheism is overall more probable in virtue of its greater simplicity (hence greater intrinsic probability) will depend crucially on whether there are any other data that theism makes more probable than does atheism. Allow me to expand a little. Where "h" stands for a given hypothesis and "e" for a given set of evidence, Bayes theorem states that:

$$\Pr(h/e) = \frac{\Pr(e/h) \cdot \Pr(h)}{\Pr(e)}$$

When comparing two rival hypotheses that are vying to explain the same set of evidence, we can set aside the intrinsic probability of the evidence, Pr(e), which will not vary between the two hypotheses. This leaves us with the likelihood of the evidence given the hypothesis in question, Pr(e/h), which is weighted according to (i.e., multiplied by) the intrinsic probability of that hypothesis, Pr(h). If it is true that simplicity is a good guide to the intrinsic probability of a hypothesis—as Swinburne (1997) has argued at length—then assuming, as we are doing, that atheism is (all-things-considered) a simpler hypothesis than theism, we should think that atheism is more *intrinsically* probable than theism. Now suppose that each of these two hypotheses renders the existence of religious belief just as likely as does the other—that is to say, suppose Pr(religious belief/theism) = Pr(religious belief/atheism). In that case, atheism wins out as the more probable

I am proposing, then, to ascertain whether there is a good argument from CSR's findings to the conclusion that theistic belief fails to count as knowledge on the assumption that theism is true. Only on that assumption is it *non*-trivial to show that theistic belief fails to count as knowledge. The fairest way to proceed seems to be, as Kelly Clark and Dani Rabinowitz put it, to "interpret the CSR objector as making the very interesting claim that despite it being true that God exists, God cannot be known to exist" (2011: 72). And as we shall see, all of the debunking arguments currently on offer are indeed aimed at the claim that even if theism is true, the truth of theistic beliefs would be a matter of knowledge-depriving luck.

One might wonder, though, whether counting as *knowledge* is the only thing that matters here. Even if theistic belief is assumed to be false and thus lacking one of the necessary conditions for knowledge, might it not be an interesting question whether theistic belief possesses other of the prerequisites for knowledge, some of which are (arguably) of value in their own right? Might it be worthwhile to try to

---

explanation *of religious belief*, since the likelihood of a piece of evidence given a hypothesis gets weighted according to the intrinsic probability of the hypothesis in question, and we are granting that atheism is more intrinsically probable than theism in virtue of being simpler—i.e., Pr(religious belief/atheism) × Pr(atheism) is greater than Pr(religious belief/theism) × Pr(theism). But to move from here to the claim that atheism is more probable than theism *simpliciter* (i.e., a more probable explanation for *all* evidence which is in need of explanation) is not legitimate. To secure that conclusion, it would have to be shown that Pr(all evidence/theism) × Pr(theism) is less than Pr(all evidence/atheism) × Pr(atheism). And even though we have granted that Pr(atheism) is greater than Pr(theism), it would still need to be shown that Pr(all evidence/theism) is not sufficiently greater than Pr(all evidence/atheism) to *overcome* the lesser weighting that the former is given. Thus, in order to reach the conclusion that theism is all-things-considered less probable than atheism, a very great deal of other things besides the existence of religious belief need to be considered.

show from CSR's findings that theistic belief lacks one or more of these other properties, on the assumption that theism is false? Let us consider those other properties. Prior to Edmund Gettier's (1963) revolutionary contribution to the analysis of knowledge, it was widely held that knowledge requires that a belief be (i) firmly held, (ii) true, and (iii) strongly internally justified (i.e., supported by the believer's evidence). In light of Gettier's counterexamples to the justified true belief analysis of knowledge, it is very widely agreed that some anti-luck condition must also be met in addition to the aforementioned three.[8] This further condition requires that there obtains some state of affairs external to the mind of the believer in virtue of which the truth of the belief is not a matter of luck, is non-accidental.[9] According to Trenton Merricks, "A belief is accidentally true for one, to a first approximation, if its being true has no relevant connection to the reasons for, or processes involved in, one's holding the belief" (1995: 843). Insisting on this further condition for knowledge, then, is motivated by the thought that the correct analysis of knowledge must block internally justified but accidentally true beliefs, of the sort that characterise Gettier cases, from being knowledge.

If a debunking argument begins with the assumption that theism is false, then which (if any) of the various epistemic properties just mentioned—being believed firmly, being justified, being non-accidentally true—might theistic belief be non-trivially shown to be lacking? Presumably not, given its falsehood, the property of

[8] As I noted in §0.0, a number of epistemologists now deny that (iii) is really necessary for knowledge, and so put forward the anti-luck condition as an alternative to internal justification.

[9] See Swinburne (2001: 199).

being non-accidentally true: it will not be interesting to show that a false belief is unreliable or improperly caused. Nor I think, will it be profitable to try to show from CSR's findings that firmness of belief is lacking in the case of theistic belief, which firmness, in any case, is regarded by many as being of *dis*value.

So how about internal justification? Well, even if it is worthwhile to show, using CSR's findings, that theistic belief is internally unjustified, the most one could establish here is that a theistic believer would lose her internal justification in the event that she becomes aware of the CSR explanation. For unless the CSR explanation is part of her evidence, it cannot be *in virtue of* the CSR explanation that her belief in God fails to be a fitting response to or is made improbable by her evidence (i.e., is internally unjustified). To put it another way, unless she is aware of the CSR explanation then that explanation cannot give her a *defeater* for her internal justification, assuming she has any in the first place.[10] John Pollock (1986: 37-39) introduced a crucial distinction between two sorts of defeater: *rebutting* defeaters and *undercutting* defeaters. One acquires a rebutting defeater for one's belief that *p* if and only if one acquires a belief that *q* where *q* implies (or makes probable) the falsity of *p*. On the other hand, one acquires an undercutting defeater for one's belief that *p* if and only if one acquires a belief that *q* where *q* implies (or makes probable) that the basis on which one believes *p* is not a good one; that

[10] Strictly speaking, the CSR explanation might constitute a *propositional* defeater (i.e., there might be a true proposition about the causal origins of her theistic belief which entails that that belief isn't knowledge) whether or not she is aware of it, but it could only constitute a *mental state defeater* (a defeater for her internal justification) once it became part of her evidence. For more on this distinction, see Bergmann (2006: 153-78).

the basis on which one believes *p* is an unreliable indicator of *p* or fails to be good evidence for *p*. Does the CSR explanation constitute a *rebutting* defeater for theistic belief? Richard Joyce suggests that "Knowledge of a belief's genealogy could show a belief to be false only if the belief implies a contrary genealogical story" (2006: 179). So, does theism imply (or at least, make probable) that theistic beliefs wouldn't have come about in the manner described by CSR? Or to put it another way, is the likelihood of the CSR story low given theism? That is a not a straight-forward question, but it is one that I shall address at length in Chapter 4, where I shall attempt to argue that it doesn't. Still, could the CSR explanation constitute a merely *undercutting* defeater? That is, could the CSR explanation show that the basis on which theism is believed is not suitably truth-conducive? Recall that we are considering the epistemic situation of theistic believers who believe in the *basic way*; *not* on the evidential basis of other beliefs. These are the sorts of believer, after all, with whom Reformed Epistemology is concerned. If such theistic believers have any internal justification in the first place which could be undercut, then it must be the case that they are *prima facie* justified simply on the basis of trusting the deliverances of their cognitive faculties (i.e., it must be the case that theistic belief is *properly* basic for them). An undercutting defeater for such believers would therefore need to consist in a reason to think that the cognitive faculties that produced their theistic beliefs are *not*, after all, trustworthy. In other words, these believers would need to be given a reason for thinking that those faculties are unreliable or are not causally connected in the right way to the truth-makers of

the beliefs they produce. But this brings us right back to the question of whether there indeed is some genetic or causal factor in virtue of which the truth of theistic beliefs resulting from those faculties would be a matter of epistemic luck. So, the question of whether CSR explanations would present a theistic believer with an undercutting defeater for her internalist justification can't be settled without first answering the question of whether theistic beliefs fail to satisfy the anti-luck condition on knowledge.

A second reason for evaluating epistemological debunking arguments as challenges to knowledge rather than internal justification, as Kelly Clark and Dani Rabinowitz note, has to do with the widely accepted (though not entirely uncontroversial) knowledge norm of assertion (KNA). According to KNA, someone is epistemically positioned to assert that *p* only if she *knows* that *p*. Support for this principle seems to derive, among other things, from the intuitive oddness of asserting "*p*, but I don't know that *p*."[11] In light of KNA, Clark and Rabinowitz suggest that "Since theistic belief is often the subject of assertion and, more importantly, influences the way theists go about living their lives, it makes sense to worry about whether theists can know that God exists in light of the CSR research, more than whether theists can rationally believe that God exists" (2011: 71-72).

Thirdly, the present thesis aims at examining the viability of Reformed Epistemology in light of CSR explanations. Since Reformed epistemologists, especially

[11] For an extended defence of KNA, see Williamson (2000: Ch. 11).

Plantinga (2000), claim that God's existence can be *known* in a non-inferential manner, it makes sense to consider whether CSR's findings show that theistic belief produced by such non-inferential mechanisms fails to count as knowledge even if true.

One final preliminary. An epistemological debunking argument can only be made to work if the religious utterances which are being targeted are taken to have cognitive content.[12] And, moreover, that cognitive content needs to be interpreted as having mind-independent truth-values.[13] Unfortunately, I do not have the space here to engage in a defence of a realist interpretation of religious language. The most I can say is that both parties to the debate into which I am about to wade are agreed that theistic beliefs do indeed have mind-independent truth-values, and I hope that is sufficient for present purposes.

§1.1.1 Classifying the arguments

A survey of the current literature suggests that there is not one single epistemological debunking argument that is being considered in the current debate, but

[12] For more on this, see Mason (2010: 777).

[13] As Justin Clarke-Doane helpfully points out, epistemological debunking arguments are conducted on the assumption of four schemas. Where "D" represents an area of discourse, these schemas are: "[D-truth-aptness] Typical D-sentences are truth-apt"; "[D-truth] Some atomic or existentially quantified D-sentences are true"; "[D-independence] The truth-value of D-sentences is relevantly independent of minds and languages"; "[D-literalness] D-sentences should be interpreted literally" (2012: 315-18).

rather, a family of such arguments. The following list is, I think, exhaustive of the various arguments under consideration.[14] Note that the following arguments are all explicitly aimed at the conclusion that even if true, theistic beliefs would be merely accidentally true in virtue of the causal genesis which CSR ascribes to them.

A. *The counterfactual insensitivity argument.* According to this argument, CSR's findings show that human beings would still have theistic beliefs even if God did not exist, and if someone would believe that *p* even if *p* were false, then she does not know that *p*.

B. *The inappropriate causal connections argument.* This argument contends that in order for a belief to constitute knowledge, the truth-makers of the belief must be causally connected to the belief in a certain way. It is alleged that, given the mechanisms of belief-formation described by CSR, theistic beliefs lack this connection to their truth-makers.

C. *The evolutionary unreliability argument.* This argument supposes that if humans can know things about a given domain of reality, then it must have been the case that evolution conferred reproductive benefits upon our ancestors in virtue of the *accuracy* of their beliefs about that domain. It is alleged that theistic beliefs lack this sort of connection with evolutionary fitness.

[14] I have included in this list the four arguments outlined by Murray (2007). Overlapping lists of arguments, also included in my own list, can be found in Leech and Visala (2011: 4), as well as in Barrett (2007: 61-70).

D. *The false positives argument*. This argument alleges that our hypersensitive agency detection device (HADD) produces many false positives in contexts where we can check its outputs against the deliverances of other cognitive faculties, and infers that in religious contexts, where we cannot independently check its outputs, it probably also produces many false positives, and is thus unreliable as a religious belief-forming mechanism.

E. *The mutually inconsistent outputs argument*. According to this argument, the various outputs of humans' religious belief-forming faculties are mutually inconsistent, and therefore those faculties are unreliable.

In what follows, I shall deal with these arguments in order, discussing argument *A* fairly briefly, while devoting more attention to argument *B*, and then a considerable amount of space to argument *C*, which is the argument that has been least scrutinised so far, despite the apparent popularity of the central epistemological principle upon which it relies. I shall in fact omit discussion of arguments *D* and *E* altogether.

Argument *E* is, in my view, so important that I will devote the whole of the next chapter to discussing it. Argument *D* has received the most extensive discussion of any of the five arguments thus far.[15] These discussions have revealed what seem to be strong reasons for thinking that it fails. Firstly, it is highly contested to what extent HADD is causally responsible for producing and sustaining religious

[15] See, e.g.: Murray (2007: 393-94; 2010: 169-71); Clark and Barrett (2011: 26-29); Leech and Visala (2011: 6-7); Barrett (2007: 67-70); Thurow (2013).

beliefs,[16] and the success of the argument turns on HADD having a fairly extensive role. But even if it turns out that HADD does play a very extensive role in producing religious beliefs, this is not sufficient to show that religious beliefs are the result of an unreliable process (and thus fail to count as knowledge), since it is evidently false that HADD is unreliable across *all* contexts where we can independently check its outputs. As Michael Murray points out, "HADD might be unreliable when I hear creaking noises in the abandoned house down the block, but might be quite reliable when I hear a whistled tune in the hall. Is HADD more like the former or the latter when it comes to religious belief?" (2010: 171). It seems, then, that more is needed in order to establish that HADD is unreliable in religious contexts.

§1.2 The Counterfactual Insensitivity Argument

The idea at the heart of this argument is that, given the CSR story, human beings have not evolved cognitive faculties aimed at producing *true* religious beliefs. Considering a very similar claim but whose target is moral beliefs rather than religious, Justin Clarke-Doane notes that

> The claim that we were [evolutionarily] selected to have true moral beliefs has counterfactual force. It implies that had the moral truths been very different, our

[16] Murray (2010: 170) notes the discordant views of several leading CSR theorists on this.

> moral beliefs would have been correspondingly different—that it would have benefited our ancestors to have correspondingly different moral beliefs. Accordingly, the key implication of the claim that we were not selected to have true moral beliefs is the negation of this counterfactual. If we were not selected to have true moral beliefs, then had the moral truths been very different, our moral beliefs would have been the same. (2012: 319).

Hence, *mutatis mutandis*, the claim about evolution's failure to select for true religious beliefs takes the form of the following counterfactual:

(1) IF God did not exist, THEN humans would still believe in God via the non-inferential belief-forming mechanisms described by CSR.

The remainder of the counterfactual argument then runs as follows:

(2) IF *S* would believe that *p* via a mechanism M even if *p* were false, THEN *S*'s belief that *p* is not knowledge when formed via M.

(3) THEREFORE, humans' theistic beliefs are not knowledge when formed via the non-inferential belief-forming mechanisms described by CSR [from (1) and (2)].[17]

[17] Notably, arguments exactly parallel to this have been run against the epistemic status of moral beliefs (where those beliefs are understood as having mind-independent truth-conditions). Michael Ruse, for instance, argues that "You would believe what you do about right and wrong,

Clark and Rabinowitz (2011: 73) note that premise (2) is the sensitivity condition for knowledge developed by Robert Nozick (1981: 172-79), which has been shown to face some considerable difficulties.[18] Joshua Thurow contends, on the other hand, that cognitive reliability does seem to require something in the neighbourhood of (2): "for perception to be reliable, it must be the case that when a certain physical object is not present in normal circumstances, for the most part, one will not believe that the object is present via perception" (2013: 85). But rather than attempt to adjudicate what is a longstanding dispute about whether something like (2) features in the correct analysis of knowledge, I shall focus my scrutiny of this argument instead on premise (1).

As has been noted in the recent literature, however, assessing the truth of (1) is significantly complicated by the fact that it is a conditional whose antecedent is *necessarily* false—at least, given the assumption of theism (which the argument needs to make) and given a classical understanding of theism according to which

---

irrespective of whether or not a 'true' right and wrong existed!" (1986: 254).

[18] One such difficulty is that the sensitivity condition, when conjoined with the anti-sceptical assumption that people know some everyday propositions, seems to require the denial that knowledge is closed under known entailment. This is because when it comes to believing the denials of sceptical hypotheses, e.g., the proposition that "I am not a handless brain in a vat being fed experiences as of having hands," agents will be insensitive because they would continue to believe such propositions in the nearest worlds in which those propositions are false. Yet an agent will generally be sensitive in believing the proposition "I have hands" (if that proposition is true) because the nearest worlds where that proposition is false will be worlds in which, say, she was born without hands or her hands were lost as a result of a horrific accident and she thus doesn't believe she still has any. The upshot is that if sensitivity really is necessary for knowledge, an agent can know she has hands but not know the entailed proposition that she isn't a handless brain-in-a-vat. For more on these issues, see Pritchard (2005: 34-66); Hawthorne (2004: 31-46).

God is a metaphysically necessary being.[19] A widely held view is that counterfactual conditionals with necessarily false antecedents are always true, but trivially so. This view was most notably spelled out by David Lewis and Robert Stalnaker. Lewis writes that

> There is at least some intuitive justification for the decision to make a 'would' counterfactual with an impossible antecedent come out vacuously true. Confronted by an antecedent that is not really an entertainable supposition, one might react by saying, with a shrug: If that were so, anything you like would be true! Further, it seems that a counterfactual in which the antecedent logically implies the consequent ought always to be true; and one sort of impossible antecedent, a self-contradictory one, logically implies any consequent. (2001: 24).

According to Stalnaker, the set of worlds should be thought of as containing one

> absurd world…in which contradictions and all their consequences are true… No other world is possible with respect to it, and it is not possible with respect to any other world… The purpose of [the absurd world] is to allow for an interpretation of 'If *A*, then *B*' in the case where *A* is impossible. (1975: 170).

On this view, if God exists necessarily then the proposition that "if God did not exist, then humans would still believe in God via the non-inferential belief-forming

[19] See Thurow (2013: 85); Clark and Rabinowitz (2011: 75).

mechanisms described by CSR" is true. But so also is the proposition that "if God did not exist, then humans would not still believe in God via the non-inferential belief-forming mechanisms described by CSR." And this is clearly the wrong result as far as the advocate of the counterfactual argument is concerned.

The advocate of this argument seems to have a couple of options here: (i) challenge the notion that God exists necessarily; (ii) employ an interpretation of counterpossibles on which they aren't all trivially true.

§1.2.1 God as a contingent being

Richard Swinburne, somewhat unusually among theistic philosophers, denies that God is a metaphysically necessary being. Rather, according to Swinburne, God is necessary only in the weaker sense that "if He exists at any time He exists at all times" (2004: 96). On this view, God's existing in some possible world entails that there is no time at which He fails to exist in *that* world, but it doesn't entail that God exists in all metaphysically possible worlds. Swinburne calls this "factual necessity," and claims that it is "the strongest kind of necessity compatible with [God]'s being a logically contingent being" (2004: 96).[20] Let us suppose Swinburne is right, and thus that it is a non-trivial matter whether or not it is true that

[20] It is this "factual necessity" that, according to Swinburne, accounts for God's existence needing no explanation despite being logically contingent.

(1) IF God did not exist, THEN humans would still believe in God via the non-inferential belief-forming mechanisms described by CSR.

Lewis and Stalnaker have offered what are very influential accounts of how to evaluate the truth-value of a counterfactual conditional such as (1). Stalnaker's proposal, summed up by Lewis, is that "*A* counterfactually implies *C* at [a world] *w* iff *C* holds at the closest (accessible) *A*-world to *w*, if there is one" (1973: 420). The closeness of worlds to one another is a matter of how similar they are in the relevant respects. As Berit Brogaard and Joe Salerno put it, "The worlds most relevantly similar to *w* will be the ones that share with *w* the most background facts held fixed in the conversational context" (2007: 1). Lewis, however, rejects Stalnaker's view that there is always just *one* closest such world. For many counterfactual situations, he points out, it is nigh-on impossible to non-arbitrarily decide which among several candidates is the uniquely most similar world to the actual world.[21] Rather, according to Lewis's (1973: 422; 2001) proposal, *A* counterfactually implies *C* at a world *w* if and only if *at least one* world where *A* and *C* are true is closer to *w* than any world where *A* is true but *C* is false. Lewis's view seems to many to be the more plausible, and I shall follow his analysis in evaluating (1). Lewis's procedure for ascertaining the truth-value of a proposition "if *p* were false

[21] As Lewis puts it, "Unfortunately, Analysis 1 [Stalnaker's analysis] depends on a thoroughly implausible assumption: that there will never be more than one closest A-world. So fine are the gradations of comparative similarity that despite the infinite number and variety of worlds every tie is broken" (1973: 421).

then *q* would be true" is to alter the truth-value of *p* from true to false and consider what would follow from that alteration, keeping a maximal amount of background facts held fixed.

The question, then, is this: if we set as *false* the truth-value of the proposition "God exists," does anything follow as far as the existence of humans with religious beliefs is concerned? To put it another way, is there an atheistic world (a world where God doesn't exist) where religiously-believing humans exist, which is *more similar* to the actual world than is any atheistic world where religiously-believing humans don't exist? That, of course, depends on whether God actually exists. If God doesn't actually exist, then the actual world will be the uniquely closest atheistic world to itself, and in the actual world humans do indeed have religious beliefs, so (1) will be true. On the other hand, if God does actually exist, then whether the closest worlds where God doesn't exist are ones in which humans still have religious beliefs depends crucially on what *difference* would be made by God's failure to exist. Consider the claim that God's non-existence makes no difference to whether there would be religiously-believing humans (and the all the prerequisites for such beings). This claim implies that all those phenomena—human cognition, complex sentient life, the physical conditions for life, and so on—are just as (logically) probable[22] given God's non-existence as they are given God's existence. Of course, the theist will protest that the closest atheistic worlds to the actual world are vastly different from the actual world. Murray writes that "I, for

[22] Here we are talking about logical probabilities, where the logical probability of *p* given *q* can be understood as the proportion of logically possible *q*-worlds in which *p* holds.

example, don't think there would be a universe if there were no God. I don't think the universe would be fine-tuned for life if there were no God. And I don't think there would be any actual life, believers, human beings, or religion either if there were no God" (2010: 175). The atheist will likely demur that God's existence doesn't (or wouldn't) make anything like that much difference. In essence, their dispute is over the *relative likelihoods*—the likelihood given theism compared with the likelihood given atheism—of a world in which there are humans capable of religious beliefs.

The upshot of all this seems to be the following. Either (a) God doesn't actually exist, in which case (1) is true. Or (b) God does actually exist, in which case *whether* (1) is true depends on whether the closest worlds where God doesn't exist are just like the actual world as regards the existence of human beings capable of having religious beliefs; that is, whether (1) is true depends on whether God's existence makes any difference regarding (i.e., increases the logical probability of) the existence of human beings with religious beliefs and of everything required for such beings to emerge. To secure the truth of (1) via (a), one would need to show that God does not actually exist. But as I argued at the outset of this chapter, an epistemological debunking argument needs to be logically independent of an argument for the falsity of theism (a *de facto* objection), so this isn't a viable option. To secure the truth of (1) via (b), on the other hand, would also require a separate argument: in this case, an argument to show that the likelihood of there being a world in which there are humans capable of religious beliefs, given God's exist-

ence, is not greater than (or not much greater than) the likelihood of such a world given God's *non*-existence.[23] Neither route, then, succeeds in securing (1) independently of other arguments.[24]

Joshua Thurow has objected, however, to any rejoinder to the counterfactual argument according to which "all contingent things necessarily depend upon God for their existence" (2013: 86). From the dependence of all contingent things on God it follows that (1) is false, since necessarily, without God there wouldn't be *any* contingent things, let alone humans with beliefs about God. But Thurow takes issue with this sort of response, claiming that

> [W]e should ignore the fact (if it is a fact) that we depend upon God for our existence when assessing the reliability of P_{BP} [the belief-forming process that generates and sustains theistic beliefs] in forming a belief that some god exists. We do this by asking what we would believe on the basis of P_{BP} if there were no gods and we still existed and used P_{BP} to form a belief about whether there are any gods. (2013: 86-87).

Firstly, though, my own strategy for assessing the counterfactual insensitivity

[23] This perhaps suggests that in order to be interesting, the counterfactual debunking argument not only depends (*arguendo*) upon theism's being true, but also upon theism's being *a priori* probable.

[24] Note that the claim about relative likelihoods (i.e. of a world containing humans with theistic beliefs given theism and given atheism) has direct implications for the overall probability that theism is true, because the likelihood of such a world given theism is a crucial factor (alongside the intrinsic probability of theism) in determining the overall probability of theism. The counterfactual argument is thus tacitly dependent upon a *de facto* objection to theism.

argument does not assume that all contingent things are necessarily dependent on God for their existence. Rather, my thought is that what we say about whether humans exist and have religious beliefs in the closest atheistic worlds to actuality has direct implications for how likely the existence of humans with religious beliefs is given theism. And so the counterfactual argument thusly construed is not independent of other arguments against theism (i.e. arguments concerning relative likelihoods). Secondly, I think Thurow is mistaken anyway in claiming we should simply ignore the question of whether the putative object of the belief in question is required for, or makes significantly more probable, the very existence of that belief. In assessing claims of the form "if X didn't exist you would still believe via a mechanism M that X exists," it seems appropriate to consider what difference would be made to the world by X's failure to exist. If one of the differences it would make is the that the belief-forming mechanism under scrutiny would not even exist (or would probably not exist), then that consideration can't properly be ignored in evaluating the counterfactual claim. For instance, consider the claim that

(B) If, contrary to fact, you did not have a body, you would still believe you had a body (because, e.g., you would be a brain in a vat under the illusion of being embodied).

In assessing the truth of (B), it seems quite proper to consider whether the actual

world in which you have a body and the nearest possible worlds in which you lack a body differ with respect to whether you even have a belief at all about having a body.[25] Presumably the *closest* worlds where you lack a body will be such that you in fact *don't* have any such belief, since a radically sceptical scenario would differ far more from the actual world than would a world where you simply lack a body as a result of having died and been cremated, for example. Using the standard Lewisian procedure to evaluate claims like (B) seems thus to give the right results. By contrast, Thurow's proposal to ignore the way in which the very existence of a belief might depend on (or be made more likely by) its object would deliver the counterintuitive result that (B) is true.

§1.2.2 Non-trivial counterpossibles

As I suggested earlier, the other option for the advocate of the counterfactual argument is to concede the logical necessity of God's existence but to challenge the Stalnaker-Lewis interpretation according to which counterpossibles are vacuously true. Some interesting challenges have in fact been mounted against that interpretation.[26]

Daniel Nolan (1997: 539-40), among others, has pointed out that very many

[25] Indeed, the desire to mitigate the threat posed by such sceptical scenarios to our knowledge of ordinary propositions whilst acknowledging the potency of sceptical hypotheses seems to have been a key motivation behind Nozick's development of the counterfactual sensitivity condition on knowledge. See Nozick (1981: 197-203).

[26] See Nolan (1997); Linda Zagzebski, cited in Wierenga (1998: 93).

philosophical debates concern theories that are either necessarily true or necessarily false. Such debates thus unavoidably involve appeals to what *would* be the case were, counterpossibly, the theory under scrutiny true. But if counterpossibles are all trivially true, then these debates are fundamentally misguided. For any true proposition of the form "if philosophical theory T were true, then C would be the case" (where theory T is either necessarily true or necessarily false, and where C is taken to be an undesirable or implausible consequence of T), there is also a true proposition of the form "if T were true, then C would not be the case"—thus making a nonsense of such reasoning. But since philosophical debates which proceed in this manner are not fundamentally misguided, says Nolan, there is something wrong with the idea that all counterpossibles are trivially true.[27]

Of course, the proponent of the counterfactual argument who wishes to deny the Lewis-Stalnaker interpretation of counterpossibles in order to advance the argument still needs a positive account of its being non-trivially true that

(1) If God did not exist, then humans would still have theistic beliefs via the non-inferential belief-forming mechanisms described by CSR.

[27] Plantinga has also, notably, put forward a very similar challenge to the Lewis-Stalnaker view. As Plantinga puts it, "A dualist might claim that if materialism were true, the content of one's beliefs wouldn't enter the causal chain leading to behaviour; a materialist might claim that if (interactive) dualism were true, an immaterial substance would (implausibly) cause effects in the hard, heavy, massy material world. One of these counterfactuals has an impossible antecedent; both, however, are properly used in the dispute between materialists and dualists" (2011: 339).

Nolan's (1997: 544) semantics for counterpossibles involves basically the same procedure as does the Lewisian approach for ascertaining the truth of counterfactuals: namely, ascertaining whether the consequent is true in all the closest worlds in which the antecedent is true, except that Nolan's approach allows also that *im*possible worlds are among the relevantly close worlds. But for the same reasons as outlined earlier, settling the truth of (1) with reference to the closest *im*possible worlds will still involve assumptions about what difference God's (counterpossible) failure to exist would make to the world. And so this argument will still not run independently of other arguments to the conclusion that God's existence would make no relevant difference.

§1.3 The Inappropriate Causal Connections Argument

Michael Murray summarises this next argument as follows:

> Cognitive psychological accounts of religion can account for the origin of religious belief in a way that makes no reference to and requires no causal connection with supernatural reality. However, properly justified belief requires that the target of the belief be causally connected to the belief itself in certain ways. Since these accounts show us that none of those ways are in fact in play in the origins of religious belief, beliefs so generated are unjustified. (2007: 395).

Indeed, in this vein, Herman Philipse remarks that "Whereas usually positing the

existence or presence of an entity, like a crocodile or chimpanzee, is the best explanation of the fact that someone believes that it exists, this is different in the case of gods" (2012: 60). Philipse seems to have in mind that were God or gods to play a role in explaining religious belief, those entities would need to have made a causal contribution to the emergence of religious belief. An initial statement of a debunking argument along these lines might be as follows:

(4) IF *S*'s belief that *p* is knowledge, THEN the truth-makers of *p* are among the causes of *S*'s belief that *p*.

(5) Given CSR explanations, God is not among the causes of humans' non-inferential theistic beliefs.

(6) THEREFORE, given CSR explanations, humans' non-inferential theistic beliefs are not knowledge [from (4) and (5)].

The claim that (4) expresses a necessary condition for knowledge has been defended most notably by Alvin Goldman (1967). To be sure, Goldman restricted the scope of his causal analysis of knowledge to "knowledge of empirical propositions only" (1967: 357), thus avoiding the problematic implications that the causal condition would have for knowledge of logical and mathematical truths.[28] Despite its not being directly empirically verifiable, the proposition "God exists" can

[28] Philip Hanson writes that "Making the causal theory a theory about empirical knowledge only is a possible if unhappy move to avoid this issue" (1978: 460).

reasonably be construed as an empirical proposition, I think, since it is liable to confirmation and disconfirmation by empirical evidence.[29] Belief in God thus falls within the scope of the causal condition.

The idea behind (5) is that the secular CSR explanation gives a causally sufficient account of the occurrence of religious beliefs, which makes no mention of God or gods. Of course, the CSR account doesn't purport to give a *genuinely* causally sufficient account of religious belief. Rather, CSR identifies mechanisms that make very probable the occurrence of religious belief on the assumption that a whole host of conditions are already satisfied: that there exists a universe capable of supporting life, that there exist the sorts of creatures with the sort of complex cognition capable of sustaining beliefs, and so on. There is immediately, then, a problem with (5). The problem is this: if God exists, then God *is* among the causes of theistic beliefs, even though God may be only an exceptionally remote cause, simply in virtue of being the first cause of all contingent things. And as Clark and Barrett put it, and as I have argued at the outset of this chapter, making this argument dependent upon the non-existence of God means that "we cannot know whether evolutionary psychology of religion undermines belief in God unless we already know there is no God" (2011: 22). To emphasise once again, such an argument, in order to make interesting the conclusion that theistic belief isn't knowledge, needs to be framed upon the assumption that theism is true. And if theism is true then God is at least the first cause of every contingent thing. So

[29] Swinburne (2004: Ch. 6), at any rate, has argued this at length.

premise (5) needs to express a causal requirement involving a particular sort of causal connection between a belief and its object: one which fails to hold between God and theistic beliefs formed in the basic way.

§1.3.1 Tightening the causal condition

The challenge for the proponent of this particular debunking argument is that of finding a causal condition which excludes theistic belief from being knowledge but which isn't so restrictive that it also excludes many other beliefs that would ordinarily be regarded as knowledge. In formulating his causal condition, Goldman himself was somewhat permissive about the sorts of causal connection that satisfy that condition: "I have not closed the list of 'appropriate' causal processes. Leaving the list open is desirable, because there may be some presently controversial causal processes that we may later deem 'appropriate' and, therefore, knowledge-producing" (1967: 370-71). It would not do, for instance, to stipulate that the object of the belief in question must be a spatio-temporally proximate cause of the formation of that belief. While such a stipulation might indeed succeed in undercutting theistic beliefs, it also looks likely to undercut memory beliefs, whose truth-makers often lie distantly in the past, as well as beliefs based upon lengthy testimony chains, and at least some beliefs about other minds.

With that said, any viable causal condition will need to exclude the following sort of case from being an instance of knowledge. Suppose you read a book about

the origins of the universe which asserts that the universe began 14 billion years ago in a Big Bang, and suppose you believe that claim. However, the book is (unbeknownst to you) a work of science fiction, and its author was writing wholly upon the basis of imagination a very long time prior to the mid-twentieth-century discoveries which confirmed the Big Bang theory. Intuitively, it seems you don't *know* on this basis that the universe began with a Big Bang 14 billion years ago. But given that every happening traces ultimately back to the Big Bang, your belief was nonetheless caused by the fact which makes it true.[30] A causal condition that could bar that sort of case from being an instance of knowledge ought, one would think, to be similarly able to bar a belief about God whose only causal connection with God obtains in virtue of God's having been the first cause of all contingent things.

A promising such strategy has been suggested by Robert Shope, who has put forward a causal condition for knowledge according to which

> links in the [causal] chain [must] (a) not involve "excessive generative potential" (roughly, that it not be the case that the beginning of the link could easily have produced some other upshot than the end of the link) and (b) not involve "excessive receptivity" (roughly, that it not be the case that the end of the link could easily have been produced by some other antecedent than the beginning of the link). (2005: 35).

[30] This example is inspired by a somewhat similar one offered by Shope (2005: 34).

Let us consider, then, whether non-inferential theistic beliefs satisfy this causal condition for knowledge. It seems that given the truth of the following proposition, they don't:

(F) God is causally connected to theistic beliefs only in virtue of being the first cause of the universe.

If (F) is true, then there are relevant alternative possibilities in which humans ended up without belief in God: possible worlds where the initial conditions were ever so slightly different but where God was still the universe's first cause. (a) thus seems not to be satisfied. Nor, I think, is (b) satisfied, since it also seems that there are relevant alternative possibilities where something other than God's merely being the universe's first cause was the cause of humans coming to believe in God.[31] (Note that this latter does not imply at all that God does not exist in the nearby possible worlds under consideration; rather, it just asserts that other, more proximate things could have been the salient cause of belief in God.) Given a

[31] Of course, if determinism is true, then the only way things could have turned out differently is if either the laws of nature or the past had been different. But the truth of determinism shouldn't be taken to mean that the beginning of the causal chain *guarantees* the end of that chain, in the sense of guaranteeing that is at issue here. Concerning this issue of determinism and luck, Williamson (2000: 124) suggests that a ball balanced on the tip of a cone in a deterministic universe is not safe from falling just because the ball doesn't fall in any possible worlds with the same laws and history up to that time. Rather, in considering whether the ball might easily have fallen, he suggests, we should consider what happens in those nearby worlds where the initial conditions (not, in this case, the initial conditions of the beginning of the universe, but the initial conditions of the situation in question) were very slightly different.

causal condition on knowledge like that suggested by Shope, then, non-inferential theistic beliefs aren't knowledge if (F) is true. But *whether* (F) is true, I want to suggest, may depend on whether or not causal determinism is true.

Let us understand causal determinism as the following thesis:

> DETERMINISM: The physical laws of the actual world α are such that any world *w* with identical physical laws and an identical history to α prior to a time *t* will also match α at any later time *t**.[32]

If determinism is true of our world, then every event is the causally inevitable consequence of the past and the laws of nature; every event has a fully sufficient set of causes preceding it. This being so, I suggest that it is very likely that (F) is true; it is very epistemically likely that God is causally connected to theistic belief only in virtue of being the first cause of the universe. Why think this? Quite simply, because it is very (epistemically) unlikely, given determinism, that God has made a non-redundant causal contribution to the occurrence of *any* event midway through the universe's history. (I say "unlikely" rather than "impossible" for reasons which will become clear shortly.) For a start, it is logically impossible in a deterministic universe for God to bring about physical events partway through history that deviate from the course that has been causally predetermined by the

[32] This definition of determinism is equivalent to one offered by Plantinga: "Let 'L' be the conjunction of the natural laws, and S(*t*) and S(*t**) be the states of the universe at any times *t* and *t**: then, Necessarily, for any *t* and *t**, if L and S(*t*), then S(*t**)" (2011: 80).

past and the laws of nature. A universe in which such a deviation occurs would by definition not be deterministic. And it is similarly impossible that God establish a deterministic universe with physical laws that underdetermine the occurrence of future events (which underdetermination would leave God some room to make a non-redundant causal contribution partway through history). A universe with such laws would, again, by definition not be deterministic.

But there is one thought about how God could make a non-redundant causal contribution partway through the history of a deterministic world, without causing any deviation from the predetermined course set down by the laws of nature and the past. Plantinga suggests the possibility that God "simultaneously suppresses some natural cause of an event, and then specially causes that event Himself" (2011: 110 fn. 30).[33] Of course, whether this really is a possibility—and to be of relevance here it will need to be a *physical* and not merely metaphysical possibility—depends on whether the suppression of a natural cause by God is consistent with the actual laws of nature, and whether that is so is not an uncontroversial matter.[34] But even supposing it *is* physically possible, do we have any reason to think that God would actually perform such acts of simultaneous suppression and special causation? It is very hard to think of one. What (literally) on earth would by achieved by such acts? It isn't as though by performing them God would

[33] In fact, this was a somewhat throwaway remark made in the context of an attempt to find a good definition of divine intervention, and the abstract possibility in question was raised by Plantinga to illustrate the inadequacy of a certain definition.

[34] For a suggestion in favour of that possibility, see van Inwagen (1988: 217). For a view according to which it isn't a possibility, see Mumford (2009).

thereby be able to guarantee any events that weren't otherwise certain to occur. And in the absence of such a reason, I suggest it is most reasonable to conclude that God's non-redundant causal contribution to a deterministic universe occurs only at the universe's beginning. Thus it is epistemically likely given determinism that (F) is true, and so it is epistemically likely given determinism that non-inferential theistic beliefs are not knowledge by the lights of Shope's causal condition.

Now, there is of course an important sense in which, according to classical theism, God's sustaining activity is necessary at every single moment to hold the universe in continued being, whether or not the universe is deterministic. Thomists have sometimes expressed this thought in terms of God's being the *primary cause* of every event; the ontological ground that holds up the entire web of efficient causation (or *secondary causation*).[35] I submit, though, that God's moment-by-moment sustaining of the world constitutes a fundamentally different kind of causal relationship than the one with which causal theories of knowledge are concerned, and so this consideration won't mitigate the success of the debunking argument under consideration.

§1.3.2 Indeterminism

If causal determinism is false, though, there may be good reason, I want to claim,

[35] See, e.g., Drees (1996: 105-06).

for thinking that (F) is false; for thinking that God is causally connected (as an efficient cause) with theistic beliefs and not merely in virtue of being the universe's first cause. The thought is this. If the universe is causally indeterministic, there will be very many causally possible futures open to it at any given time.[36] Given just the indeterministic laws of nature and the state of the universe during its early moments, the objective probability that the universe would eventually give birth to life on earth and to humans capable of complex cognition was presumably some way short of 1. Even shortly prior to the occurrence of the genetic mutations in virtue of which humans came to have a natural disposition towards supernatural belief, the occurrence of those mutations was not wholly inevitable, given indeterminism. So, if God wanted to *guarantee* that humans come to have some sorts of beliefs about Him—if God wanted to increase the objective probability from, say, highly probable to *certain*, so as to eliminate any risk whatsoever—then God would need to act as a causal agent during those crucial junctures in cosmic history.[37] But even given God's proximate causal activity in establishing in early humans those dispositions towards religious belief, would both clauses of Shope's causal

[36] Concerning this causal underdetermination, van Inwagen writes that "Suppose…that a rolling billiard ball exhibits the position-momentum and time-energy uncertainties predicted by Heisenberg. For an object as big as a billiard ball, these uncertainties are miniscule indeed. Nevertheless, the capacity of the collisions of rolling spheres to magnify slight deviations is astounding: Within a few minutes the arrangement of balls on two tables [which were initially as close to being perfect duplicates of one another as is physically possible] will be entirely different" (1988: 225).

[37] This is not at all to suggest that such proximate divine activity needs to be invoked to explain these occurrences, whose physical probability may well have been quite high anyway.

condition be satisfied? That is somewhat hard to say. As regards clause (a), could humans easily have ended up failing to believe in God? Presumably they couldn't, if that proximate activity on God's part was deliberately aimed at guaranteeing the emergence of religious beliefs. (a) is thus met. As regards clause (b), though, could it easily have been the case that humans came to have religious beliefs *without* such proximate causal involvement on God's part? Here the answer is less clear. It depends upon how high was the objective probability, just prior to God's acting, that humans would wind up with religious beliefs anyway. If that probability was already fairly high then it seems that humans could quite easily have had religious beliefs anyway, and so (b) wouldn't be met.

With that said, the situation in which God is involved as a proximate cause in establishing the human religious belief-forming disposition seems to be relevantly analogous to a case which is, intuitively, an instance of knowledge. Murray (2007: 396) has suggested the following such case. Jones is a candidate in a local election and wishes to make himself known to voters, but doesn't want to go to the trouble of making phone calls to every last one of them. So he sets up an automatic dialling machine which makes calls to homes throughout his constituency informing voters of who he is and what his policies are. Suppose further that these automated phone calls are the only source of information most of the voters have about Jones and his policies. It seems that they do *know* the things they learn by way of these automated calls set up by Jones, and by analogy, that humans know at least some of the things they believe about God if those things are true and if

the dispositions towards those beliefs were established in the aforementioned manner. What's more, given causal indeterminism it may be that God acts in an even more direct manner to produce beliefs about Himself in humans: namely, by acting as the proximate cause of religious experiences, which—at least, if William Alston's (1991) lengthy defence is near the mark—can be a source of non-inferential warrant for theistic beliefs.[38]

§1.4 The Evolutionary Unreliability Argument

The clearest statement of this next argument has been offered by John Wilkins and Paul Griffiths.[39] That said, the epistemological principle upon which it relies—or something like it—seems to have been first articulated by W.V. Quine (1969: 69-90) in his "Epistemology Naturalized." Summarizing Quine's proposal, E.J. Lowe writes that "any kind of knowledge attainable by human beings, including anything that might deserve to be called 'metaphysical' knowledge, must be compatible with our status as a kind of natural creature—in fact, a species of animal—that has arisen through wholly natural processes of biological evolution" (2002: 5). Taking their cues from Quine, Griffiths and Wilkins contend that a necessary condition for our being epistemically reliable concerning a given do-

[38] Of course, a religious experience is not a sufficient condition for belief in God; humans would also need to possess some mechanism by way of which a religious experience leads to belief in God.

[39] See Griffiths and Wilkins (2010); Wilkins and Griffiths (2013).

main of reality is that reproductive advantages accrued to our evolutionary ancestors as a consequence of the *accuracy* of their beliefs about that domain. And they are by no means alone among contemporary philosophers in taking this view. William Ramsey argues in favour of what he calls "evolutionary reliabilism," an epistemological position which appears to be very close to that advocated by Griffiths and Wilkins. Ramsey writes,

> Putting things very roughly (and, no doubt, too simplistically), evolutionary reliabilism claims that cognitive mechanisms come to be adaptive, in part, by generating desires that correspond with reproductive fitness (e.g., a desire to reproduce, to stay alive, etc.) and accurate representations that enable the organism to get around and satisfy those desires. (2002: 19).

Take visual perception, for example, which is held by evolutionary reliabilists to be a paradigm instance of a cognitive faculty favoured by natural selection for its accuracy.[40] The idea, in short, is that the accuracy of a creature's beliefs about medium-sized physical objects in its spatio-temporal vicinity will significantly

[40] In a recent and frequently-cited paper arguing that evolutionary explanations debunk moral beliefs (at least, where the truth-makers for such beliefs are taken to be mind-independent moral facts) Sharon Street uses beliefs about medium-sized physical objects as prime examples of beliefs which cannot be evolutionarily explained without reference to their truth-makers: "In order to explain why it proved advantageous to form judgments about the presence of fires, predators, and cliffs, one will need to posit in one's best explanations that there were indeed fires, predators, and cliffs, which it proved quite useful to be aware of, given that one could be burned by them, eaten by them, or could plummet over them" (2006: 160-61, fn. 35).

affect its chances of reproducing, and thus evolution will in the long run favour creatures with more accurate visual perception systems. So, suppose that when running towards the edge of a cliff with a sheer drop into an ocean full of lethal submerged rocks, Fred believes, based upon his being appeared to by a depthless expanse of green below and blue above, that there is a calm lake ahead of him. Fred, it seems safe to say, will not be long for this world. George, by contrast, approaching the same cliff edge, believes, based on his being appeared to in such a way that where the green and blue meet, the blue looks much further away than the green, that there must be a very big drop at the end of the green. So long as creatures like Fred and George generally strive to avoid falling hundreds of feet,[41] George's prospects look rather better than Fred's. George will very likely live longer than Fred and so will have more opportunities to reproduce, thus passing on at a greater frequency his genes, which include those genes that dispose his offspring to have depth perception which is more accurate than Fred's. George's offspring and those of creatures with the same visual abilities as George (all else being equal) will, in the coming generations, increasingly outnumber the offspring of Fred and of creatures with the same visual abilities as Fred (all else being equal). After hundreds of generations of this evolutionary winnowing, the population of creatures of which Fred and George were once members can reasonably be

[41] Ramsey (2002: 19) contends that this is a tendency which evolution would favour, though Plantinga (2002) has questioned this. Plantinga has claimed that false beliefs conjoined with misguided desires can just as successfully lead to adaptive behaviour as can true beliefs conjoined with survival-aimed desires. For a response to Plantinga's argument, see Law (2011).

expected to consist overwhelmingly of creatures whose depth perception is, on the whole, reliable. This, at any rate, seems to be the picture that Griffiths and Wilkins have in mind. And not only they, but also notable advocates of the evolutionary debunking argument against moral realism, such as Sharon Street, seem to have in view a similar notion. Thus Street: "A creature obviously can't run into such [mind-independent moral] truths or fall over them or be eaten by them. In what way would it have promoted the reproductive success of ancestors to grasp them?" (2006: 130-31).

Such a connection between the accuracy of a given type of beliefs and the reproductive success of our ancestors is necessary, according to Griffiths and Wilkins, in order for it to be reasonable to accept and act upon beliefs of that type. It is fair, I think, to read them as claiming that we should conclude that humans are reliable in forming a given type of beliefs only if this sort of connection with evolutionary fitness obtains.[42] We should conclude that humans are unreliable with respect to a given type of belief, then, if the reproductive advantages which accrued to our ancestors in virtue of having the type of beliefs in question did so *regardless* of the accuracy of those beliefs (this needs some careful qualifying, of course, to which I shall come shortly). Griffiths and Wilkins thus offer the following epistemological principle: "The *X* facts [must be] related to the evolutionary success of *X* beliefs in such a way that it is reasonable to accept and act on *X*

[42] The sort of reliability that is at issue here, I take it, is the sort of reliability with which process reliabilism is concerned: that is, with the ratio of true beliefs to false beliefs that are the product of a given type of belief-forming process.

beliefs produced by our evolved cognitive faculties" (2010: 2). They call this the "Milvian bridge" principle, so named because the Emperor Constantine's victory over Maxentius at the battle of Milvian bridge in 312AD was, according to a certain tradition, due to the truth of the Christian faith under whose banner he fought. Pragmatic success on the battlefield was thus the result of true belief. Analogously, if we are to reasonably conclude that our beliefs about some domain are sufficiently reliably formed to count as knowledge, then it needs to be the case that the success of our ancestors in the evolutionary struggle was due (in part, at least) to the accuracy of their beliefs about that domain.

Having outlined their Milvian bridge principle, Griffiths and Wilkins then proceed to argue that religious beliefs fail to measure up to it:

> No Milvian bridge is available for religious beliefs, because none of the leading accounts of the evolution of religious belief makes any reference to the truth or falsity of those beliefs when explaining their effects on reproductive fitness. This is true both of evolutionary theories which explain religion as a side-effect of other adaptations, and those which explain it as an adaptation in its own right… David Sloan Wilson's account shows the same lack of concern with the specific content of religious beliefs; arguing that religion evolved through multi-level selection, driven by the benefits which social cohesion and prosocial behavior provide at the level of the group… Theories which explain the evolution of religion as a side-effect are equally undiscriminating. (2013: 8).

§1.4.1 Refining the argument

Let us see, then, if we can state this debunking argument in a more precise manner. In order to do that, a clearer statement of the Milvian bridge principle is needed, I think. If it is to serve as a premise in an argument to the conclusion that we should think we are unreliable with respect to religious beliefs, this principle needs to specify the relation beliefs are required to have to their truth-makers, otherwise it is simply unclear whether or not religious beliefs measure up. The foregoing discussion of evolutionary reliabilism suggests how we ought to specify the relation. A first approximation might be the following:

(M) IF *S* is reliable with respect to X beliefs, THEN X beliefs contributed to the evolutionary success of *S*'s ancestors in virtue of their accuracy at representing X facts.

Taking (M) as the major premise, the remainder of the argument might run as follows:

(8) Theistic beliefs *didn't* contribute to the evolutionary success of our ancestors in virtue of their accuracy at representing the facts that would make theistic beliefs true.

(9) Humans are not reliable with respect to theistic beliefs.

(10) IF *S* is not reliable with respect to X beliefs, THEN *S*'s X beliefs are not knowledge.

(11) Humans' theistic beliefs are not knowledge.

One need not look too closely to see, however, that (M) is not up to the job. It implies the unreliability of beliefs that made no contribution to evolutionary success, but this will include many sorts of beliefs which Griffiths and Wilkins do not intend to debunk. It seems safe to say that scientific beliefs, for instance, made no contribution to the reproductive success of the vast majority of our ancestors. And even if, as seems rather doubtful, beliefs about scientific theories made a positive contribution to the reproductive fecundity of our most recent forbears, a few centuries is nowhere near long enough for natural selection to effect a very substantial remoulding of the faculties we use for forming scientific beliefs. But as far as Griffiths and Wilkins are concerned, the debunking of scientific beliefs is the wrong result.

Now, Griffiths and Wilkins do think that they can keep scientific belief safe from being debunked. They hold that there is what they call "an *indirect* Milvian bridge" which connects the accuracy of scientific belief to evolutionary success in an indirect manner—indirect because it involves the faculty of "commonsense" (which is alleged to have a direct connection with evolutionary success in virtue of accuracy) acting as the ultimate arbiter of the results of science. They write,

> The reasons we have to think that our scientific conclusions are correct and that the methods we use to reach them are reliable are simply the data and arguments which scientists give for their conclusions, and for their methodological innovations. Ultimately, these have to stand up to the same commonsense scrutiny as any other addition to our beliefs. (2013: 6).

Whether or not commonsense really can keep our scientific theorizing in check, that Griffiths and Wilkins take themselves to have to offer such a defence of scientific beliefs suggests the following modification of the Milvian bridge principle:

> (M*) IF *S* is reliable with respect to X beliefs, THEN *either* X beliefs contributed to the evolutionary success of *S*'s ancestors in virtue of their accuracy at representing X facts *or* X beliefs are held by *S* on the basis of Y beliefs, where Y beliefs contributed to the evolutionary success of *S*'s ancestors in virtue of their accuracy at representing Y facts.

One problem with (M*), so far as it is intended to serve as the major premise in an argument that debunks religious beliefs, is that it allows in principle that religious beliefs held on the basis of other sorts of beliefs (beliefs which *are* adaptive in virtue of accuracy), may be able to meet Griffiths and Wilkins criterion. Griffiths and Wilkins acknowledge, indeed, that "if there are independent reasons for religious belief, their cogency is not removed by the fact that religious beliefs have evolu-

tionary explanations" (2013: 10-11). This is not a problem, however, if the target for debunking is theistic beliefs held in the *basic* way—that is, beliefs about God which are not held on the basis of any other beliefs. And basic theistic beliefs are precisely the sorts of beliefs whose epistemic status Reformed Epistemologists are concerned to vindicate.[43] I suggest, then, that (M*) fairly captures the intent of Griffiths and Wilkins.

The following seems to me to be an adequate reconstruction of the debunking argument put forward by Griffiths and Wilkins:

(7) IF *S* is reliable with respect to X beliefs, THEN *either* X beliefs contributed to the evolutionary success of *S*'s ancestors in virtue of their accuracy at representing X facts *or* X beliefs are held by *S* on the basis of Y beliefs, where Y beliefs contributed to the evolutionary success of *S*'s ancestors in virtue of their accuracy at representing Y facts.

(8) Theistic beliefs which are held in the basic way *neither* contributed to the evolutionary success of our ancestors in virtue of their accuracy at representing the facts that would make them true *nor* are they held on the basis of other beliefs which did contribute to the evolutionary success of our ancestors in virtue of their accuracy at representing the facts which would make them true.

(9) Humans are not reliable with respect to theistic beliefs that are held in the

[43] See, e.g., Plantinga (2000, 1983b); Clark (1990); Wolterstorff (1983).

basic way.

(10) IF *S* is not reliable with respect to X beliefs, THEN *S*'s X beliefs are not knowledge.

(11) Humans' theistic beliefs that are held in the basic way are not knowledge.

§1.4.2 Dialectical considerations

This debunking argument is put forward in a dialectical situation in which Griffiths' and Wilkins' Milvian bridge principle, premise (7), is questionable, or so I shall now argue. The dialectical situation here seems to be this: Griffiths and Wilkins intend to attack the epistemic status of theistic beliefs without assuming anything about their truth-value—that is, they intend to show that theistic beliefs are not sufficiently reliably formed to count as knowledge, but without assuming that theism is false or probably false. I have already argued that this is the only viable way to construe an epistemological debunking argument from CSR's findings. But assuming (for the sake of the argument) the truth of theism, is the Milvian bridge principle likely to be true? We arrived earlier at the following formulation of that principle:

(M*) IF *S* is reliable with respect to X beliefs, THEN *either* X beliefs contributed to the evolutionary success of *S*'s ancestors in virtue of their accuracy at representing X facts *or* X beliefs are held by *S* on the basis of Y beliefs, where

> Y beliefs contributed to the evolutionary success of *S*'s ancestors in virtue of their accuracy at representing Y facts.

We first need to consider how to interpret the conditional here. If we interpret it as a strict conditional, it will be true if and only if the consequent is true in every metaphysically possible world where the antecedent is true. That is a tall order indeed. But it is widely held, nonetheless, that epistemological principles ought to be necessarily true, in the *metaphysical* sense of necessity. As Matthias Steup notes, "A proper analysis of knowledge should at least be a necessary truth. Consequently, hypothetical thought experiments provide appropriate test cases for various analyses" (2010). But if the entailment relation in (M*) is one of metaphysical necessity, then (M*) is false. This is because there are metaphysically possible creatures able to reliably form beliefs of a sort that neither contributed in any way to their ancestors' evolutionary fitness nor are held on the basis of beliefs that did so contribute. Some such creatures did not even evolve at all but were generated instantaneously, or if they did evolve, had their evolved cognitive faculties augmented such that they reliably form beliefs about matters wholly remote from anything having to do with life in their ecological niche. For such creatures, reliable belief-formation about at least some propositions in no way depends upon a connection between evolutionary success and true belief. In the metaphysically possible worlds in which such creatures live, the antecedent of (M*) is true while the consequent is false, and so, interpreted as a strict conditional, (M*) is false.

On the other hand, taken as a merely material conditional, (M*) is true just as long as in the *actual* world (or, perhaps, in a nearby region of worlds centred on the actual world) the consequent is true whenever the antecedent is true. Let us grant Griffiths and Wilkins that it is acceptable to put forward an epistemological principle that involves only this weaker sort of entailment. Taken as a material conditional, (M*) asserts that it is true just in the actual world (or perhaps also nearby worlds) that believing subjects reliably form beliefs about a given domain only when the aforementioned connection with evolutionary fitness obtains. It asserts that humans in the actual world reliably form beliefs only about matters that impinged on the reproductive prospects of our ancestors or that can be checked with reference to the sorts of beliefs that did. But is it legitimate to assert this in the current dialectical context, in which the truth of theism is being assumed (if only for the sake of the argument)? That depends crucially on whether theism makes it probable that humans would be capable of knowing and thus reliably forming beliefs about things that were irrelevant to their survival and that can't be checked against the sorts of beliefs that *are* thus relevant.

Let us understand theism according to what might be called the classical definition: there exists necessarily an immaterial person, God, who is omnipotent, omniscient, and perfectly good, and who created and sustains the universe. Now, it is hard, I think, to see how theism so-defined entails the falsity of (M*). It is consistent with theism, on this definition, that God refrained from in any way steering the evolutionary processes that gave rise to human beings and their

cognitive capacities. And so it is consistent with this sort of theism that evolved human cognitive faculties ended up being reliable only with respect to the sorts of things about which true (or approximately true) belief was needed for survival and reproduction. Still, perhaps the probability of (M*) given theism is low. Here we are talking about logical probability, where the logical probability of *p* conditional upon *q* can be understood as the proportion of logically possible *q*-worlds in which *p* is true.[44] To say that the logical probability of (M*) given theism is low is thus to say that there is only a small proportion of theistic worlds in which (M*) is true. But why think that is the case? Given the definition of theism we are currently working with, it is unobvious to me why we should expect that God would give humans cognitive abilities that are reliable *beyond* the sorts of homely matters that impinge upon survival and reproduction. This definition of theism, let us call it *bare theism*, does not imply that humans are special in some way that all other organisms are not, and bare theism gives us no reason to think that all those other organisms have cognitive abilities that go beyond what is needed to get by in their respective ecological niches. (After all, we generally take it that other animals have

[44] Michael Rea has objected to this way of understanding objective probability. He writes "It is tempting to try to understand objective probability in terms of proportions of possible worlds. Intuitively, the probability of a fair coin landing heads is 0.5 because half of the *relevant possibilities*—i.e. half of the worlds that share our history up to the time of the flip, our natural laws, and so on—are worlds in which heads appears… But, of course, this intuitive suggestions will not do; for (presumably) there are infinitely many possible worlds, and it makes no sense to talk about halves, or majorities, or any other proportion of an infinite set" (2002: 180). This is a somewhat puzzling objection, though. It seems perfectly intelligible to speak of dividing a circle in half or into thirds—or whatever proportions one wishes, even though the circle consists of an infinite number of points.

the ability to know—if that is the right term—only what is needed in order to feed, fight, flee, and fornicate,[45] and theism so-defined does not imply that humans are any different.) So not only does bare theism *not* entail the falsity of (M*), it doesn't even render (M*) improbable. It is thus legitimate in the dialectical context in which Griffiths and Wilkins put forward their argument to assert (M*), provided the target for debunking is basic belief of the *bare* theistic variety.

That said, things might look differently given a thicker definition of theism. Plantinga writes that

> If [theism] is true, then there is, indeed, such a person as God, a person who has created us in his image (so that we resemble him, among other things, in having the capacity for knowledge), who loves us, who desires that we know and love him, and who is such that it is our end and good to know and love him. But if these things are so, then he would of course intend that we be able to be aware of his presence and to know something about him. (2000: 188-89).

Plantinga is working with an understanding that is undoubtedly richer than the bare theism we were previously considering. Let us call Plantinga's definition *theism*+. In working with this understanding I think Plantinga intends to capture the common core that is shared by the major monotheistic faiths—the Abrahamic faiths. That common core includes that humans are significant to God in a way that would lead us to expect that humans have some cognitive abilities which go

[45] These are Patricia Churchland's (1987: 548) famous "Four F's."

beyond what was needed to get by in our ancestral environment; in particular, having at least a minimal sort of natural knowledge of God.[46] Theism+ does indeed, then, make (M*) improbable. And if the target for debunking is a belief about the sort of God entailed by Plantinga's thicker concept of theism, then in attacking the epistemic status of such a belief (on the assumption of its truth), it seems highly dubious to assert an epistemological principle that is likely to be false given the truth of the targeted belief.

§1.4.3 Self-referential problems for the Milvian bridge principle

Michael Murray (2007: 398) has argued that the Milvian bridge principle should be rejected as it debunks many more sorts of beliefs than Griffiths and Wilkins intend it too—it cuts too wide a swathe. In essence, Murray's claim echoes a key premise of Plantinga's evolutionary argument against naturalism (EAAN):[47] that there is no reason to think that beliefs about *any* domain are evolutionarily selected for in virtue of their accuracy; the demand imposed by the Milvian bridge principle leads to radical scepticism. But I shall now put forward an objection to Griffiths and Wilkins that relies upon nothing quite as bold as that premise of Plantinga's EAAN. I shall argue that while it may allow that perceptual, and perhaps even scientific beliefs (as Griffiths and Wilkins contend) are reliably

[46] Indeed, that he so intends is suggested by his use elsewhere of the term theism to denote what is "central to the great monotheistic religions—Christianity, Judaism, Islam" (Plantinga 2011: ix).

[47] See Plantinga (2011: 308-49; 2002; 1993a: 216-38).

formed, the Milvian bridge principle renders epistemically suspect the very sort of reasoning required to adjudicate the truth of that principle. We do not reliably form beliefs, by the principle's own lights, about whether that principle is true, or so I shall now argue.

Thomas Crisp (2011) has recently offered an argument similar to the EAAN, but relying on weaker premises, to the conclusion that it is unlikely that our faculties for *philosophical reasoning* are reliable given that our cognition is the product of unguided atheistic evolution. He deploys this thesis in an attempt to undercut the justification for key premises in atheistic arguments from evil: premises which are "recondite philosophical claims" such as that "No reason would justify an extremely powerful, extremely wise, wholly good creator of all things in permitting the existence of evils of the sort contained in our world" (2011: 126). Crisp's argument for this thesis of unreliability is that

> Given unguided A&E [unguided atheistic evolution], the main explanation why our evolutionary ancestors evolved cognitive faculties of the sort they did was that those faculties were adaptive: useful for feeding, flying, fighting, reproducing and so forth. But why would cognitive faculties selected for their success at those tasks have required reliability with respect to abstruse philosophical matters, matters quite unconnected to the concerns of everyday life? From a fitness point of view, such cognitive reliability seems wholly unnecessary. (2011: 116-17).

Crisp notes the possibility that "perhaps reliability on recondite philosophical

matters far removed from the everyday concerns of life is a 'spandrel'—a non-adaptive by-product of some adaptively selected trait," but he considers that "since at this point anyway, no one has been able to produce a convincing argument that we should *expect* reliability on recondite philosophical matters as a by-product… perhaps the best thing to say about the probability…is that it is inscrutable" (2011: 117).

Now, it is my contention that the belief that "(M*) expresses a true epistemological principle" is indeed a philosophical belief—and a recondite one at that. But doesn't this assume something that has been shown to be thoroughly problematic: namely, that a token belief can sensibly be subsumed under just one type? The problem of generality, as it is known, is held by critics of reliabilism to be a fatal difficulty for that theory.[48] That problem is normally framed in terms of how to assign a token belief-forming process to a type of process, rather than in terms of how to assign a token belief to a type of belief, but the problem applies, *mutatis mutandis*, to the latter. Roughly, the problem is alleged to be that any given token belief-forming process can be subsumed under a number of different types of process that vary considerably in their reliability, and there is no non-arbitrary way to decide which type is the relevant one. Consider my belief, formed at 5:00 PM on a misty December 12th 2012 in poor lighting conditions, that there is a sparrow sitting atop the nearby lamppost. The token process that produced this belief can be characterised at one extreme as an instance of perception as such—a

[48] See Swinburne (2001: 13-20).

generally fairly reliable type of process; or as an instance of visual perception in misty and poorly lit conditions—a considerably less reliable type of process; or at the opposite extreme, as an instance of *my* visual perception in misty and poorly lit conditions at 5:00 PM on December 12th 2012—a type of process which has just one token, and so whose reliability depends on whether the belief is true.[49] The problem in the current context is that of ascertaining the *relevant type* of belief to which belongs the token belief that "(M*) is a true epistemological principle." In fact, for my present purposes this isn't really a problem. If there really is no non-arbitrary way to assign tokens to types, as critics of reliabilism allege, then Griffiths and Wilkins' key premise, (M*), is sunk anyway and there is no need to further criticise it. On the other hand, Griffiths and Wilkins clearly assume that for the purposes of their argument they *can* sensibly assign token beliefs to broad types: commonsense, perception, moral, religious, scientific, and so on. And so it seems fair in the present dialectical context that in attacking their key epistemological principle, I too make this assumption. And on the basis of the broad categorizations that Griffiths and Wilkins employ, the broad type of belief to which belongs the belief that "(M*) is a true epistemological principle" would seem to be that of *philosophical beliefs*.[50]

But why think that philosophical beliefs are unreliable by the lights of Griffiths and Wilkins' criterion, (M*)? Recall that according to that criterion,

[49] For a classic exposition of the generality problem, see Feldman (1995).

[50] A narrower type, such as *epistemological beliefs*, would also suffice for the argument I am about to make.

(M*) IF *S* is reliable with respect to X beliefs, THEN *either* X beliefs contributed to the evolutionary success of *S*'s ancestors in virtue of their accuracy at representing X facts *or* X beliefs are held by *S* on the basis of Y beliefs, where Y beliefs contributed to the evolutionary success of *S*'s ancestors in virtue of their accuracy at representing Y facts.

I stated the criterion in this way to reflect Griffiths' and Wilkins' contention that there are two sorts of "Milvian bridge" which can connect reliability to evolutionary fitness: a direct one and an indirect one. A direct Milvian bridge obtains just if a given type of beliefs contributed to the reproductive success of our ancestors in virtue of their accuracy; just if our ancestors were evolutionarily selected to have true beliefs of this sort. Now, I contend that this direct connection clearly doesn't hold for philosophical beliefs. It is extremely hard to see what reproductive advantage there would have been for our ancestors in having true beliefs about matters such as the correct analyses of knowledge and justification, the correct normative ethical theory and the correct applications thereof,[51] the correct theories about the nature of time, identity, composition, substance, necessity, consciousness, and so on. Their having utterly false beliefs on all these matters seems completely consistent with their being highly successful in performing the

[51] To be sure, evolutionary psychology of morality is indicating that certain intuitive moral belief-forming dispositions were better suited than others for the ancestral environment (see Joyce 2006), but this in no way shows that our ancestors needed to have true beliefs about which *overarching normative theory* best accounted for all their considered moral judgements.

"four F's." A small caveat is in order here, though. Clearly there are a *few* philosophical beliefs which it would be positively maladaptive to hold, for instance: the belief that "it is right and proper to believe logically contradictory propositions," or that "rationality demands that beliefs be formed on the basis of cracking open fortune cookies," or that "the world around us is an artificial simulation and the best way to return to reality is to eat poisonous mushrooms." It is not too hard to imagine how having such beliefs as these could rather quickly get a creature into trouble (though it seems pretty unlikely that our Pleistocene ancestors thought too much about such matters). But all this shows is that evolution would select against philosophical beliefs that lead to biologically costly errors in navigating the everyday world. And only a tiny proportion of all possible philosophical beliefs are such that they would lead to this sort of maladaptive behaviour. Most philosophical beliefs don't lead to any behaviour at all—not, at least, behaviour of the sort which is likely to have any substantial impact upon a creature's reproductive prospects (heavy drinking may be an exception). Suppose, for instance, there is certain philosophical problem, and humans have come up with a variety of theories purporting to offer a solution to the puzzle—be that problem the search for the correct analysis of knowledge, or the laws of nature, or whatever. Call those theories T_1, T_2, T_3,…, T_n. Suppose T_1 and T_2, if believed, will lead to maladaptive behaviour, as with the belief that "rationality consists in believing logically contradictory propositions." On the other hand, all the many remaining theories, if believed, will make no significant difference at all as regards a crea-

ture's reproductive prospects. Belief in T_1 and T_2 (or the tendency to believe in T_1 and T_2) will eventually be winnowed away by natural selection. But beyond that, natural selection can get no further "grip," as it were, on the truth-values of the remaining theories. It is simply unable to adjudicate between the many theories that are left standing. And that indeed seems to be just the situation as regards the array of competing theories that are currently on offer for any given philosophical problem: belief in any of one them won't make the least bit of difference to reproductive success compared with belief in any other.

But is there what Griffiths and Wilkins call an "indirect Milvian bridge" connecting philosophical beliefs to evolutionary success? An indirect Milvian bridge obtains, recall, just if the type of belief in question can in some way or other be checked against beliefs of a type that *did* have a direct connection to evolutionary success in virtue of accuracy. I doubt that there is even this indirect connection, however, when it comes to philosophical beliefs. Consider the other sorts of beliefs on the basis of which philosophical beliefs are formed and held, or against which philosophical beliefs can be checked. Beliefs about the proper rules of inference are certainly among them. Using the sort of broad categorization of beliefs that Griffiths and Wilkins seem to employ, we might call these *logical beliefs*.[52] It seems quite plausible to think that logical beliefs (at any rate, beliefs about first-order

[52] One might protest that beliefs about the proper rules of inference are themselves philosophical beliefs, and are the sort of beliefs that it seems likely that evolution would select for in virtue of their accuracy. This may well be so, but nothing turns on it for the purposes of my argument, for no one would suggest that most philosophical beliefs are *merely* beliefs about the proper rules of inference.

logic) do indeed have a direct connection with evolutionary fitness in virtue of their accuracy. Believing that *modus ponens* and *modus tollens* are correct rules of inference would seem to stand a creature in much better stead, as regards evolutionary fitness, than the beliefs that affirming the consequent or denying the antecedent are correct ways to reason.[53] (Or, strictly speaking, perhaps what would be evolutionarily favoured is correct *use* of the rules.) But this isn't nearly sufficient to establish that there is an indirect Milvian bridge for philosophical beliefs. If philosophy were about nothing more than applying the correct rules of inference, there surely wouldn't exist anything like the kind of expert disagreement that there in fact is concerning philosophical matters.

In addition to logical beliefs, of course, the formation of philosophical beliefs also involves what we might call *conceptual beliefs*. Crisp writes that a conceptual belief is a belief concerning "a sentence S [that is] *a priori* knowable by anyone with an adequate understanding of the concepts expressed by the constituent terms of S. So take "no married male is a bachelor." It is a conceptual truth in our sense: anyone with an adequate grasp on the concepts bachelor, married, and male can see a priori that it's true" (2011: 126). As with logical beliefs, it seems fairly plausible that evolution would select in favour of accurate conceptual beliefs. As Crisp points out, "Plausibly, reliability of this sort would have been fitness enhancing for our ancestors. A creature with the concepts of, say, danger or food, who lacks a reliable grip on the conditions under which such concepts are satisfied

[53] For examples, see Law (2011: 247-49).

or on the logical relations holding among related concepts will fare poorly in the evolutionary game" (2011: 127).

But being reliable with respect to both logical beliefs and conceptual beliefs is still not sufficient for being reliable with respect to philosophical beliefs. A person could be fully able to apply the correct rules of inference and have an adequate grasp on all the concepts relevant to the questions of (say) whether the doctrine of temporal parts is true, whether knowledge requires a sensitivity condition, whether there are true propositions about future contingents, whether preference-satisfaction is the sole axis of value—and yet still be unable to settle those questions. Philosophers can and do routinely point out (what they take to be) the untoward logical consequences of one another's positions. But unless those consequences include an outright logical contradiction, doing so is hardly ever sufficient to conclusively settle a philosophical dispute. And even when it *is* shown that a logical contradiction is implied by a certain set of philosophical propositions, there is still frequently scope for reasonable disagreement about *which* of those propositions is thereby shown to be false.[54]

[54] A good example of such a situation is the debate over *existentialism*, the doctrine that singular propositions entail the existence of the objects that they are about—even when such propositions deny the existence of the objects that they are about. Plantinga (1983a) has mounted what he takes to be a reductio of this claim. He argues that the proposition (entailed by existentialism) that "necessarily, if the proposition that 'Socrates does not exist' is true, then Socrates exists," when conjoined with a number of very plausible propositions concerning the nature of truth and necessity, leads to the contradictory proposition that "possibly, Socrates exists and Socrates does not exist." However, Jeff speaks (2012) has argued that it is not existentialism that is thereby refuted, but another of the propositions with which Plantinga conjoins it for his reductio, namely,

Of course, just occasionally a philosophical theory will imply a proposition whose truth can be checked against *scientific* beliefs, which may have an indirect "Milvian bridge" connection with evolutionary fitness (assuming Griffiths and Wilkins are right that scientific beliefs are fully accountable to commonsense and perception). Some have claimed, not uncontroversially, that presentism, the doctrine that only the present time is real, implies things which are inconsistent with the empirically-confirmed theory of Special Relativity: namely, that there is a privileged frame of spatio-temporal reference in relation to which things are either past, present, or future.[55] Even if this is right, though, I submit that such cases are very rare. The great majority of philosophical questions for which scientific data are relevant are such that the data significantly underdetermines which of a variety of philosophical theories is the correct one.

In addition to invoking the findings of science, applying the correct rules of inference, and adequately grasping the relevant concepts, it seems that philosophical reasoning frequently requires judgments to be made concerning the intuitive plausibility of the propositions which are implied by the philosophical theory under scrutiny. In particular, intuitions are in very many cases indispensable for adjudicating whether a certain concept or property is instanced in a certain situation. Should we think these intuitions have a connection to evolutionary

---

the proposition that "if possibly Socrates exists, then the proposition that Socrates exists is possibly true." The question at issue in this debate is this: for *which* premise is there a reductio? (My thanks to Eddy Chen for drawing my attention to this example.)

[55] For more on this debate, see Lowe (2002: 322).

fitness in virtue of their accuracy? That seems doubtful. As Michael Rea puts it,

> [M]y intuitions report that each of the following propositions is necessarily true: that no two material things occupy the same place at the same time, that removing one grain from a heap never leaves one with a nonheap, that if Socrates exists he is able to think, that no one is morally responsible for something they were forced to do, that torturing small children is morally wrong, that it is possibly true that a necessary being exists, and that the axioms of the S5 modal system are true. There are many who share my intuitions on these matters, but there are at least as many who do not; and from an evolutionary point of view, it seems not to matter one bit who is correct. (2002: 195).

Rea argues, further, that an appeal to intuition's track record to establish its reliability will not help here since philosophical intuitions outside the domains of logical and conceptual truths diverge so considerably. (Put in Griffiths' and Wilkins' terms, an appeal to the observed track record of intuition would constitute the claim that there is an indirect Milvian bridge for intuition; that intuition's reliability can be checked on independent grounds.) In any case, an appeal to intuition's track record in delivering us *true* philosophical beliefs will, in the present context, beg the question of how we know that it has delivered us true philosophical beliefs. It is hard to see, then, what other sorts of track record could be appealed to apart from intuition's success in *bringing about consensus*. But firstly, it is far from clear that intuition *has* tended to bring about philosophical consensus

very much of the time.[56, 57] And secondly, it is in any case very hard to see how the success of philosophical intuitions at producing consensus would suffice in the present context to vindicate their reliability, even if it were true that intuitions concerning philosophical puzzles have generally tended to converge. Take the case of moral beliefs, as a closely analogous example. Morality is a domain for which, according to some,[58] it is very unlikely that we were evolutionarily selected to have *accurate* beliefs—if the truth-makers for those beliefs are mind-independent, that is. (Note that opponents of the evolutionary debunking of morality do not deny this claim about the lack of a connection between the accuracy of moral beliefs and evolutionary fitness, but rather, seek to show that moral beliefs with mind-independent truth-conditions can nonetheless constitute knowledge.[59]) In that dialectical context—a context in which an evolutionary challenge has been posed concerning the global reliability of a certain type of belief—it would hardly be sufficient to rescue moral beliefs from being debunked simply to point out that humans by and large *agree* with respect to their reflexive moral judgments. I conclude, then, that there is no Milvian bridge of either sort

[56] Rea writes that "A survey of the history of philosophy reveals that intuition-based theorizing has an extremely poor track record with respect to producing anything like the progress and consensus we find in the empirical sciences" (2002: 196).

[57] I do not wish to deny that there is *ever* consensus to be found within philosophy concerning matters of intuitive judgment. There is considerable agreement, to take one prominent example, regarding the failure of the justified true belief analysis of knowledge in light of Gettier cases, but such instances of consensus are noteworthy for their scarcity.

[58] See Joyce (2006); Street (2006); Griffiths and Wilkins (2010); Wilkins and Griffiths (2013); Ruse (1986, 2006).

[59] See Wielenberg (2010); Enoch (2010); White (2010).

for the intuitions to which philosophical reasoning appeals.

Putting this all together, I have argued that belief in the Milvian bridge principle, (M*), fails to satisfy the principle's own conditions for epistemic reliability. It is therefore self-defeating. I noted that Griffiths and Wilkins allow for two ways in which a given type of beliefs might be connected with evolutionary fitness in virtue of accuracy: a direct connection and an indirect connection. I first tried to show that there is no direct connection. Despite the maladaptive behaviour to which a very few philosophical beliefs could conceivably lead, the vast majority of philosophical beliefs seem highly unlikely to bring about any behaviours which significantly affect evolutionary fitness. I subsequently tried to show that there is not even an indirect connection between fitness and the accuracy of philosophical beliefs. An indirect connection would require that the other types of beliefs upon which philosophical beliefs are based have a connection with fitness in virtue of their accuracy. I suggested that such a connection does plausibly obtain in the cases of logical beliefs and conceptual beliefs, and perhaps even in the case of scientific beliefs. But reliability with respect to all three of these types of belief is still not sufficient for reliability with respect to philosophical beliefs, since in addition, appeals to intuition are required in order to make judgments about philosophical theories, and I argued that intuition lacks the requisite connection with evolutionary fitness.

§1.5 Conclusion

The verdict I have delivered on the success of three currently popular epistemological debunking arguments has been a largely negative one. I suggested that the counterfactual insensitivity argument can't be made to work unless it smuggles in tacit assumptions about the difference—or lack thereof—that would be made to the world by God's failure to exist. In short, that argument rests upon the success of separate arguments to show that the likelihood of a world containing humans with religious beliefs is not greater given God's existence than given God's non-existence. The inappropriate causal connections argument fares somewhat better, I suggested, but probably only given two assumptions: namely, that a more restrictive causal condition on knowledge of the sort offered by Shope is correct, and that causal determinism is true. Given both these assumptions, it seems that God would be causally connected with theistic belief only in the sense of being the originator of the universe, which falls short of Shope's condition. I argued that given *in*determinism, however, it is likely that God would need to have acted much more proximally in order to guarantee the emergence of dispositions towards religious beliefs in humans, and moreover, that God could act as a proximate cause of religious experiences. Finally, I considered the evolutionary reliability argument of Griffiths and Wilkins, suggesting both that the principle upon which it relies is quite likely false given a certain sort of theism, but moreover, that that principle is ultimately self-defeating.

Chapter Two

A Debunking Argument from Diverse Deliverances

> Mere accident has decided which of these numerous worlds is the object of [a person's] reliance… the same causes which make him a churchman in London would have made him a Buddhist or a Confucian in Peking.
>
> John Stuart Mill

§2.0 Introduction

We have seen that the standard model of CSR postulates that human beings possess various cognitive mechanisms that dispose us to form religious beliefs under certain widely realized circumstances. According to CSR, however, this cluster of cognitive mechanisms is responsible for generating doxastic states that have only a rather slim propositional content, while the remaining content is filled in by cultural influences—hence the kaleidoscopic array of religious beliefs we find around the world today.[1] The epistemological principles that were at play in the

[1] According to Michael Murray,"When environmental conditions stimulate the various mental tools that are taken to be involved in religious cognition—HADD, Theory of Mind, memorability

debunking arguments considered in Chapter 1 were such that this datum couldn't be exploited. The present chapter, however, will attempt to find a viable debunking argument that does attempt to utilise this fact of diverse deliverances; an argument to the conclusion that a set of religious beliefs generated by this combination of bio-psychological and cultural belief-forming mechanisms are unreliably formed, and so fail to constitute knowledge even if they are true. An argument of this sort is the best of the debunking arguments on offer.

CSR theorizes that a set of universal religious belief-forming tendencies find their eventual expression through a variety of cultural religious frameworks. If correct, this seems to lend weight to a frequently expressed thought: that a typical religious believer would very likely have arrived at a significantly different set of religious beliefs had she been raised in a significantly different culture. As Alvin Plantinga notes, in the vicinity "there is an oft-repeated pluralistic argument (an argument that goes back at least to John Stuart Mill's *On Liberty* and possibly all the way back to the third century) that seems to be designed to appeal to reliabilist intuitions" (1995: 211). According to Mill's version of that argument, "mere accident has decided which of these numerous worlds is the object of [a person's] reliance… the same causes which make him a churchman in London would have made him a Buddhist or a Confucian in Peking" (1991 [1859]: 229-30).[2] The

---

and transmissibility via minimal counter-intuitiveness, etc.—the outputs of these tools are still highly non-specific… no one doubts that divergent cultural traditions play an enormous role in giving religious concepts their specific contours" (2010: 172).

[2] For a thorough discussion of Mill's argument, see Mawson (2009).

lesson Mill appears to have drawn is that even if the churchman's beliefs are true, it would be a matter of sheer epistemic luck that the relevant belief-producing mechanisms led him to truth rather than falsehood. Something very like this argument has found more recent expression, notably in the writings of John Hick, who suggests that

> [R]eligious allegiance depends in the great majority of cases on the accident of birth: someone born into a devout Muslim family in Pakistan is very likely to be a Muslim, someone born into a devout Hindu family in India to be a Hindu, someone born into a devout Christian family in Spain or Mexico to be a Catholic Christian; and so on. The conclusion that I have myself drawn from this is that a "hermeneutic of suspicion" is appropriate in relation to beliefs that have been instilled into one by the surrounding religious culture. (1997: 281).

And, again, writing in a very similar vein, Philip Kitcher has argued that

> Most Christians have adopted their doctrines much as polytheists and the ancestor worshipers have acquired theirs: through early teaching and socialization. Had the Christians been born among the aboriginal Australians, they would believe, in just the same ways, on just the same bases, and with just the same convictions, doctrines about Dreamtime instead of about the Resurrection… Given that they are all on a par, we should trust none of them. (2011: 26).

In what follows, I shall make a rigorous attempt to reconstruct an argument of this sort, drawing upon the latest and most promising accounts of knowledge.

§2.1 Formulating a Debunking Argument from Diverse Deliverances

The first thing to note about the sort of argument we are considering, as it is typically presented, and as it is presented by Mill and Kitcher, is that it involves a counterfactual conditional: You believe that *p*, but *if you had been raised in different cultural circumstances, you would instead have believed q* (where *q* is inconsistent with *p*). Of course, as Alvin Plantinga (1995: 212) has pointed out, a counterfactual of this sort is true not just of many religious beliefs, but also of all sorts of other beliefs, including scientific beliefs. Had I been raised in first-century Athens, I would have believed that the earth is flat, which is obviously inconsistent with my actual belief that the earth is round. There is, however, a standard fix for this issue: namely, to emphasise the similarity of belief aetiology between the actual and the counterfactual cases.[3] Whereas my belief that the earth is round is the causal result of seeing countless images of the earth taken from space and of the unanimous testimony of countless scientific experts, the belief I would have held if I were raised in ancient Athens—the belief that the earth is flat—would have been the causal result of something very different: accepting mere speculation about the earth's shape. On

[3] For more on this, see Bogardus (2013: 377-78).

the other hand, the thought goes, when it comes to most religious beliefs, the difference between the causal history of the actual belief and that of the counterfactual belief is not nearly so marked. Mill and Kitcher seem implicitly to be aware of this point. Thus, Mill: "the *same causes* which make [a man] a churchman in London would have made him a Buddhist or a Confucian in Peking [emphasis mine]" (1991 [1859]: 229-30). And Kitcher: "Had the Christians been born among the aboriginal Australians, they would believe, *in just the same ways, on just the same bases, and with just the same convictions,* doctrines about Dreamtime instead of about the Resurrection [emphasis mine]" (2011: 26). In our reconstruction of the argument from diverse deliverances, then, the counterfactual conditional we employ should have the following form:

(*C*) IF *S* had been raised under different cultural circumstances, THEN *S* would have held a belief that is inconsistent with her actual belief, whose causal history would have been *relevantly similar* to that of her actual belief.

I suggest, then, that we may formulate an argument from diverse deliverances in the following way:

Core argument

(1) For many religious believers, *R*, it is true that *C*: IF *R* had been raised under different cultural circumstances, THEN *R* would have held different and

inconsistent religious beliefs whose causal histories would have been relevantly similar to those of her actual religious beliefs.

(2) IF *C* is true for *R*, THEN *R*'s religious beliefs are unreliably formed.

(3) THEREFORE, for many *R*, *R*'s religious beliefs are unreliably formed [from (1) and (2)].

Knowledge defeat argument

(4) IF one's belief is unreliably formed, THEN that belief is not knowledge.

(5) THEREFORE, for many *R*, *R*'s religious beliefs are not knowledge [from (3) and (4)].

In what follows, my focus will be upon the core argument. Before I can go any further, however, I must distinguish two kinds of reliability.

§2.1.2 Modal reliability and process reliability

On the one hand, there is what might be termed *modal reliability.* Modal reliability is what is lacking in a Gettier case, in which a subject is lucky, in a certain sense, to wind up with a true rather than false belief. Farmer Jones, for example, looks into his field and forms the true belief that his prize cow Daisy is in the field, though what he is in fact looking at is a black and white sheet caught up in some trees that looks from a distance just like Daisy, obscuring the real Daisy from view. Given

how he forms his belief, Farmer Jones appears to be very lucky to get a true rather than false belief. Put in terms of the framework of possible worlds, there are some very close worlds in which Farmer Jones gets a false belief: worlds in which Farmer Jones still forms his belief by looking at the black and white sheet caught up in the trees, but in which Daisy has in fact wandered into the neighboring field. Duncan Pritchard (2005: §6.1) has termed this kind of luck "veritic luck." A modal reliability condition on knowledge aims to exclude precisely this sort of luck. Two candidates for such a condition have been given serious consideration in the recent epistemological literature: the sensitivity condition (which we encountered in Chapter 1) and the safety condition. Sensitivity can be formulated in the following way:

> SENSITIVITY: IF *S* knows that *p* in a world *w*, THEN in the nearest world(s) to *w* in which *p* is false and *S* forms a belief about *p* in the same way as in *w*, *S* doesn't believe that *p*.[4]

I noted in the previous chapter that sensitivity suffers from several notorious difficulties, including its failure to preserve deductive closure and to allow for inductive knowledge.[5] On the other hand, the much more popular safety condition appears to be capable of excluding veritic luck whilst avoiding the

[4] Dretske (1971) and Nozick (1981: 172-79) have both defended a condition along these lines.

[5] For more on these issues, see Hawthorne (2004: 31-46) and Pritchard (2005: 33-66).

aforementioned difficulties:[6]

> SAFETY: IF *S* knows that *p* in a world *w*, THEN there is no nearby world *w** in which *S* arrives at a false belief about *p* with a relevantly similar causal history.[7, 8]

Now, I suggest that we understand "belief about *p*" somewhat loosely, so as to encompass belief that *p*, belief that ~*p*, and any belief in a proposition that is obviously inconsistent with *p*. The reason for this is to ensure that cases in which *p* is a necessary truth are given intuitively correct verdicts.[9] An example might be a

---

[6] Bogardus (2013) interprets the argument from diverse deliverances in terms of safety.

[7] It is true that cruder versions of safety, which merely require that *S* doesn't falsely believe *p* in a nearby world, cannot give sensible results concerning badly formed beliefs in necessary truths. The more sophisticated version of safety that I have offered is close to formulations presented by Pritchard (2012: 256-57) and Williamson (2009: 235).

[8] Instead of my clause concerning "a relevantly similar causal history," some authors prefer the phraseology of "by way of the same method." I don't have any quarrel with this way of putting things, but the notion of "holding the method fixed" is really a shorthand way of saying that, for the purposes of evaluating a belief's safety, we should only look at those nearby worlds in which the belief's causal history is at most slightly different than it is in the actual world.

[9] Pritchard pursues a similar approach to extending the safety condition so as to be able to deal with beliefs in necessary propositions: "[I]t 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 nearby possible worlds where *that* belief is false, there is a wide class of nearby possible worlds where that belief-forming process brings about a doxastic result that is false (e.g., a possible world in which one in this way forms the belief that 2 + 2 = 5)" (2009: 34). See also his (2012: 256-57; 2014: 158-59).

case in which Josh forms a belief about the square root of 9 by rolling a dice. Suppose that in the actual world Josh is lucky: the dice comes up "3," and so he correctly believes *p* ("the square root of 9 is 3"). Now, it may well be that there isn't any nearby world in which Josh explicitly believes ~*p* as a result of his dice-rolling "calculation," but still, there are nearby worlds in which, using the same dice-rolling method (and hence forming a belief by way of a relevantly similar causal history), Josh ends up believing propositions that are obviously inconsistent with *p*, for instance, that "the square root of 9 is 4." Given the way we have formulated safety, Josh won't count as knowing *p* in the actual world, and more generally, safety thus formulated seems able to handle cases of belief in necessary truths.

At the same time, some epistemologists have suggested that modal reliability may not be sufficient for knowledge by itself, and that there is a distinct kind of reliability that is also necessary for knowledge.[10] One reason for thinking this is that there is a certain sort of case in which a subject forms a modally reliable belief, despite her belief having been formed by way of a type of causal process that intuitively seems not to be capable of producing knowledge. Suppose, for instance, that Alfie forms various beliefs out of wish-fulfillment, but that unbeknownst to him, he has an epistemic guardian angel who was assigned to him at birth and who has been tasked with making Alfie's every wish come true. Alfie's belief that "Santa Claus is real," formed by way of wish-fulfillment, is safe, given

[10] See, for example, Pritchard (2012) and Becker (2008).

that the guardian angel acts to ensure its truth: there is no nearby world in which Alfie forms a belief out of wish-fulfillment and yet ends up with a false belief. But intuitively there is something wrong with Alfie's forming his beliefs on the basis of wish-fulfillment. What's wrong, perhaps, is that he forms his beliefs by way of a *type* of process that is generally unreliable—that is, a type of process that generates an insufficiently high proportion of true beliefs when used under normal circumstances.[11, 12] The process he uses—*wish-fulfillment*—is one that, when used under normal circumstances in the actual world (and in worlds with broadly similar physical setups to the actual world), generates considerably more false than true beliefs. It has a low *truth-ratio*. Alfie's failure to know might thus be explained in terms of his failure to satisfy the following condition:

> PROCESS RELIABILITY: IF *S* knows that *p*, THEN the salient process type that produced *S*'s belief has a high truth-ratio in the actual world and in worlds

[11] It should be noted that although many epistemologists consider knowledge to be inconsistent with belief that is the product of a generally unreliable process type, Maria Lasonen-Aarnio (2010) has defended the view that cases like that of Alfie are in fact cases of knowledge (knowledge, on her view, being mere safe belief), but that our reluctance as onlookers to *attribute* knowledge in such cases is explained by our wish not to reward the subjects in such cases for their employment of an unreasonable belief-formation policy—a policy that *in general* does not yield knowledge.

[12] I should also note that there is a slightly different but related diagnosis of Alfie's failure to know, which has to do with the fact that his cognitive success (i.e., the fact that he arrives at the truth) is not properly creditable to his cognitive abilities. As I shall explain in Chapter 3, this sort of virtue-theoretic approach to thinking about knowledge is fundamentally reliabilist in character, though differs somewhat from a straightforward process reliabilist approach.

with similar physical laws.[13]

Of course, appealing to process reliability to explain why Alfie's beliefs aren't knowledge depends on its being the case that the process he uses is rightly classified as *wish-fulfillment* rather than *wish-fulfillment in the presence of an epistemic guardian angel who ensures the truth of one's wishes*. The latter process type presumably has a very high truth-ratio, while the former doesn't. Alfie's token belief-forming episode is of course an instance of *both* the aforementioned process types (and of indefinitely many other process types at greater and lesser levels of generality), and so, if we are to appeal to the unreliability of the process type that Alfie uses, there needs to be a fact concerning *which* process type is the salient one. The notorious difficulty of determining the salient process type in a principled manner is known as the generality problem, and it is a problem that will become relevant later.

In what follows, I shall examine the core argument outlined in §2.1, interpreted first as targeting modal reliability—specifically, safety—and then as understood in terms of process reliability.

§2.2 Arguing that Diverse Deliverances Imply Unsafety

Let's now try to restate the core argument, outlined in §2.1, so as to secure the

[13] Becker (2007: Chs. 2, 5) and Goldman (1979) have argued for including in the analysis of knowledge something like this condition.

conclusion that religious beliefs formed by way of culturally contingent causal processes are not *safe*. For the purposes of illustration, let's consider Mill's Victorian churchman, who believes that "God is triune," and who, by way of similar causal processes as led him to his Trinitarian belief, would have believed that "God is non-triune" had he been raised among Confucians in Peking. We should assume for the sake of the argument that the churchman's actual belief happens to be true. The argument needs to be able to succeed even on this assumption, because it is trivial and uninteresting to show that what is assumedly a false belief is not an instance of knowledge. To say that the churchman's Trinitarian belief fails to be safe is to say that there are nearby worlds—including some of those worlds where he was raised in Peking—in which the churchman has a false belief whose propositional content and causal history are relevantly similar to those of his actual-world belief. In general, then, a proposition of the following form needs to be true of a person's religious beliefs if those beliefs are to be deemed unsafe:

(C^*) There are nearby possible worlds in which *S* has a false belief about *p* with a relevantly similar causal history to her actual-world belief that *p*.

Accordingly, our core argument, understood in terms of a failure of safety, can be stated as follows:

(1*) For many religious believers *R* and many religious propositions *p*, it is true

that C^*: there are nearby worlds in which R has an (*ex hypothesi*) false belief about p, which has a relevantly similar causal history to her actual-world belief that p.

(2*) IF a person safely believes that p in a world w, THEN there is no nearby world w^* in which she has a false belief about p with a relevantly similar causal history.

(3*) THEREFORE, for many religious believers R and many religious propositions p, R's belief that p is not safe [from (1*) and (2*)].

From there, the rest of the argumentation—the knowledge defeat argument—can proceed as before, at least, given the plausible assumption that knowledge requires safety.

What I want to argue now is that premise (1*) of the above argument is false, or at any rate, that unless we deem it false we will be committed to an unwarranted degree of scepticism about many ordinary cases of testimony. In general, when presented with the claim that S's belief that p is unsafe, there are a couple of ways in which one might try to challenge that claim: (i) one could try to show that the possible worlds in which S has a false belief about p by way of a relevantly similar causal history to her actual belief are not in fact *close* worlds; (ii) one could try to show that, even supposing there are close worlds in which S has a false belief about p, the causal history of S's belief in those worlds is not relevantly similar to the causal history of S's actual belief. I am going to attempt both (i) and (ii).

Now, clearly, the terms "close" and "relevantly similar" are significantly vague.[14] In disputed cases, such as cases of religious belief, it won't really do to appeal to intuitive judgments about the similarity of pairs of cases of religious beliefs that are mutually inconsistent with one another. The intuitive judgments of interlocutors in the atheist-theist debate are not likely to form a consensus on this matter. But what we can do, it seems to me, is use our intuitive judgments about cases on which all parties to the debate are agreed as a kind of anchor, in a manner that will become clear shortly. The anchor I propose to use is a sort of ordinary non-religious case of testimony-based belief, which I think nearly everyone will grant is a case of knowledge.

Let's term coin a term of art: an *error world* is a possible world in which a subject *S*, who correctly believes *p* in the actual world, instead has a false belief about *p*. An error world is a world either in which *p* is false and yet *S* continues to believe *p*, or in which *p* continues to be true and yet *S* incorrectly believes ~*p* (or some other proposition that is obviously inconsistent with *p*). Assuming the truth of the churchman's Trinitarian belief in the actual world, the world in which the same individual was instead raised in Peking among Confucians is an error world: in that world he (*ex hypothesi*) falsely believes that "God is non-triune." My contention, then, is that for some ordinary testimony cases that virtually everyone will deem

[14] The seeming inescapability of such vagueness in formulating central epistemological notions is apparently part of the motivation for Timothy Williamson's knowledge-first approach to epistemology, which eschews the project of locating non-circular necessary and sufficient conditions for knowledge, reliability, or justification (see his 2000: 2-5; 2014).

knowledge, the closest error worlds are *at least as close* to the actual world as the closest error worlds are to the actual world for fairly typical cases of religious belief, such as that of Mill's Victorian churchman. What's more, I contend that for such testimony cases, the causal history of belief in the closest error worlds is *at least as similar* to the belief's causal history in the actual world as is true for many cases of religious belief. It's worth noting that the closeness of an error world to the actual world is in part, though by no means wholly, determined by the similarity of the belief's causal history across worlds. The two variables are not completely independent of one another. In particular, a difference between two worlds with respect to the environment in which the believer forms her belief will entail differences with respect to the causal history of her belief, as well as entailing other differences between the two worlds that don't directly bear upon the causal history of her belief—for instance, the way the wind is blowing her hair.

So, consider the following sort of testimony case. Jill gets off the train in Truthland and obtains a correct belief about the directions to the British embassy by way of the testimony of someone who seems to Jill, and in fact is, both honest and knowledgeable about the whereabouts of the embassy.[15] What's more, everyone else in Jill's nearby vicinity is both honest and knowledgeable about the whereabouts of the embassy. I take it that many of us will be inclined to attribute

[15] This is inspired by a case presented by Jennifer Lackey (2007: 352), in which a man named Morris asks a trustworthy looking stranger for directions to the Sears tower in Chicago, who in fact is trustworthy and gives Morris the correct directions. I concur with Lackey's judgment that Morris acquires knowledge.

knowledge to Jill here. But it seems to me that for Jill in this case, the nearest error world is *at least as close* to actuality as the nearest error world is to the actual world for Mill's churchman. And relatedly, the causal history of Jill's belief in the closest error worlds is *at least as similar* to her belief's causal history in the actual world as the churchman's belief's causal history is to the causal history of his counterfactual belief in the world in which he is raised in Peking. Put another way, the degree to which Jill's environment must be counterfactually altered so that she instead winds up with a false belief seems to be *no greater than* the degree of counterfactual alteration involved in switching the churchman's entire cultural setting. One way to change Jill's environment so that she gets a false rather than true belief is by switching her location to Falsityland, where she would be surrounded largely by people who are mistaken (though sincere) about the directions to the British Embassy. But if we want to maintain that Jill acquires knowledge in the actual world then we'll of course have to say that it is *isn't* relevant (to the safety of her actual belief) that she would have fallen into error had she been located in Falsityland; we'll have to say that the world in which Jill is located in Falsityland is *not* a sufficiently close world, or that the causal history of her belief formed in Falsityland is *not* relevantly similar to the causal history of the belief she forms in Truthland. And yet the argument against religious beliefs from diverse deliverances, cast in terms of safety, will have to maintain that it *is* relevant for safety that a religious believer would have had false religious beliefs had she been located in religious "Falsityland" rather than in religious "Truthland." And so, if one insists

that such a believer's beliefs aren't safe even if they're true, then one will also be committed to saying that Jill (and people in situations like that of Jill) has an unsafe belief. But that seems implausibly sceptical.[16]

An important caveat to the above must be entered at this point. In using the example of Mill's churchman, I assumed that the individual in question was fairly far removed from the possible influence of a different religious tradition, meaning that a considerable counterfactual alteration of his environment would be required in order to place him under such influence. Now, I take it that the situation of Mill's Victorian churchman is fairly typical of believers who live in largely religiously homogeneous societies. In some sections of contemporary Western societies, though, there are presumably at least some individuals whose circumstances are such that very little would have to have been different in order for them to have ended up with a different set of religious beliefs.[17] That is, the nearest error worlds for the religious beliefs of such individuals are very close indeed. A child growing up in a multicultural district of a Western European city, say, might very easily have been exposed to the influence of a different religious tradition than the one to which she was actually exposed. Consider Aaliyah, who

[16] Note that this is different from Tomas Bogardus's (2013: 379-82) recent objection to a safety-based contingency argument against religious beliefs. Bogardus's objection is twofold: (i) safety is not necessary for knowledge; (ii) in any case, the epistemic luck that would be involved in happening to be born in a culture with true religious beliefs is a benign kind of luck. That sort of epistemic luck is analogous, he claims, to safely crossing a bridge that was nearly rendered unsafe: first-order safety is compatible with being lucky to have used a safe method.

[17] Thanks to Chris Kyle for drawing my attention to this consideration.

would have been sent to a school with a wholly different religious orientation, but for the whimsical decision of an educational administrator. Supposing (*arguendo*) that Aaliyah's Islamic beliefs are true, it looks as though not very much *at all* needed to go differently in order for Aaliyah to wind up with (*ex hypothesi*) false religious beliefs—indeed, less than had to go differently in order for Jill, who was fairly geographically far from Falsityland, to have ended up with a false belief. So it looks as though we may be able to say that Aaliyah's belief is unsafe, seemingly *without* risking the kind of epistemic "collateral damage" that we faced if we said the same of Mill's churchman.

But this isn't much of a consolation to the advocate of the argument from diverse deliverances. Mill and Kitcher and others who advocate such an argument are presumably thinking of their arguments as applying in full force to religious believers who come to their particular religious beliefs by way of religious enculturation, *regardless* of whether their circumstances are quite as modally precarious as those of Aaliyah. Now, to be clear, the extent to which a religious believer is *modally* distant from the influence of a different religious tradition is not straightforwardly determined by her *geographical* remoteness from a different tradition, though geographical remoteness will typically contribute to modal distance. The two will come apart in the case in which a believer has been raised in a location that is geographically far removed from the influence of a different religious tradition but in which it is true that if (say) the believer's parents had died or had not managed to stay financially above water (states of affairs that could fairly easily

have come about, let's say), she would have been sent to be raised by relatives living in a far-flung land and thus exposed to a wholly different religious outlook. Nonetheless, for many—perhaps most—individuals who are geographically far removed from the influence of other religious traditions, it is true that not easily could they have been subject to the influence of another tradition. Furthermore, even for a good many believers who are living in fairly close physical proximity to adherents of another tradition, a significant amount would have had to go differently in order for them to have been subject to those other religious influences. Religious believers in a situation as precarious as Aaliyah's, vis-à-vis which religious beliefs they end up holding, are presumably going to be in a considerable minority, even among inhabitants of moderately pluralistic Western societies. But in order to avoid the sorts of unintended sceptical consequences sketched above (i.e., regarding Jill), the argument from diverse deliverances would have to restrict its focus exclusively to cases like Aaliyah's; exclusively to those religious beliefs that are less modally reliable than any beliefs that we take to be uncontroversial cases of knowledge. Such a narrowing of the argument's scope would be seriously disappointing as far as Mill and Kitcher are concerned. Let's see, then, how an argument from diverse deliverances fares if conducted in terms of the reliability of belief-forming process types. What I will argue is that there *is* a way to individuate process types such that religious beliefs can be shown to be the product of an unreliable type, though I will go onto to note some of the potential costs of individuating types in the suggested manner.

§2.3 Arguing for the Unreliability of Religious Belief-Forming Process Types

If one wants to run the core argument in terms of process type reliability, without assuming the falsity of any particular set of religious beliefs—an assumption that would render the argument uninteresting—then one would think the best way to proceed is to try to show that the salient process type is one that is responsible for a variety of mutually inconsistent religious beliefs. After all, a process type is not going to have a truth-ratio high enough to satisfy a process reliability condition on knowledge if too many of the beliefs for which it is responsible are inconsistent with one another. Even without making any assumptions about the truth-values of any *particular* beliefs that are the outputs of a given process type, we can still infer that the truth-ratio of that process type is no greater than the largest number of the type's outputs that could all be true at once divided by the total number of the type's outputs. I suggest, then, that if construed in terms of process types, the core argument ought to run in the following way:

(1**) For many religious believers *R* and religious propositions *p*, the salient process type that is responsible for producing *R*'s belief that *p* is such that a significant proportion of its outputs are mutually inconsistent.[18]

[18] Note that when the core argument is run in terms of process reliability, the argument's first premise doesn't need to appeal to a counterfactual claim, though it *could*. The relevant counterfac-

(2**) IF the salient process type that is responsible for producing one's belief is such that a significant proportion of its outputs are mutually inconsistent, THEN that process type does not have a truth-ratio sufficiently high to count as reliable.

(3**) THEREFORE, for many religious believers *R* and many religious propositions *p*, the salient process type that is responsible for producing *R*'s belief that *p* does not have a truth-ratio sufficiently high to count as reliable [from (1**) and (2**)].

Again, the remainder of the argument—the knowledge defeat argument—can proceed as originally stated in §2.1, on the (not uncontroversial) assumption that process type reliability is necessary for knowledge. My focus once again will be the first premise, (1**). The issue that now rears its ugly head is that of picking out the *salient* process type. As I mentioned earlier, any particular belief will be the result of a token causal process that can be classified under virtually indefinitely many process types, and these types will vary widely in their reliability. But rather than try to resolve the thorny problem of how the salient process type ought to be selected (the generality problem), the approach I'll take instead is to explore the prospects for two broad approaches to this issue that could be taken by the advocate of the argument from diverse deliverances. One approach is to pick out a

---

tual would be something like: if *R* had been raised under different cultural circumstances, then *R* would have had different and (*ex hypothesi*) false religious beliefs that would have been the product of the same salient process type as produced her actual religious beliefs.

fairly narrow, though nonetheless repeatable, process type. I'll suggest, however, that this approach faces a serious difficulty if monotheistic religious beliefs are the target of the argument. Another approach is to select a rather wider (i.e. more general) process type, something like *accepting the testimony of a religious tradition*. I shall argue that this latter approach may in fact succeed, and that it is the most promising debunking argument available, though I shall outline some of its potentially costly consequences for moral epistemology and science.

§2.3.1 Narrower process type individuation

Consider that what determines the narrowness of a process type is the extent to which the specification of that type involves particular individuals—including particular objects, times, and events. The process type *forming a belief by way of accepting the testimony of Huw Edwards via the TV set in my lounge between 6 and 6.30PM on Saturday 8th March 2014* is a very narrow process type compared with the type *forming a belief by way of accepting the testimony of a TV news anchor*, which mentions no particular objects, events, or times. The latter type has vastly many more instances than the former. What I want to suggest is that the argument from diverse deliverances faces a serious difficulty if the process type picked out as the salient one is specified in a way that involves particular individuals. As I said, the argument from diverse deliverances needs to assume for the sake of the argument the truth of the religious beliefs that are being targeted. What the defender of the argument

from diverse deliverances is then looking to show is that the relevant process type, as a result of which a given believer formed her religious beliefs, is such that it yields many beliefs that are inconsistent with the believer's actual beliefs; the defender of the argument is trying to show, that is, that the relevant process type has an insufficiently high truth-ratio. But if the process type in question is specified in a way that involves the *particular* chain of individuals whose testimony was actually causally efficacious in bringing about a person's religious beliefs, or in a way that involves the *particular* religious texts or rituals that actually contributed to her belief-formation, then it looks as though that process type will *not*, on the whole, yield very many beliefs that differ from her actual beliefs. Now, it is worth emphasising at this point that we need not restrict our attention just to the *actual world* truth-ratio of process types (i.e. the ratio of true beliefs to false beliefs produced in the actual world by a given process type); rather, we can—and according to Kelly Becker, *should*—also take into account the truth-ratio of a given process type across a nearby region of possible worlds centred on the actual world.[19] Given this, one might think that even if we do mention in our specification of the process type things like a particular set of religious scriptures or a particular testimony chain, we will nonetheless get to take into account worlds where, say, that set of scriptures or that testimony chain yields many claims which fail to

[19] The reason for this is that on a narrow type-selection approach such as the one we are currently considering, and which Becker (2007) advocates, a given process type may not have many actual-world instances, and so in order to adequately assess its truth-ratio, we need to consider to take into account the process type's possible outputs throughout a nearby region of worlds; worlds that are fairly similar to the actual world. See Becker (2007: 89).

match the facts.[20] One might think that when all these worlds are factored in, it could well turn out that even a narrow process type, one whose specification includes a particular set of religious scriptures or a particular testimony chain, nonetheless yields a significant proportion of mutually inconsistent outputs and thus a significant proportion of false beliefs.

I think this line of reasoning is flawed, however. Suppose that the beliefs of a certain monotheistic believer are the target of the argument from diverse deliverances on this occasion, so that the truth of her various religious beliefs is granted for the purposes of the argument. In that case, the defender of the argument from diverse deliverances must reckon with the fact that given the truth of a monotheistic belief-system (according to which God created humans and wishes to relate personally to them), it is pretty epistemically probable that God has providentially guided or overseen the processes by which the scriptures of her religion have been written, assembled, and subsequently transmitted down through history, such that not easily could those scriptures have ended up asserting more than a small proportion of false claims. If God has indeed guided these transmission processes, and did so as a result of a settled intention rather than a whim—something which also seems highly epistemically probable given the truth of a monotheistic belief-system on which God wishes to relate to humans—then a possible world where a

[20] Issues about modal fragility may threaten the notion that a particular set of religious scriptures could have included substantially different propositional content and yet have remained *the very same* set of scriptures. The argument I am about to offer, though, doesn't rely on assuming the modal fragility of scriptures or testimony chains.

large proportion of scriptural assertions (as they are in that world) fail to match the facts (as they are in that world) will be extremely far from the actual world, and so they will barely, if at all, register in our evaluation of the process type's truth-ratio.

Now, there is an obvious objection to the argument I just offered: there exist disagreements (hence, mutually inconsistent beliefs) not only at the *inter*-religious level, but also, frequently, at the *intra*-religious level.[21] There is diversity of belief even among believers who formed their beliefs as a result of the *same* processes of historical transmission, by way of the *same* scriptures and traditions. What can be said about this? Well, to begin with, it is important to note that we implicitly relativize our reliability judgements about process types to a particular domain and a particular level of specificity. Here's what I mean. Suppose *The Daily Post* prints a fair number of errors with respect to the precise times at which events occurred. Nonetheless, that, by itself, doesn't seem to licence the conclusion that reading *The Daily Post* is an unreliable way to form the belief that some particular event has occurred at all, just that it's an unreliable way to form a belief with more specific content concerning the precise time at which some event occurred. Again, suppose *The Daily Post* frequently prints mistakes regarding the sports scores. This doesn't appear, by itself, to licence the conclusion that *The Daily Post* is an unreliable way to form beliefs about some domain other than sports scores, say, political current affairs. Analogously, then, the fact that some religious process type (say, the type *forming beliefs by way of reading the Qur'an*) leads to a good deal of mutually

[21] Thanks to Brian Leftow and Chris Kyle for drawing my attention to this objection.

inconsistent beliefs concerning the various *specific ways* in which the core tenets of a religious system could be true, doesn't licence the conclusion that the process type in question is an unreliable way of forming the mere belief that those tenets are true. And it does look like a great deal of intra-religious disagreement—not *all* instances of it, to be sure—occurs at the level of beliefs with more specific content; beliefs not merely that the core tenets of a religion are true, but concerning the *specific ways* in which those core tenets are true. For instance: about exactly how it is that Jesus' death achieved the reconciliation of God with humans; about the precise nature and scope of the commandment to love one's neighbour; about the exact manner in which the Hajj is to be undertaken; about the means by which God created the world; and so on. The claim can still stand, then, that conditional on the truth of a given monotheistic belief system, it is very epistemically likely that God has guided the processes by which at least the core tenets of that system have been transmitted down through history, such that not at all easily could those processes have yielded more than a small proportion of false beliefs whose content is no more specific than that the core tenets of that system are true.

With that said, this won't be a claim that can be made on behalf of religious belief-systems that don't avow belief in a deity capable of such supernatural guidance, and so perhaps a version of the argument from diverse deliverances cast in terms of narrower process types can still succeed against *those* sorts of religious beliefs. I take it, though, that the argument from diverse deliverances has usually been thought of as threatening *all* varieties of culturally instilled religious beliefs.

That is certainly how Mill and Kitcher seem to see things. Accordingly, let us consider an approach on which a wider process type is held to be the salient one.

§2.3.2 Wider process type individuation

An appropriately wide process type whose specification involves no mention of particular testimony chains or particular religious texts, traditions, or rituals might be something like *accepting the testimony of a religious tradition*. A process type like *this* will surely be wide enough so as to be instantiated by a large variety of mutually inconsistent religious belief tokens, and thus have an insufficiently high truth-ratio. Indeed, consider again Mill's claim that "*the same causes* which make [a man] a churchman in London would have made him a Buddhist or a Confucian in Peking [emphasis mine]" (1991 [1859]: 229-30). As far as I can tell, *accepting the testimony of a religious tradition* is about as narrow a process type as can be selected here without rendering Mill's claim false; any more detail—any mention of *particular* religious teachers, texts, rituals, or experiences—and the pair of cases Mill has in mind won't rightly be described as involving the *same type* of causal process. Similarly, Kitcher would have to have in mind something like this fairly wide process type in order for it to be true that, as he claims, "Had the Christians been born among the aboriginal Australians, they would believe, *in just the same ways, on just the same bases,* and with just the same convictions, doctrines about Dreamtime instead of about the Resurrection [emphasis mine]" (2011: 26). Again,

Aboriginal and Christian ways of coming to believe their respective doctrines seem to have in common nothing more specific than that they both involve a tendency to accept the passing on of religious teachings and practices from one generation to the next. This strategy we are now considering, then, does indeed appear to fulfil the desideratum of picking out a process type that encompasses many mutually inconsistent religious beliefs. But I now want to explore two potentially problematic implications of this approach, and see what can be done to mitigate the seeming difficulties.

§2.3.2.1 Implications for moral epistemology

One implication of advancing the argument from diverse deliverances on the basis of a wide type-selection approach is that *moral* belief-forming process types at an equivalent level of generality, some of whose outputs we tend to be very reluctant to deem non-knowledge, yield about as high a proportion of mutually inconsistent outputs as does the type *accepting the testimony of a religious tradition*. If the type *accepting the testimony of a religious tradition* doesn't have a high enough truth-ratio to yield knowledge, then nor do these other process types. I won't dwell long on this point, since it has already received significant discussion. Plantinga (1997: 297-98), for instance, gives the example of someone's belief that "racial discrimination is wrong," which has been instilled by her cultural upbringing. Peter van Inwagen (1995: 238) and Jerome Gellman (1993: 350) make similar points concerning

political beliefs, which typically one has arrived at partly by way of upbringing, such as the belief that "democracy is a fairer system of government than totalitarianism." The (wide) process type at issue here, something like *accepting cultural moral/political values*, looks to be responsible for nearly as many mutually inconsistent belief outputs as does *accepting the testimony of a religious tradition*. By parity reasoning, then, it looks like the defender of the argument from diverse deliverances should conclude that one doesn't know, on that basis, that (say) racial discrimination is wrong. Of course, one might bite the bullet regarding culturally instilled moral beliefs, and yet maintain that moral beliefs formed by way of rational reflection are in good shape, nonetheless. The problem is that the equivalently wide process type here, *rational moral/political reflection*, appears to issue about as many mutually inconsistent outputs as the aforementioned types.

A way around this problem might be to adopt some form of moral anti-realism, according to which moral propositions are true or false just in virtue of facts about the moral beliefs of the relevant moral community. On such a view, one could plausibly get the result that moral belief-formation by way of one's upbringing can often result in knowledge: assuming that moral facts supervene upon the consensus of a moral community, process type reliability would require only that the type *accepting cultural moral/political values* (and, indeed, the type *rational moral/political reflection*) doesn't produce more than a small proportion of mutually conflicting beliefs *within* a given moral community; it wouldn't require such consistency *across* moral communities. There is perhaps the worry with this constructivist approach

that one will be committed to saying that an appalling (say, fascistic) moral belief that is held by the majority of a certain community will both be true and constitute knowledge for members of that community. Such a consequence could perhaps be avoided by enlarging moral communities both spatially and temporally to the point that such beliefs are held only by a minority, thus rendering them false and so non-knowledge. This enlargement of moral communities will, however, presumably also have the effect of increasing the overall degree of disagreement occurring within them, thus decreasing the overall truth-ratio of the process types *accepting cultural moral/political values* and *rational moral/political reflection*, perhaps—though not inevitably—to the point that they won't be counted reliable.

§2.3.2.2 Implications for science

Another implication of running an argument from diverse deliverances using a wide process type individuation approach concerns the deliverances of science. The issue is roughly that two process types at a level of generality approximately equivalent to *accepting the testimony of a religious tradition*, namely, *constructing and testing scientific hypotheses* (for the scientists themselves) and *accepting expert scientific testimony* (for the laypeople), encompass scientific belief tokens from *all* eras of scientific research. Given that many scientific beliefs from bygone eras are inconsistent with modern scientific beliefs,[22] the overall truth-ratio of the types *constructing and testing*

[22] For a colourful list of now abandoned scientific theories, see Laudan (1981).

scientific hypotheses and *accepting expert scientific testimony* is not going to be all that high; that truth-ratio can *at most* be equal to the greatest number of scientific beliefs that could all be true at once divided by the total number of scientific beliefs. Even without assuming the falsity of any particular scientific beliefs, we can see that, given such a wide type-selection policy, the truth-ratio of the aforementioned scientific process types will get dragged down due to the mutual inconsistencies between scientific beliefs of different eras. Of course, a much more fine-grained individuation of scientific belief-forming process types—which, for instance, includes in the description of a process type the specific experimental methods, techniques, background theory, and equipment used—wouldn't lump all scientific beliefs together under one type, and thus wouldn't have the upshot that scientific beliefs all stand or fall together with respect to the process type reliability condition. But were the defender of the argument from diverse deliverances to adopt this fine-grained approach to type selection, she would have to do likewise for religious beliefs, thus mentioning specific religious texts and testimony chains in her description of the relevant process type, and I have already suggested that that approach won't work (not against monotheistic religious beliefs, at any rate), given the needed assumption of the truth of the targeted religious beliefs.

With that said, one fairly striking difference between religious diversity, on the one hand, and the diversity of expert scientific opinion, on the other, is that the latter phenomenon is principally *diachronic*. That is to say, disagreement between scientific experts, though to some degree *synchronic* (occurring between contempo-

raries *at* a given time), is much more pervasive *across* historical time-periods. On the other hand, religious diversity is clearly both diachronic and synchronic: it is both the case that modern people have substantially different religious beliefs to peoples living thousands of years ago, *and* that various peoples hold substantially different religious beliefs from one another at any given point in history. Accordingly, one move the defender of the argument from diverse deliverances might make is to index process types to *times*, or time periods. The details would need some working out, in particular, concerning how times are to be individuated. At any rate, the basic thought is that by indexing process types to time periods, the epistemically salient process type employed by Ptolemaic astronomers—*constructing and testing scientific hypotheses during* t_x—can be distinguished from the epistemically salient process type used by modern day cosmologists—*constructing and testing scientific hypotheses during* t_y—and thus the unreliability of the former need not spread to the latter. Of course, this would mean that during what Thomas Kuhn (1996) has termed "revolutions" in the history of science, during which a once-dominant theoretical paradigm is overthrown and various theoretical frameworks compete to become the new dominant paradigm, it will be harder for a scientific practitioner to know the truth of a given hypothesis (assuming the hypothesis in question is even true at all). After all, a process type such as *constructing and testing scientific hypotheses during* $t_{\textit{some particular scientific revolution}}$ would be responsible for a significant proportion of mutually inconsistent outputs and thus would *at best* have only a moderate truth-ratio. But this seems in fact to be a fitting consequence. Now, consistency

demands that religious belief-forming process types also be indexed to times, but since religious diversity is a synchronic phenomenon to a much greater degree than is scientific expert disagreement, this shouldn't undermine the argument from diverse deliverances. A type such as *accepting cultural religious testimony during* $t_{\textit{the early twenty-first century}}$ will still yield the large proportion of mutually inconsistent outputs that the defender of the argument from diverse deliverances needs it to.

Still, one might rightly worry that even with process types indexed to time periods, scientific beliefs whose epistemic credentials are seemingly very strong will get unfairly dragged down by the fact that, at any given time, some areas of scientific research lag considerably behind others, and yet on this wide type-selection approach they are all lumped together under one process type. Why, for instance, should the epistemic status of Charles Darwin's (intuitively, well grounded) beliefs about biological heredity be brought down by the fact that at the time he was working, some fairly far removed areas of science such as astronomy had produced many false beliefs? One solution to this worry would be to index process types to individual *propositions*; to specify types according to the following schema: *forming a belief about whether p by way of a causal process type C during t.*[23] In this way, the

[23] One might try to index a process type to something broader than a single proposition, such as a "subject matter," but it is very hard to see how the individuation of subject matters won't simply end up being entirely *ad hoc*. On this point, Becker (2007: 104; 2008: 363-58) proposes that we specify a process type in terms of the "narrowest specific-content-neutral" description of the token belief's subject matter. I take it that specific-content-neutrality is a matter of the description of the process not entailing the propositional content of the process's outputs. As far as I can see, though, Becker's examples of such descriptions (e.g., "belief about birds" as the description of the subject matter of the belief that "there is a bird in the tree at which I am looking") are *not* in fact

process type reliability pertaining to a given scientific belief will be determined only by the truth-ratio among beliefs held during the same time period concerning just that one proposition. This move would go some way to alleviating the concern about degenerate fields of scientific research dragging down progressive fields. Again, *ideally* one would have recourse in one's specification of the causal process type to the specific experimental methods and theoretical background employed, but this isn't available on the wide type-selection approach that the defender of the argument from diverse deliverances needs. Also, note that this move of indexing process types to particular propositions has the following cost for the defender of the argument from diverse deliverances. It is admittedly true that a process type such as *forming a belief about whether "God is triune" by way of accepting the testimony of a religious tradition during* t_n will yield a pretty high proportion of mutually inconsistent outputs and so at best have an insufficiently high truth-ratio for knowledge, and it's true that this will in general be the case whenever the proposition is one believed only by some but not other religious traditions. But when it comes to propositions believed by virtually all religious traditions, the proportion of mutually inconsistent outputs of the relevant process type will be very low, and so given the truth of such a proposition, the type reliability condition will be satisfied.

---

the narrowest specific-content-neutral descriptions (a point made forcefully by Brueckner and Buford (2013)). Rather, the narrowest such description for the belief that *p*, for any *p*, will always be "belief about *whether or not p*."

§2.4 Conclusion

In summary, I have attempted to identify a stronger debunking argument against religious beliefs than any of those arguments considered in Chapter 1. I have suggested that even this strongest of debunking arguments suffers from some shortcomings, in the sense that it has unattractive implications for moral epistemology—unattractive, at any rate, if one wishes to hold onto moral realism, which admittedly some naturalists are happy enough to surrender. What's more, I suggested that in order to avoid certain sceptical consequences regarding science, process types should be indexed to times and propositions, which means that the argument from diverse deliverances will at best only be capable of debunking beliefs in propositions that are not believed by a majority of religious traditions. In the next chapter, I am going to try to show that even supposing a debunking argument of the sort explored in this chapter is successful against culturally contingent religious beliefs, there is nonetheless a "core" religious content in which humans are disposed to believe across all cultures, and which is such that humans can come to know that content if theism is true, given various currently popular analyses of knowledge including those featured in this chapter.

Chapter Three

Knowledge of Core Content

> Two things fill the mind with ever-increasing wonder and awe, the more often and the more intensely the mind of thought is drawn to them: the starry heavens above me and the moral law within me.
>
> Immanuel Kant

§3.0 Introduction

Let us take stock of where we have now arrived. In Chapter 1, I canvassed three different attempts at formulating a debunking argument against non-inferential theistic beliefs, and found those attempts largely wanting. I then proceeded in Chapter 2 to present what I consider to be a more powerful debunking challenge, one which exploits facts about the diverse and mutually incompatible contents of the beliefs that arise from the mechanisms described by CSR. I argued that the fact of diverse contents can be used to formulate a kind of debunking argument on the assumption of a process reliabilist condition for knowledge. My contention was that this way of framing the *de jure* objection against religious beliefs is the strongest available version of such an objection. The task of this chapter will be to try to

show that even supposing the debunking argument presented in Chapter 2 is successful against religious beliefs that are partly the product of the cultural influences to which an individual happens to be exposed, nevertheless, beliefs about a certain class of what I shall call "core" religious propositions—that is, religiously relevant propositions in which humans are naturally disposed to believe across all cultures—can often amount to knowledge if theism is true. The content of core propositions includes that there is a personal explanation for the cosmos, that the cosmos is providentially ordered, and that there is moral law with a transcendent personal source. These core propositions are true if the God of Abrahamic monotheism exists, and moreover, knowledge of such propositions will amount to knowledge *de re* of the God of monotheism if He exists. I consider two distinct ways in which core propositions might be known. Firstly, in §3.2, I suggest that core propositions might be known entirely non-inferentially. Secondly, in §3.3, drawing upon some recent literature on knowledge from falsehood, I also suggest that an adherent of a religious tradition that is significantly mistaken may be able to arrive at knowledge of core propositions via simple deductive inferences from certain false propositions that she has accepted from her tradition. The significance of these claims will become fully clear in Chapter 4, where I shall argue that knowledge of core propositions would provide someone with the cognitive contact with God that is sufficient for having a personal relationship with God, even if only *de re* relationship.

§3.1 Preliminaries

§3.1.1 Core religious content

Core propositions are those propositions in which we are disposed to believe as a result of the CSR-described cognitive dispositions alone. Non-core propositions are those religious propositions that are not such that intuitive cognition alone inclines us towards belief if them; that is, non-core propositions are those propositions that are believed at least partly as a result of specific cultural inputs. In the Introduction to this thesis, we saw that the current CSR literature seems to indicate that belief of the following propositions is part of what Harvey Whitehouse (2008: 20) has termed the "cross-culturally recurrent religious repertoire":

(a) There is a personal first cause of the universe.

(b) There exists at least one disembodied superhuman agent who possesses beliefs, intentions, and desires, and who:

(c) possesses super-properties such as immortality, epistemic infallibity (i.e. never having false beliefs), and unlimited informational access;

(d) is interested in and aware of the moral conduct of human beings, including their secret conduct, and is liable to issue rewards and punishments based upon such conduct; and

(e) is responsible for designing apparently purposeful features of the universe.

We saw that non-reflective beliefs of such propositions are thought to be the causal product of a variety of cognitive dispositions working in concert, including the Hypersensitive Agency Detection Device (HADD), the Theory of Mind (ToM) module, a preference for concepts that are minimally counterintuitive, intuitive morality, and intuitive teleology. Let's give the name "HADD+" to this cluster of cognitive dispositions.

C. Stephen Evans (2010) has offered what is to my mind a helpful way of summarising the content of this pan-cultural awareness of divinity that arises from HADD+, suggesting that it might be seen as consisting of at least three fundamental intuitions, namely: a sense of "cosmic wonder"—that is, an intuition that there is a personal explanation for why there is a cosmos at all; an intuition that the cosmos is providentially ordered, which is to say, the product of a designing mind; and an intuition that there is an objectively-binding moral law with a transcendent personal source.[1] The core propositions that are the content of these intuitions will include both existential generalizations of the form "something is F" and propositions that predicate things of an entity designated by definite descriptions of the form "the *x* which is uniquely F is G."

[1] Notably, Tim Mawson (2005: Ch. 10) and Gary Gutting (1983: Ch. 5) have both argued that something like the core content I have just outlined is roughly the extent of what can be known (or at any rate, justifiably believed) on the basis of the collective testimony of human religious experience.

§3.1.2 Two ways of knowing

I said that I'm going to try to show that beliefs in core propositions that are arrived at in either of two ways will be able to count as knowledge if theism is true. Allow me to illustrate a little more clearly how I am thinking about these two ways of knowing, by way of a case that parallels the sort of religious cases with which I am concerned.

In the context of a dispute over Alvin Plantinga's (2002) evolutionary argument against naturalism, Jerry Fodor (2002) asks us to consider a race of creatures just like ourselves except that they falsely presuppose that all the objects in their environment are witches. Fodor suggests that this is the sort of scenario Plantinga has in mind when he claims that creatures with systematically false beliefs could nonetheless get by perfectly well in their environment. As Fodor describes the situation, "We think that tree is very big; they think that witchtree is very big; we think the cat on the mat is asleep; they think the witchcat on the witchmat is asleep… What you end up with is a creature who is about as adaptive as ourselves, but most of whose beliefs are false" (2002: 34). Fodor points out that creatures such as these could well end up with a lot of true beliefs even in spite of their error in thinking that every object is a witch: "If a creature believes *that appletree witch is blooming*, then it presumably believes that *that's an appletree* and that *that's a witch* and that *that's blooming*. And two of these are true beliefs that the creature shares with us" (2002: 34). Now, there are plausibly a couple of ways in

which these bewitched creatures could arrive at belief of the true propositions that Fodor mentions. These two ways correspond to the two ways in which core religious propositions may be known in spite of someone's being in significant error about many non-core religious propositions, so I shall argue.

Firstly, it could be that in addition to forming reflective beliefs in propositions whose content involves the concept *witch*, these creatures also non-inferentially form non-reflective (or intuitive) beliefs in many entailed propositions that don't involve the concept *witch*. Recall Justin Barrett's characterization of non-reflective beliefs, which we encountered in §0.2. Barrett writes that

> Nonreflective beliefs map closely onto what might be called tacit or intuitive knowledge. They are products of the intuitive system, and are likewise cognitively natural in McCauley's sense of having a high degree of automaticity and requiring little conscious effort to produce. These nonreflective beliefs are representations that we have whether or not we know we have them. They are nonreflective in that they do not require conscious, deliberate, reflective resources to form them. We might never be aware of many of our nonreflective beliefs even though they guide information processing, speech, and other actions… Nonreflective beliefs include such ideas as "Snakes are scary," "Rainbows have bands of color," "People have minds," "I am," "A definite article does not precede a proper name," "Unsupported objects fall," "The sun moves relative to the earth," and "Dogs have puppies." When cognitive scientists talk about what infants or chimpanzees think or know, they generally are referring to nonreflective beliefs. (2011: 48-49).

If it is cognitively constituted as we are, a creature who is looking at an appletree and who reflectively believes that "there is an appletree witch before me" will also, at least non-reflectively, believe a host of other, true propositions such as that "there is a solid object before me," "there is an object that is smaller than the sun before me," "there is a non-manmade object before me," and so on. Similarly, a religious believer who reflectively believes false propositions to which her culture testifies such as that "Zeus is the supernatural creator of the heavenly bodies" will also, if the CSR story is correct, have various non-reflective beliefs of less fine-grained propositions, such as that "there is a supernatural creator of the heavenly bodies," or "there is a supernatural creator," or even less specifically "there is a supernatural being."

There is a second way in which Fodor's bewitched creatures could arrive at true beliefs in spite of their persistent error. They might consciously draw very simple deductive inferences from witch-involving propositions to entailed, non-witch-involving propositions: for instance, from "that appletree witch is blooming" to "that appletree is blooming." Similarly, a religious believer who believes a false non-core proposition, such as that "Apollo is the author of the moral law," could draw a simple deductive inference from that proposition to the core proposition that "there is a supernatural, personal author of the moral law"—a proposition which is true if the God of monotheism exists.

In what follows I shall argue that both of these ways of arriving at beliefs of core propositions are liable to generate knowledge, given the truth of theism.

§3.1.3 A virtue-theoretic condition on knowledge

Now for one final preliminary. We have already encountered process reliabilism, the view that in order for one's belief to count as knowledge, the salient causal process type that produced one's belief must have a sufficiently high truth-ratio among its outputs. In recent years, several leading epistemologists have defended a very noteworthy variant on process reliabilism that draws upon the Aristotelian notion of achievement as success through ability, an approach known as *virtue epistemology*.[2] According to John Greco, one of the foremost defenders of virtue epistemology, "knowledge is a kind of success from ability… Put another way, knowledge is a kind of *achievement*, or a kind of success for which the knower deserves credit" (2010: 7). Ernest Sosa (2007, 2009), another key advocate of the view, employs the analogy of an archer shooting an arrow at a target. The archer's shot can be evaluated along several dimensions, including: (i) *accuracy*: how close the shot is to the bull's eye; (ii) *adroitness*: how competent is the shot; (iii) *aptness*: whether the shot is accurate *because* it is adroit—i.e., whether its accuracy *manifests* the archer's competence. The archer's shot counts as an achievement for which the archer deserves credit just in case his shot is accurate *because* it is adroit; just in case it is *apt*. According to Sosa, belief too is a performance that can be

[2] As we shall see in §3.2.2, there are in fact two rather different sorts of project that both go by the name of "virtue epistemology." So-called "virtue reliabilists," such as Greco (2010) and Sosa (2007, 2009), think of intellectual virtues primarily in terms of truth-conducive cognitive powers or dispositions, whereas "virtue responsibilists" such as Zagzebski (1996), Roberts and Wood (2007), and Baehr (2011), think of them rather in terms of habits of a person's intellectual character.

evaluated along these dimensions.[3] Knowledge, on this view, is apt belief: belief that is accurate because adroitly formed.

How does this view differ from simple process reliabilism? In at least two crucial ways, I suggest. Firstly, virtue reliabilism urges that process types be individuated at the level of a given individual's cognitive abilities or dispositions, meaning that the salient type that Jack instantiates when he forms a perceptual belief in good lighting at close proximity is different from the salient type that Jill instantiates when she forms a perceptual belief in good lighting at close proximity. As Greco notes, "Because abilities are reliable dispositions, the account is a version of reliabilism. Because abilities are person-level dispositions, the account is a version of agent reliabilism" (2010: 10). Secondly, whereas simple process reliabilism merely requires that the *existence* of one's belief be due to a reliable causal process, virtue reliabilism requires that the *accuracy* of one's belief be due to one's employment of a reliable process.[4] The distinction is an important one: as

[3] It might seem unintuitive to say that belief is a performance, given that beliefs are states and are typically unintentional. But Sosa (2011: 86) wants to say that the sense in which holding a belief is a performance is roughly the sense in which a bridge might be said to perform well under strain, and thus be creditable for withstanding.

[4] Kallestrup and Pritchard (2012b) note that one encounters in the literature two notably different readings of the all-important "because of" relation. On the view of Greco (2010: 71, 74-75), to say that the agent's cognitive success is "because of" her cognitive ability is to say that the agent's abilities form a chief part of the causal explanation of her success. By contrast, Sosa (2007: Ch. 5) and Turri (2011) hold that the "because of" relation is to be understood in terms of the manifestation of a disposition. On this view, to say that the agent's cognitive success is "because of" her relevant cognitive dispositions is to say that that success *manifests* those dispositions, in the way

Sosa notes, "Something may explain the existence of a certain entity, however, without even partially explaining why it has a given property" (2007: 95). I suggest that we state the virtue-theoretic condition in the following way:

> COGNITIVE ACHIEVEMENT: IF *S* knows that *p*, THEN the accuracy of *S*'s belief that *p* is (at least in significant part)[5] due to *S*'s stable and reliable cognitive abilities.

§3.2 Knowing Core Propositions—Way I: Intuitive, Non-Inferential Belief of Core Propositions

In what follows, then, I shall seek to show that beliefs of core propositions, arrived at in the two ways sketched out, are likely to satisfy the safety condition, the cognitive achievement condition, and the simple process reliability condition. Let's begin by considering the non-inferential way of arriving at knowledge of such propositions. The claim I will defend in this section is that beliefs of core propositions that are the product of maturationally natural intuitive cognition are very likely, given God's existence, to be safe, to count as cognitive achievements, and to

---

that a devastating forehand smash manifests Roger Federer's tennis abilities, whereas his great wealth doesn't manifest those abilities even though it is "because of" them in a different sense.

[5] The reason for the "at least in significant part" qualification is to allow for cases of testimonial knowledge, wherein the recipient of testimony shares the credit for her true belief with the person(s) who testified to her. For more on this, see Greco (2010: 80-84); Sosa (2007: 93-97).

be the products of a process type with a high truth-ratio.

§3.2.1 Safety

Recall how we stated the safety condition in the previous chapter:

> SAFETY: IF *S* knows that *p* in a world *w*, THEN there is no nearby world w^* in which *S* has a false belief about *p* with a relevantly similar causal history.[6]

My contention is that non-reflective beliefs of core propositions that are the product of HADD+ will likely satisfy the safety condition if theism is true. That is, if God exists, then it is (epistemically) likely to be the case that someone's belief of a core proposition *p*, formed in the aforementioned manner, is such that there is no nearby world in which she has a false belief about *p* by way of a relevantly similar causal history. Now, if God exists, a given core proposition *p*—say, the proposition that "there exists a transcendent moral lawgiver" or that "the person who created the universe is very powerful"—could not easily have been false. Presumably, God's existence and possession of the attributes in question is not, if actual, something that could easily have failed to be the case, even supposing God

[6] Recall that in the previous chapter, I said that in order for our safety condition to be able to deal properly with cases involving beliefs of necessary truths, we should understand "belief about *p*" somewhat loosely, so as to encompass belief that *p*, belief that ~*p*, and belief that *q* where *q* is obviously inconsistent with *p*.

doesn't exist necessarily or possess those attributes essentially. From the modal stability of a proposition *p* it follows that someone who correctly believes *p* could not easily have had a false belief *that p*. But, of course, the safety condition with which we are working is not so crude as to be satisfied just because the proposition that the subject believes is modally stable. It also matters for the safety of a belief of a core proposition *p* that the subject could not easily had a false belief *about p* (which, as I have circumscribed "about *p*," could be the belief that ~*p* or the belief that *q*, where *q* is obviously incompatible with *p*) by way of a similar causal mechanism as produced her actual belief that *p*. And it is this latter for which I need to argue. That is, I need to show that humans could not easily have come to have false beliefs in the negations of core propositions (or things that obviously entail the negations of core propositions) by way of causal mechanisms relevantly similar to those that in fact produce beliefs of core propositions. Suppose that theism is true and that a human subject *S* non-reflectively believes a true core proposition *p* as a result of HADD+. I can think of two potential reasons why one might conclude that there *are* close possible worlds in which *S* has false beliefs that are the product of a relevantly similar causal mechanism, and hence, that *S*'s belief that *p* is not safe.

One reason for thinking so is that among the various cognitive modules that comprise HADD+ is one that is evidently error-prone in certain contexts, namely, the Hypersensitive Agency Detection Device (HADD). Michael Murray summarizes the thought thus: "in light of the fact that HADD is hyperactive…it

generates many false positives and is thus unreliable" (2007: 394). If Jenny were staying the night in a big old house, for instance, and heard creaking floorboards, her HADD might well generate the belief that there is an agent moving about the place—perhaps a burglar, perhaps a poltergeist, depending on Jenny's background beliefs—even if there in fact isn't. So the thought goes, then, there are nearby worlds in which *S* has false beliefs that are the product of causal mechanisms relevantly similar to those that produced her beliefs of core propositions in the actual world. Now, since this sort of worry has received a good deal of discussion already, especially by Murray (2007, 2010), I shall be rather brief here. One thing to note is that a false belief such as that "there is an agent moving around the house" won't count as a belief *about p*, where *p* is a core religious proposition such as that "there is a supernatural creator of the universe." It isn't even obvious that the two propositions concern "a similar subject matter." But even if we set that issue to one side, there is another reason for thinking that the former belief isn't relevant to the safety of the latter. The tendency to believe in a supernatural creator and moral law-giver only arises given the disposition to postulate agency (i.e. HADD) *in conjunction with* a number of other cognitive dispositions including, crucially, the dispositions to believe that moral obligations have a transcendent source and that various features of the world exhibit purpose and design.[7] It is very doubtful, therefore, that a belief that is the product of HADD but not of these other cognitive modules should count as having a relevantly similar causal history

[7] See Barrett (2011).

(for the purposes of safety) to a belief that is the product of HADD+ —that is, HADD in conjunction with theory of mind, intuitive morality, intuitive teleology, and the rest. To underscore this point, it is worth noticing that many belief-producing mechanisms that produce beliefs that are (intuitively) instances of knowledge are comprised of elements that *on their own* would fairly easily produce false beliefs. For instance, suppose that someone forms a belief about roughly how far away a lightning strike is by way of a combination of hearing the crack of thunder and seeing the flash. Neither seeing the flash nor the hearing the crack of thunder taken in isolation from the other would produce a safe belief about the approximate distance of the lightning strike, but when used together, the resulting belief could not easily have been false. I suspect a number of testimony-based beliefs are like this too—at least, when one's belief is the result of receiving testimony from multiple sources. Suppose one hears on separate occasions from three different friends each with independent sources that so-and-so has had an affair recently, and suppose that each of these friends has something of a fondness for idle gossip. Had one's belief been based upon any one of these three bits of testimony but not on the other two as well, then one's belief may well have been unsafe. As it is, though, one could not easily have had a false belief about whether so-and-so had an affair by way of receiving the testimony of *all three* of these friends. I contend, then, that the fact that HADD+ contains an element that on its own would produce unsafe beliefs in certain contexts does not mean that true beliefs produced by HADD+, when all its elements are taken together, are unsafe.

Here is the other potential reason for thinking that there are close possible worlds in which *S* has false beliefs that are the product of a relevantly similar causal mechanism as produced her beliefs of core propositions in the actual world. The religious belief-dispositions to which HADD+ gives rise are held to be a mere by-product of evolution—an "evolutionary accident"—rather than something for which evolution directly selected.[8] One might think, then, that evolution could quite easily have thrown up a slightly different set of religious belief-dispositions in human beings, such that humans would instead have been disposed to mistakenly believe the negations of at least some core propositions (mistakenly, that is, given the truth of theism). Let's give the label HADD+* to those alternative dispositions.

Let's grant that knowledge-undermining luck—what Pritchard (2005: §6.1) calls "veritic luck"—would afflict true beliefs that are the product of HADD+ if there indeed are nearby possible worlds in which humans evolved HADD+* rather than HADD+. But what I want to argue is that, in fact, those worlds in which humans evolved HADD+* are *not* close worlds if theism is true. Rather, those worlds are very distant worlds if God exists. Put simply, if the God of classical theism exists, then it is epistemically highly likely that God has providentially acted in such a way that it isn't *lucky* that humans came to possess HADD+ rather than HADD+*; acted in such a way, that is, that there are no close worlds in

[8] As Paul Bloom puts it: "Religion is not an adaptation, then; it is not like colour vision or the love we feel toward our children. Rather, it should be seen as similar to the ability to understand calculus or the pleasure we get in playing video games—something that did not emerge due to its adaptive value, but rather is a by-product (or spandrel) of pre-existing adaptations" (2010: 119).

which humans evolved HADD+*. I appreciate, however, that there are a variety of proposals on the table concerning the metaphysics of divine providence, and I don't wish to anchor my contention to any one such theory.[9] I shall therefore try to secure this claim with respect to all three major accounts of providence: theological determinism, Molinism, and open theism.[10]

Before I go any further, I must define some terms that will be important in the discussion that follows: namely, *strong actualization*, *weak actualization*, and *probabilistic actualization*.[11] For God to strongly actualize a state of affairs *A* in a world *w* is for God to issue a decree *D* (e.g., "Let there be light!") in *w*, such that in every metaphysically possible world *w** in which God issues *D*, the state of affairs *A* (which is specified in *D*) obtains in *w**. Put another way, for God to strongly actualize *A* in a world *w* is for God to issue a decree that broad-logically entails *A*.[12] For God to weakly actualize a state of affairs *A* in a world *w* is for God to strongly actualize a state of affairs *A** in *w*, such that were *A** to obtain in *w* then *A* would obtain in *w*. That is, weak actualization of *A* involves strong actualization of a state of affairs *A**

[9] For an excellent survey of such theories, see Flint (2009).

[10] As I shall explain in §3.2.1.3, I intend open theism to be understood in such a way that it is compatible with the view that God is timeless, provided divine timelessness is not supplemented with theological determinism or exhaustive middle knowledge.

[11] The definitions that follow are largely those found in Plantinga (1974b: 173).

[12] Note that I am using the terms "metaphysical" possibility/necessity and "broad-logical" possibility/necessity interchangeably to denote one and the same modality. The modality in question is also sometimes spoken of in terms of "absolute" possibility/necessity. For a helpful account of various modalities and their interrelation, see Leftow (2012: Ch. 1).

that counterfactually entails *A*.[13] As we shall see shortly, weak actualization plays a crucial role in the Molinist account of providence, according to which God strongly actualizes sets of circumstances such that were a free creature to be in those circumstances, he or she would freely perform a certain action. Finally, for God to probabilistically actualize a state of affairs *A* is for God to strongly actualize a state of affairs A^* at *t*, such that were A^* actual at *t* then it is objectively probable at *t* that *A* will become actual, and such that *A* does in fact come about as a non-deviant causal result of God's strong actualization of A^* at *t*.[14] With these definitions in hand, we can now proceed to carve up the space of theories of divine providence.

§3.2.1.1 Theological determinism

According to theological determinism (known variously as Augustinianism, Thomism, and Calvinism), every contingent state of affairs that is actual is strongly

[13] If one is attracted to the Stalnaker-Lewis account of counterfactual conditionals (Stalnaker 1975; Lewis 2001), one might say that for God to weakly actualize *A* in a world *w* is for God to strongly actualize a state of affairs A^* in *w*, such that no A^* & ~*A*-world is closer to *w* than is any A^* & *A*-world.

[14] Talking about a state of affairs as coming to obtain *at* a particular time obviously requires that we individuate states of affairs in such a way that they are capable of obtaining at some times and not others, rather than as being indexed to times (e.g., the state of affairs of *JFK's being assassinated on November 22nd 1963*) and hence at all times either obtaining or failing to obtain.

actualized by God.[15] Put another way, on this view there are no contingencies that have not been divinely decreed.[16] It is important to note that theological determinism is fully compatible with *either* physical determinism or physical indeterminism. The conjunction of theological determinism and physical determinism entails that God has strongly actualized a set of physical laws which, in conjunction with the initial physical state that God has strongly actualized, broad-logically entails the complete physical state of the world at all future times. We can assume, plausibly I think, that strong actualization is closed under broad-logical entailment (at least, where contingent states of affairs are concerned), so that if God strongly actualizes any two contingent states of affairs *A* and *B*, then any contingent state of affairs *C* which is broad-logically entailed by the conjunction of *A* and *B* is also strongly-actualized by God.[17, 18] Hence, if God strongly actualizes an initial physical state and a set of laws (both of which are contingent states of affairs) that together broadly-logically entail the complete physical state of the world at all future times, then God thereby strongly actualizes the complete physical state of the world at all future times. That is how things stand if God has

[15] As Derk Pereboom puts it, "Theological determinism is the position that God is the sufficient active cause of everything in creation, whether directly or by way of secondary causes such as human agents… Everything that happens to us, to the last detail, is in accord with God's will" (2011: 262-63).

[16] I leave it open whether, on this view, necessary states of affairs are also divinely decreed. Nothing I am about to say turns on this issue.

[17] For a defence of this thought, see van Inwagen (1988: 218).

[18] The reason for stating this closure principle in a way that applies only to contingent states of affairs is to avoid the consequence that God strongly actualizes necessary states of affairs—an issue on which I wish to stay neutral.

created a deterministic world. But how do things look given the conjunction of theological determinism and physical *in*determinism? Well, in that case, the laws that God has strongly actualized are such that those laws in conjunction with the initial state *don't* broad-logically entail the complete state of the world at all future times; rather, they underdetermine the way in which history will unfold. But crucially, it is open to God to close the causal gaps by issuing decrees that fix the outcomes of physically indeterministic events. As Peter van Inwagen notes,

> The Greek atomists held that atoms—what are now called elementary particles—could swerve in the void, and something very much like this is true according to modern physics. If God's causal relations with the world are confined to continuously holding the elementary particles in existence and continuously supplying them with their [indeterministic] causal powers, then He does not decree the outcomes of such "swerves in the void," since such "swerves" are not determined by the causal powers of the particles. (1988: 224).

The crucial point is that, if theological determinism and physical indeterminism are both true, then God's causal relations with the world are *not* confined merely to continuously holding the elementary particles in existence and continuously supplying them with their indeterministic causal powers, but rather, God *additionally* decrees the outcome of every single physically indeterministic "swerve in the void."

On Duncan Pritchard's (2005: §5.1) analysis of luck, a state of affairs that is

lucky is one that actually obtains and yet fails to obtain in a wide class of nearby possible worlds with the same or very similar conditions prior to the time at which the state of affairs actually obtains.[19, 20] Given theological determinism, the extent to which it is lucky that some contingent state of affairs actually obtains is a matter of whether—all else remaining as equal as possible—God fails to strongly actualize that state of affairs in some nearby worlds, or put in more ordinary terms, a matter of *how easily* God could have failed to strongly actualize that state of affairs. And how easily God could have failed to strongly actualize a state of affairs is a function of the *depth* of God's actual desire for that state of affairs. What do I mean by talking of the "depth" of a desire? Well, what I have in mind is the affective analogue of what Plantinga terms the "depth of ingression" of a belief in someone's noetic structure. Plantinga writes that

> [A]n account of S's noetic structure would include something like an index of *depth of ingression.* Some of my beliefs are, we might say, on the periphery of my noetic structure. I accept them, and may even accept them firmly, but I could give them up without much change elsewhere in my noetic structure… On the other hand,

---

[19] Pritchard's (2005: 125-33) other necessary condition for luck is that the event's occurrence is important to some agent.

[20] The reason we shouldn't insist that the prior conditions be kept *exactly* fixed is that this would entail that no event in a deterministic universe is lucky. Yet, as Williamson notes, "determinism does not trivialize safety [the converse of luck]" (2000: 124; 2009: 325). He points out that the ease with which a ball balanced on the tip of a cone could have fallen off is not lessened just because it is located within a deterministic universe. After all, the initial conditions of that universe could so easily have been slightly different.

> … suppose I were to come to think there had not been much of a past (that the world was created just five minutes ago, complete with all its apparent memories and traces of the past) or that there were not any other persons: these changes [to his noetic structure] would have great reverberations; these beliefs of mine have great depth of ingression into my noetic structure. (1983b: 50).

Analogously, a desire is *deeper* the closer it is to the centre of a person's "affective structure"; that is, the more deeply ingrained it is in her system of desires. A shallow desire—a whim—is a desire that is on the periphery of one's affective structure, and which is such that getting rid of that desire would have little overall impact on one's affective structure. Note that a whim can nonetheless be a *strong* desire, as, for instance, when a child is momentarily captivated by a toy in a shop window to the point of being overwhelmed by the desire to have it, and yet has forgotten about it the very next day. I would suggest, then, that it is the *depth* of a desire, much more than its strength, which determines a desire's modal stability. A deep desire is one that a person could not easily have failed to have, whereas a strong desire may or may not be one that a person could easily have failed to have.

Suppose that in actuality I have 120,000 hairs on my head, and that God has decreed that this be so. I have no idea whether this is so, but let's imagine that it is pretty much a matter of indifference to God whether I have 100 hairs more or less than the number I actually have. Were that the case, then the world in which God decrees that I have 111,990 hairs (and as many other things as possible remain the

same) would be a very close world. On the other hand, let's suppose that in the actual world God has a very deep desire that there be rational, embodied creatures, and has decreed that it be so. Plausibly, if we're talking about the God of Christian theism, such a desire would indeed be a deep desire of God's—a fact having to do with God's perfectly loving character, His being triune and so in some sense instantiating personal relationship within Himself, His deep desire to share loving personal relationship with others besides Himself, and so on. In that case, the nearest world in which God creates no such creatures (and in which everything else remains as similar as possible) is very far from the actual world, because God's desires would have to have been very different from what they actually are in order for God to have instead actualized that world. In general, then, we can say that on theological determinism, the closeness of a world *w* to the actual world is a matter of how different God's desires would have to have been in order for *w* to have been actual. There is a very real sense in which the modal landscape, on theological determinism, is shaped by God's desires and their relative depths of ingression in God's affective structure. The question, then, is how deep is God's actual desire for humans beings to evolve HADD+ rather than HADD+*? Plausibly, that desire is deep indeed; God's strong actualization of the state of affairs in which humans have HADD+ rather than HADD+* would not plausibly have been the result of mere whim. Indeed, God would plausibly be deeply averse to the state of affairs in which humans existed but were naturally inclined to believe falsehoods concerning core propositions—a fact that has to do

with God's deep desire to relate personally to human beings, and the way that human beings cannot relate personally to God without possessing at least a minimal cognitive grasp of God of the sort involved in knowing core propositions (this latter is something I shall argue for in Chapter 4). And if it is true that God would have a deep desire for humans not to evolve HADD+*, then the world in which humans evolved HADD+* is not close to the actual world, regardless of how lucky it may appear to be from the point of view of evolutionary science that humans wound up with the cognitive traits that we have.

§3.2.1.2 Molinism

On the Molinist view, God's providential governance of the world proceeds by way of a mixture of strong actualization and weak actualization. Molinists, as noted earlier, hold that when deliberating about which world to create, God finds Himself confronted with an array of brute contingent truths over which He has no control. As Dean Zimmerman puts it, "it is as though God 'wakes up' to find certain contingent things true—there is an independent source of contingent fact at work 'before' God has a chance to do anything about it" (2009: 46). The truths in question are subjunctive conditionals, and they can be grouped into two categories. A counterfactual of creaturely freedom (CCF) is a conditional of the following form: if an agent *S* were in the circumstance C, then *S* would freely perform action

A.[21] And a counterfactual of physical indeterminacy (CPI) is a conditional of the following form: if a particle of physical matter P were in the circumstance C, then P would indeterministically follow trajectory T. Molinism holds both that every such counterfactual proposition is definitely either true or false, and that God has exhaustive "middle knowledge" of the truth-values of all such counterfactuals.[22] In deliberating about which world to create, God decides which initial set of circumstances to strongly actualize, knowing—from His exhaustive middle knowledge—which other states of affairs are counterfactually entailed by, and thus weakly actualized in virtue of, those states of affairs that He strongly actualizes. Hence, on

[21] See Zimmerman (2009: 48-68) for a detailed exploration of how to give an account of the truth-conditions for such conditionals.

[22] In fact, virtually no Molinist writings of which I am aware make any mention of CPIs, which is odd considering that a Molinist account of providence is radically incomplete without an appeal to God's exhaustive middle knowledge of CPIs. So, in claiming that Molinism "holds" that there is an exhaustive set of CPIs, what I really mean is that Molinists are in fact committed to the claim that there is such a set and that God has exhaustive epistemic access to it. The reason that Molinism is committed to these claims is as follows. Given that Molinists are supposed to be libertarians about free will, they must hold that God has created a physically indeterministic universe. And so, if God is to know with certainty at the moment of creation *which* sets of physical circumstances will eventuate at some future time *t* (such that were a human agent to be in those circumstances at *t*, he or she would freely perform a certain action), then God will have to have a way of knowing with certainty how the indeterministic physical universe will unfold up until *t*, and hence, He will have to have access to an exhaustive set of CPIs at the moment of creation. The only reference I can find to the Molinist's need to invoke brute, contingent CPIs is in a paper by Zimmerman, who writes that "The Molinist…denies that this [the Stalnaker-Lewis account] is the right way to think about the truth conditions for subjunctives describing what would happen in indeterministic settings... There can be [according to the Molinist] 'brute facts' about what would happen if this or that indeterministic situation were to obtain—facts that are not settled by the nearness of worlds, at least if nearness is measured by the categorical facts about the past plus the (indeterministic) laws" (2009: 55).

the Molinist picture, every single contingent state of affairs that is actual is either strongly or weakly actualized by God.

How, then, are we to think about the question of whether it is *lucky* that some state of affairs is actual, given Molinism? I think it will be helpful at this point to employ a pictorial analogy. Imagine that prior to creation, God is confronted by an absolutely vast array of on-off switches. These switches represent all the various possible states of affairs that God might choose to strongly actualize. Above the switch array is an equally vast row of lights, accompanied by a wiring diagram indicating how the switches have been wired to the lights, such that a given switch (or combination of switches) would activate a given light (or combination of lights)—i.e., indicating that if such-and-such switches were activated then such-and-such lights would light up. The activated lights represent the states of affairs that are weakly actualized in virtue of God's strong actualization of states of affairs that counterfactually entail them, while the wiring configuration represents the CCFs and CPIs that are brutely, contingently true. The Molinist thought is that God had no control over how the wiring configuration was set up in the first place, and hence, no control over which such counterfactuals are true. Now, imagine that above the row of lights, there is a second row of lights, and a wiring diagram indicating which combinations of activated lights on the first row would lead to which combinations of activated lights on the second row—indicating, that is, that if such-and-such lights on the first row were activated then such-and-such lights on the second row would be activated—and imagine that that there is yet a

third row of lights, again wired up so that certain combinations of activated lights on the second row would activate certain combinations of lights on the third row, and so on and on. We can imagine that the lighting board consists of vastly many such rows, wired up in the aforementioned manner. Each overall configuration of the switch array and lighting board represents an entire world history—a possible world. We can see why, on the Molinist picture, God's creative choice is significantly constrained: confronted as He is by the brute contingent facts of the wiring setup that happens to obtain, God is limited to choosing between lighting configurations that are available to Him *given* that wiring setup[23]—He is limited, that is, to choosing between "feasible" configurations (which represent what Molinists term "feasible" worlds; worlds that God can actualize *given* the CCFs and CPIs that happen to be true).

How, then, should we think about the question of whether, given Molinism, there is a close world in which humans evolved HADD+*? Let's use the term *God's plan* to denote the totality of God's preferences concerning the activation state of each switch and each light. Now, it is epistemically likely that, since He had no control over it, the wiring setup that God got lumped with didn't permit Him to realize the entirety of His plan. Indeed, it is epistemically rather likely that God had to make some "trade-offs." Suppose God desires the activation of two particular lights, call them *x* and *y*, but that there is no feasible lighting configuration that

[23] This thought is crucial to Plantinga's famous Free Will Defence (1974a; 1974b: Ch. 9), according to which it is broad-logically possible that the CCFs that God is confronted with are such that there is no world He can actualize which contains moral good but no moral evil.

includes the activation of both *x* and *y*. In that case, God will have to choose between *x* and *y* based on (*inter alia*) the relative depth of His desires for them.

In general, if there is any light or set of lights, *y*, such that *x* and *y* cannot simultaneously be activated within any feasible overall configuration, and which is such that God's desire for the activation of *y* is almost as deep as His desire for the activation of *x*, then God's desires need not have been very different in order for *y* to have been activated instead of *x*. Putting things back in terms of worlds and states of affairs, and supposing that the state of affairs *A* obtains in the actual world, α: if there is any state of affairs or set of states of affairs *B*, such that *B* and *A* do not both obtain in any single feasible world, and such that God's desire for the actuality of *B* is almost as deep as His desire for the actuality of *A*, then the nearest feasible world to α in which B obtains instead of *A* is *close* to α.

Now, the issue we are concerned with isn't simply whether there is a close world in which humans failed to evolve HADD+, but rather, whether there is a close world in which, instead of evolving HADD+, humans evolved HADD+* and hence were disposed to form false beliefs about core propositions.[24] So what we are interested in is whether there is a close world in which humans evolved HADD+*. Let's give the label "~*A*" to the state of affairs in which humans don't

[24] As Pritchard (2005: §5.2) notes, knowing that *p* is compatible with its being lucky that one is even in a position at all to form a belief about *p*. That is, one can know that *p* in the actual world even though in a wide class of nearby possible worlds, one fails to form a belief about whether *p*. The troublesome variety of epistemic luck—what Pritchard (2005: §6.1) terms "veritic luck"—is present only if there are nearby worlds in which the subject forms a false belief about the target proposition by way of a relevantly similar method.

evolve HADD+*. ~*A* obtains in the actual world, α. The question is: given that God exercises providence in the manner envisaged by Molinists, is there a world close to α in which *A* (the state of affairs in which humans *do* evolve HADD+*) obtains? Since we mere mortals have no more than the tiniest glimpse at the truth-values of the CCFs and CPIs, we don't know what dilemmas and trade-offs God may have faced in making His creative decision. The most we can really do is to speculate about the relative depths of God's desires for various states of affairs, and on the basis of such speculation, to make judgments about how different God's desires would need to have been from what we suppose they actually are—assuming trade-offs between ~*A* and various other desirable states of affairs had to be made—in order for God to have instead actualised a world in which ~*A* fails to obtain (and hence, in which *A* obtains). Now, presumably there indeed are non-actual states of affairs that God desires and that, for all we know, had to be sacrificed for the sake of securing ~*A*. The question is: how much deeper is God's desire for ~*A* than for those other states of affairs? I shall argue in Chapter 4 that the ability of humans to relate personally to God depends upon their having knowledge at least of core propositions. If this is so, and if divine-human relationship is indeed as important to God as the major monotheistic traditions have claimed, then God's desire that humans not have false beliefs about core propositions, and hence that ~*A* obtains, will be very deep indeed. Moreover, I submit that it is hard to think of any non-actual state of affairs, *B*, which is such that (supposing that God could not, given the CCFs and CPIs that obtain, have both *B*

and ~*A* be actual) God's desires need only be *slightly* different than they actually are in order for God to have opted for *B* over ~*A*. After all, no matter how bad is the "hand" that God gets dealt with respect to the CCFs and CPIs, the option always remains for God not to create a world at all, as does the option to create a world without any humans or similar agents.

§3.2.1.3 Open theism

The open theist view of providence denies that every single state of affairs that is actual is either strongly or weakly actualized by God. Open theism affirms that God has created a physically indeterministic universe, but denies that God is in the business of decreeing the outcome of every single physically indeterministic occurrence (contra theological determinism) and denies that there is an exhaustive set of determinately true or false CCFs and CPIs on the basis of which God can weakly actualize free creaturely actions and indeterministic events (contra Molinism). Rather, when God wishes to bring about a certain state of affairs *A* that is such that God cannot simply strongly actualize it—prime examples being states of affairs involving free human actions, which could not be genuinely free if God were to strongly actualize them—the best God can do is to attempt to probabilistically actualize *A* by strongly actualizing a state of affairs A^* at *t* such that if A^* were actual at *t* then it is objectively probable that *A* would become actual. As Thomas Flint puts it,

> Knowledge of such "would-probably" conditionals, say Openists, gives God enough information to operate providentially. Take Framboise… Perhaps God believes that Framboise and others would benefit greatly were she to meet a marriageable young gent. And maybe he sees that there's a situation he can put her in such that, if in that situation, she'd *probably* freely decide to order a raspberry sundae. And, upon ordering it, she'd *probably* strike up a conversation with that nice young sundae-maker (and raspberry afficionado) Maraschino, who'd *probably* ask her out on a date, which (since God sees that the two are a perfect match) she'd *probably* accept, and so on. (2009: 271).

Importantly, this so-called "risky" view of providence is quite compatible with the view that God is timeless and hence has timeless knowledge rather than foreknowledge[25]—provided, that is, that divine timelessness is not supplemented with theological determinism or exhaustive middle knowledge. The advocate of divine timelessness may quite consistently hold that God has timeless knowledge and yet takes risks in governing the world (i.e. through failing to either strongly or weakly actualize every state of affairs that becomes actual), by holding that God has *stages* in His timeless knowledge, such that in making providential decisions concerning some particular stage of the world's history, "God can ignore or somehow 'bracket' parts of what he knows, rendering them irrelevant to his decision to include this or that in his overall plan for the world" (Zimmerman 2009: 42).

As I said, on the view of providence we are now considering, there are some

[25] For a book-length defence of divine timelessness, see Leftow (1991).

states of affairs that God just *cannot* strongly or weakly actualize. On this view, a free creaturely action cannot be strongly actualized by God because in that case it simply would not be a *free* action. And in vastly many cases, neither can a free creaturely action even be weakly actualized by God, because the relevant CCFs are neither true nor false, there being nothing in virtue of which it is true that the creature *would definitely* freely do this or that in a given situation, prior to her actually performing the relevant free actions.[26] That is why, on this view, God can be said to take risks with respect to free creaturely actions in particular. With that said, there doesn't seem to be any barrier in principle to the idea that God might—perhaps somewhat often—strongly actualize the outcomes of physically indeterministic events that do not directly bear upon free creaturely actions. Accordingly, even on this view of providence it is plausible to think that God would act so as to *guarantee*—i.e., strongly actualize—the occurrence of events about which He cares greatly, insofar as His guaranteeing those events does not compromise human freedom. And I would suggest that the emergence of HADD+ in human beings is one such occurrence.

[26] The rejection of the Molinist notion that there is an exhaustive set of CCFs each of which is either true or false is motivated in part by the grounding objection, a powerful version of which states that since—very plausibly—truth supervenes upon being (i.e., a change in the truth-value of a proposition necessarily corresponds to a change in the state of the world), and since, by the Molinist's own admission, there is nothing in the world upon which the truth-values of most CCFs supervenes prior to the performance of the relevant free creaturely actions, there is no reason to think of all CCFs as having truth-values. For more on this see Perszyk (2013: §3.1). Brian Leftow has suggested (in conversation) that the grounding objection vis-à-vis CPIs is perhaps even stronger.

Perhaps there is an objection in the vicinity, though. Might it be true that by "intervening" in the workings of the physical world in the manner I am suggesting, God in fact *does* compromise human freedom, which requires that there be stable natural regularities? According to George Ellis, "[I]t seems probable that fixed laws of behaviour of matter, independent of interference by a Creator or any other agency, is a requisite basis of existence of independent beings able to exercise free will, for they make possible meaningful complex organized activity without outside interference" (Cited in Plantinga 2011: 99). Presumably the thought here runs something like this: morally significant freedom requires that one have the ability to choose between various possible courses of action without coercion and with a high degree of confidence about what the immediate consequences of one's actions will be. It requires, for instance, having confidence that a bullet fired from a gun at another person is very likely to kill or dreadfully injure that person, and not at all likely to morph into marshmallow in mid-air and harmlessly tickle her. And this in turn requires that there be strong physical regularities.

All of this seems very plausible. But there is in fact no reason at all to think that God would have to do anything that was remotely noticeable by humans in order to ensure that human beings end up with HADD+.[27] All that activity of determin-

[27] And indeed, Plantinga argues that even if humans *did* notice instances of special action, their morally significant freedom need not be compromised. He writes that, "sensible free action does not require that God never intervene. Ric [a rock climber] reaches the top [of a rock face]; the fastest way down would be to jump; he's not tempted though, because he knows that a 150-foot fall

ing otherwise indeterministic outcomes could take place at the level of genetic replication, and over enormously long aeons—vast stretches in which humans did not yet exist, or existed but were thousands of years away from developing the means for finding out about DNA replication and the like. By the time humans did acquire the wherewithal for studying the biochemical world, the human religious belief-disposition had long been in place.

So, I conclude that on any of the major accounts of providence, it can be shown that it is epistemically highly likely that God would providentially act in such a way that it is not lucky that humans possess HADD+ rather than HADD+*. If theism is true, it is very likely the case that intuitive beliefs in core propositions that are the product of HADD+ are safe.

§3.2.2 Cognitive achievements through maturationally natural abilities

Can someone's belief in a true core proposition count as a cognitive achievement when that belief is the product of the non-inferential cognitive mechanisms described by CSR, namely, HADD+? Recall how we stated the cognitive achievement condition earlier:

---

would kill or injure him. Now suppose he also believes that God occasionally intervenes, causing someone who takes such a fall to survive unhurt; that still won't tempt him to jump. All that's required for purposeful free action is reasonable confidence in substantial regularity in the neighbourhood of the proposed action. And that's certainly compatible with God's sometimes intervening" (2011: 104).

> COGNITIVE ACHIEVEMENT: IF *S* knows that *p*, THEN the accuracy of *S*'s belief that *p* is (at least in significant part) due to *S*'s stable and reliable cognitive abilities.

What I'm going to claim is that if true, beliefs of core propositions that are the product of HADD+ will count as cognitive achievements to at least the extent that various other true, non-inferential beliefs that are the product of intuitive cognition count as cognitive achievements. Now, it is hard to avoid the fact that a "cognitive achievement" sounds like something that requires some deliberate effort, whereas forming non-reflective beliefs via one's maturationally natural cognitive faculties involves little or no conscious effort. Recall Barrett's characterization of non-reflective beliefs:

> They are products of the intuitive system, and are…cognitively natural in McCauley's sense of having a high degree of automaticity and requiring little conscious effort to produce. These nonreflective beliefs are representations that we have whether or not we know we have them. They are nonreflective in that they do not require conscious, deliberate, reflective resources to form them. (2011: 48)

Clearly, there is an important question concerning whether the accuracy of beliefs that arise in this more-or-less automatic fashion can be said to be due to the believer's abilities in the relevant sense. But before I tackle that question, there are two other issues I want to deal with briefly.

One issue is that of cognitive integration. On the virtue-theoretic picture advocated by Greco, Sosa, and others, a belief-producing mechanism won't count as a cognitive *ability* at all unless it is integrated into one's cognitive system in a certain way. According to Greco, "cognitive integration concerns the range of outputs—if the products of a disposition are few and far between, and if they have little relation to other beliefs in the system, then the disposition is less well integrated on that account" (2010: 152). Consider Plantinga's (1993b: 199) example of a freak brain lesion that causes its possessor to believe that he has a brain lesion. Intuitively, we want to say that the subject's true belief here is not an item of knowledge. And Greco thinks that the explanation why it isn't knowledge is that the brain lesion is not sufficiently well-integrated into the subject's cognitive architecture to count as an ability:[28] "The process produces only a single belief, for example, and it is unrelated and insensitive to other dispositions governing the formation and evaluation of belief" (2010: 152). Now, as we saw in the Introduction to this thesis, HADD+ in fact comprises a cluster of cognitive dispositions, including a disposition to postulate agency (HADD), a disposition to ascribe beliefs, desires, and intentions to putative agents (theory of mind), a disposition to believe in objective-

[28] This also appears to be Pritchard's judgment, when he writes that the brain lesion subject "does not have knowledge in this case, and the reason for this is that his beliefs are true *despite* his cognitive abilities and not because of them" (2012: 263). As Pritchard notes, the subject's belief about the presence of the lesion appears to be safe, since there is no nearby world in which the subject is mistaken about whether he has a lesion as a result of the same causal process. Hence, cases like this motivate the addition of a cognitive ability condition to an account of knowledge that already features a safety condition.

ly-binding moral obligations (intuitive morality), a disposition to ascribe design and purpose (intuitive teleology), and so on. Each of these dispositions has a considerable range of outputs; each such disposition is well integrated into our cognitive architecture, and as such, each could be considered a faculty in its own right. Many paradigm cognitive abilities are built up out of a number of more basic abilities. For example, the ability to read involves visual perception in conjunction with various recognitional abilities as well as abilities to comprehend language. Insofar as HADD+ is built out of more basic abilities that are themselves well integrated into our cognitive architecture, it seems plausible to think that HADD+ itself is well integrated.[29]

The other issue is this. Some of the deliverances of intuitive cognition—and some non-reflective beliefs of core religious propositions may be like this[30]—are such that there is no event of coming to believe; rather, they are such that one simply believes them from the start. The various non-reflective assumptions that John Searle (1995) contends are among "the Background"—the set of assumptions that are a precondition of our engaging with the world—would seem to be like this. Searle has in mind such assumptions as that there is a reality that exists

[29] Plantinga (2000: Ch. 6) in fact seems to think that God would endow human beings with a *sui generis* faculty for producing beliefs about God—a *sensus divinitatis*. Were that the case, though, it might be rather less plausible that the faculty in question is well-integrated into our cognitive architecture.

[30] This would seem to be what Barrett's "preparedness" account suggests, according to which children even from a very early age have non-reflective beliefs about superhuman agency possessed of such properties as immortality, omnipresence, and epistemic infallibility. See Barrett and Richert (2003).

independently of our representations of it, that causation is a genuine mind-independent relation, that physical objects are spread out in space, that truth is a matter of correspondence with reality, and so on. Can such non-reflective assumptions, if true, meet the cognitive achievement condition? Greco has offered a suggestion regarding this matter:

> Here is a proposal for dealing with this sort of case. Knowledge is a kind of success from ability, we continue to say, but this can be taken in two ways. In the standard case, knowledge is true belief resulting from the exercise of one's intellectual abilities. In the limit case, knowledge is true belief that (partly) *constitutes* one's intellectual abilities. The idea is that human beings might very well have some of their cognitive faculties structured by information that is hard-wired. (2010: 85)

This doesn't seem too implausible. On the other hand, perhaps a state that one maintains through ability, even though one never underwent a process of arriving at that state, can count as a success that is the result of (rather than being *part of*) one's abilities. Here's an example. Imagine that a weightlifter simply pops into being in something like the fashion in which Donald Davidson's "swampman" comes about.[31] Imagine, further, that this weightlifter pops into existence whilst holding up a very hefty set of dumbbells, and that due to his great strength he continues for a very long time to keep them aloft. Now, it seems plausible that the

[31] Swampman is a molecule-for-molecule duplicate of Donald Davidson, who pops into being *ex nihilo* as a result of lightning strike in a swamp. See Bergmann (2006: 147-49).

weightlifter's holding the dumbbells up is an achievement; that it is a kind of success that is down to his abilities. But, clearly enough, there was no event during which the weightlifter *began* to lift the dumbbells. We might think of innate beliefs as being like this. Even though the believer never underwent a process of arriving at such beliefs, nonetheless, her continuing to hold the belief is creditable to her cognitive abilities.

To return to the question I raised earlier, is it possible for a cognitive success to count as a cognitive achievement if it is the result merely of the exercise of maturationally natural cognitive dispositions that one has not consciously worked at cultivating? I think the answer is *yes*. But this is a question that in fact touches upon an issue that crucially divides two sorts of virtue epistemologists.[32] Virtue reliabilists such as Greco and Sosa are content to allow that among our knowledge-yielding cognitive abilities, or intellectual virtues, are various sub-personal processes of which the subject may not be reflectively aware and for which she need not be able to take responsibility, in the sense that their reliable operation need not contribute to what Jason Baehr (2011: Ch. 6) terms a subject's "personal intellectual worth." Sosa writes that "there are also relevant competences that operate sub-personally, are not reason-based, and yet yield knowledge. Our basic perceptual and mnemonic abilities can operate largely sub-personally while still delivering much knowledge" (2011: 86). Virtue responsibilists, by contrast, understand intellectual virtues not in terms of reliable cognitive faculties,

[32] For a very informative account of what divides virtue reliabilists from virtue responsibilists, see Baehr (2011: Ch. 4). See also Greco's (2000) review of Zagzebski (1996).

but rather, in terms of person-level traits or habits of a certain sort. According to Baehr, "an intellectual virtue is a character trait that contributes to its possessor's personal intellectual worth on account of its involving a positive psychological orientation toward epistemic goods" (2011: 102). Examples of such traits might be things like open-mindedness, conscientiousness, intellectual honesty, self-awareness, love of the truth, intellectual courage, intellectual humility, and so on.[33] In very much this vein, Linda Zagzebski defines knowledge as "a state of true belief arising out of acts of intellectual virtue" (1996: 271). She further spells out this definition in the following way:

> [I]t requires the knower to have an intellectually virtuous motivation in the disposition to desire truth, and this disposition must give rise to conscious and voluntary acts in the process leading up to the acquisition of true belief (or cognitive contact with reality), and the knower must successfully reach the truth through the operation of this motivation and those acts. (1996: 273).

A crucial feature of this way of thinking about intellectual virtues is the notion that the virtues at issue are the sorts of things that need to be cultivated. Indeed, Jay Wood and Robert Roberts contend that "An intellectual virtue is an *acquired* base of excellent epistemic functioning [emphasis mine]" (2007: 95).

Now, if the responsibilist understanding of intellectual virtue is correct, and if

[33] Roberts and Wood (2007) offer excellent chapter length explorations of each of a number of such intellectual character traits.

satisfying the cognitive achievement condition entails arriving at the truth through acts of intellectual virtue understood along responsibilist lines, then clearly there is a problem for my claim that religious believers can satisfy the cognitive achievement condition in coming to believe true core propositions as a result of (*inter alia*) the operation of maturationally natural cognitive dispositions in whose development they have played little or no effortful role. Allow me to say three things by way of response.

Firstly, it seems to me that virtue responsibilists are right to be concerned about the way in which so much of the focus of contemporary epistemology is on the epistemic status of cases in which the intellectual character of the believing subject is more or less irrelevant.[34] And they are right, moreover, to be concerned about the heavy focus upon propositional knowledge, largely to the neglect of other epistemic goods such as knowledge by acquaintance, understanding, and wisdom.[35] But as far as I can see, there is no reason *in principle* why the virtue reliabilist accounts of Sosa and Greco cannot fully accommodate the responsibilist's intellectual character traits, by seeing them as virtues that are genuinely necessary for attaining knowledge and understanding in *certain* domains—domains in which getting at the truth is intellectually demanding. Examples might be the domains of astrophysics, archaeology, literary criticism, crime scene investigation,

[34] Zagzebski points out that "Because the examples often used as paradigms are on the low end of the scale, little effort is made to distinguish between, say, a person who has real understanding of her environment and one who merely knows that the room she is in has four walls" (1996: 273).

[35] For an exploration of this wider range of goods to which epistemologists might turn their attention, see Roberts and Wood (2007: Ch. 2).

economic forecasting, and indeed, philosophy. In fact, Greco explicitly welcomes the idea that a full-orbed virtue epistemology will appeal to both faculties and character traits, when he affirms what he sees as Aquinas's understanding of intellectual virtues: "some *powers* are virtues, but…some *habits* are virtues too" (2000: 180). Notably, Sosa (1991: 225, 78) employs a distinction that may hold part of the key to seeing how to integrate these two aspects. He distinguishes between *fundamental virtues*, responsible for producing non-inferential or foundational knowledge—for instance, eyesight, memory, and logical intuition—and *derived virtues*, which are involved in drawing inferences that produce inferential knowledge. Virtue responsibilists stress that one cannot analyze "high-grade" knowledge—that is, knowledge that is acquired through processes of intellectual inquiry of some sort—without making reference to the subject's intellectual character.[36] As far as I can see, the virtue reliabilist can happily accept this point, noting that knowledge of this loftier sort can indeed be attained only through the exercise of the derived virtues (in addition, presumably, to the fundamental virtues), which will almost certainly include a range of intellectual character traits. That is, everything the virtue responsibilist wants to affirm regarding high-grade knowledge and understanding and the importance of intellectual character in attaining those epistemic goods can be affirmed also by the virtue reliabilist, and *without* having to deny the status of knowledge to cognitive successes that aren't the product of intellectual character virtues but rather of merely reliably functioning

[36] See, e.g., Roberts and Wood (2007: Chs. 1 and 2).

faculties. This leads me onto my second point.

Whilst, as I have just indicated, I have no quarrel with the virtue responsibilists' plea for epistemologists to pay more attention to the phenomena of high-grade knowledge and understanding, this should not come at the expense of writing off less impressive cognitive successes—ones that are due merely to the reliable operation of one's maturationally natural cognitive faculties—as wholly unworthy of the title of "knowledge." As Greco rightly notes, "there are paradigmatic cases of knowledge that seem not to involve that sort of virtue [i.e. traits of intellectual character] at all. It is a hard sell, for example, that such virtues are always and essentially involved in cases of perceptual knowledge" (2010: 10). What's more, it seems that these less hard-wrought cognitive successes are almost always essential to the attainment of high-grade knowledge. Consider Sherlock Holmes' eventual discovery that Jack Stapleton was a potential heir to the Baskerville estate and hence would have had a powerful motive for trying to dispense of Henry Baskerville. This is piece of high-grade knowledge if anything is. But it isn't as though Sherlock's inquiry in *The Hound of the Baskervilles* yields just one big piece of high-grade knowledge and nothing else. Rather, it involves the accumulation of a wealth of "low-grade" cognitive successes—all manner of episodes of perception, receipt of testimony, memory recall, consulting intuitions about counterfactuals, and so on, which successes are attained for the most part through the reliable operation of maturationally natural cognitive dispositions. To be sure, Sherlock's intellectual character traits *do* play an essential role in explaining how he comes to

his astounding overall conclusion—after all, Watson has most of the same items of low-grade knowledge to work with, and yet comes nowhere near to seeing how to fit them all together. But the crucial point here is that the raw material for an intellectual inquiry, for which intellectual character virtues are undoubtedly essential, consists in items of knowledge many of which are the product merely of the reliable operation of the subject's maturationally natural cognitive dispositions. This less impressive kind of knowledge is nonetheless essential for attaining the kind of knowledge that the responsibilists so prize, and ought to be recognized as such.

My third point is closely related to the previous. In fact, I doubt that the virtue responsibilist will have any qualm with my claim that high-grade knowledge cannot be attained independently of a range of cognitive successes that are the result merely of the reliable operation of one's native cognitive dispositions. It is perhaps just that she desires the recognition that there is more than one kind or quality of knowledge, and that knowledge that results from the exercise of intellectual character traits at which one has worked hard at acquiring is *more valuable* than knowledge that results merely from the reliable functioning of one's native faculties or dispositions. Indeed, I have no difficulty in accepting such a claim. As Sosa (2007: 24) rightly notes, thinking that there is more than one kind or quality of knowledge need not commit one to a linguistic thesis according to which

"knowledge" is an ambiguous term.[37] And crucially, the thought that some items of knowledge represent greater achievements than others, and hence are more valuable than others, is quite consistent with our thinking about the nature of achievements in general. Recall that in the film *The Matrix*, Neo and his comrades have various combat abilities simply "downloaded" into themselves before they embark upon a hazardous mission to rescue Morpheus, who is being tortured by Agent Smith inside the Matrix. At the press of a few buttons, Neo comes to possess a fully-fledged set of Kung Fu fighting abilities—no painstaking training over many years is required. These abilities of Neo's are certainly not "acquired" in the sense that virtue responsibilists care about. Once inside the Matrix, Neo successfully incapacitates numerous agents through the use of his fearsome Kung Fu abilities. Does the fact that Neo didn't work at acquiring and cultivating those abilities detract from the credit that he gets for taking out those bad guys? Intuitively, it does. But at the same time, it doesn't seem right to say that Neo gets *no* credit, nor that his taking out those agents is in *no* way an achievement on his part. As I said, achievement is a gradable phenomenon. If knowledge is a cognitive achievement, then it too is gradable—one can know more or less well, which seems to me to be a rather fitting consequence of a virtue-theoretic account of

[37] The remark of Sosa's to which I am referring is made in the context of a discussion of his distinction between "animal knowledge" and "reflective knowledge." Reflective knowledge for Sosa is apt belief (i.e. animal knowledge) that one has an apt belief (i.e. an item of animal knowledge)—in other words, reflective knowledge is higher-order animal knowledge (for more on this, see Sosa 2009: Ch. 7). The distinction between high-grade and low-grade knowledge doesn't neatly map onto Sosa's distinction, but they are clearly not wholly unrelated.

knowledge. Even if one arrives at true belief in core propositions through the employment of faculties at which one has not worked at cultivating, one may still achieve knowledge, even if the achievement is less of an achievement compared with cognitive successes that are the result of hard-wrought habits of intellectual character.[38]

§3.2.3 Process reliability

In Chapter 2 I noted that there is a longstanding difficulty concerning how we are to identify in a principled manner the salient process type—of the indefinitely many types that are instantiated by any given belief token—in virtue of which the process type reliability of a belief token is determined. This is known as the generality problem. Sadly, I don't have a proposal for how to resolve the generality problem.[39] Therefore, my strategy here won't be to try to demonstrate that a

[38] As I said, attaining high-grade knowledge very often seems to involve the virtuous handling of a variety of items of low-grade knowledge. Accordingly, one way in which a theist might attain high-grade religious knowledge is by systematizing and thinking through her various items of low-grade religious knowledge (including the deliverances of HADD+) in a manner that exhibits the various virtues of intellectual character discussed by the virtue responsibilists. This would be one way to understand what natural theology is about.

[39] Kelly Becker's recent proposal for solving the generality problem seems to be attracting the most attention at the moment. Becker's contention is that "the relevant process type is the narrowest, *content-neutral* process that is causally operative in belief production" (2008: 363). As Brueckner and Buford (2013) have shown, however, even this proposal fails to get the right results all the time. Becker (2008: 357) mentions a case in which a person's office colleague, Haveit, wants her to mistakenly believe that another of their colleagues, Nogot, owns a Ford, and so hypnotizes

particular process type is the salient one and then try to show that the process type in question is reliable (i.e. has a high truth-ratio). Rather, in the absence of a solution to the generality problem, I shall settle for the somewhat weaker contention that the process type that the religious sceptic will most likely want to claim is the salient one should not, in fact, be deemed unreliable. Now, the religious sceptic who wants to claim that beliefs of core propositions are unreliably formed when they are the product of intuitive cognition will want to show that the salient process type here is such that, either: (i) it yields a high proportion of mutually inconsistent outputs, in which case no assumptions about the truth-values of any particular beliefs are required in order to show that the overall truth-ratio of the type in question is low; or (ii) a significant proportion of the type's outputs are beliefs that, all parties to the debate (in this case, the atheist-theist debate) agree are false beliefs. The best candidate for such a process type, as far as I can see, is the type *trusting the deliverances of intuitive cognition*. A process type of such width is surely instantiated by a lot of false belief tokens, given that all manner of false

her into believing it, whereupon she correctly deduces by existential generalization that *someone* in the office owns a Ford—Haveit is the said owner, and would not have hypnotized the subject had he not owned a Ford. Becker thinks that he can diagnose why this isn't a case of knowledge by appeal to a process type reliability condition according to which the relevant type is picked out in the aforementioned manner. However, as Brueckner and Buford point out, "What is the relevant process-type in the Haveit case? Following Becker's narrow-specification content-neutral (in his sense) lead: *forming a Haveit-induced hypnosis-based car-ownership-in-the-office belief and then reasoning from the belief by using existential generalization*. What is the actual-world truth-ratio of this process-type? If the only hypnosis-based belief that Haveit gives her is that Nogot owns a Ford, then the foregoing method has an actual-world truth-ratio of 1/1. Then the reliable process condition is *satisfied*, contrary to what Becker wants to claim, at least regarding actual-world reliability" (2013: 173-74).

naïve physics and biology beliefs are included.

Firstly, though, it is not in fact obvious that the type *trusting the deliverances of intuitive cognition* has an insufficiently high truth-ratio for knowledge. Truth-ratio is of course a statistical matter, and hence, ascertaining a process type's truth-ratio is a matter of counting up all the instances of the type and then dividing by the number of those instances that yield true belief. In reality there isn't any precise way to do this, and so we just have to use the coarsest of estimates. Proceeding in that vein, then, it should be borne in mind that *trusting the deliverances of intuitive cognition* is also instantiated by all manner of belief tokens whose content concerns the presence of other minds and the states that those minds are in, the properties that various kinds of entity possess, the reality of the external world, the reality of the past, and so on. On the anti-sceptical assumption that many of these belief tokens are true, I submit that it is not clear, at any rate, that the type *trusting the deliverances of intuitive cognition* really does have an insufficiently high truth-ratio.

Secondly, recall that in §2.3.2.2 I argued that a process reliability-based debunking argument against religious beliefs should keep process types indexed to particular propositions, in order to protect a range of scientific beliefs from unintentional debunking. If what I argued there is correct, then even supposing *trusting the deliverances of intuitive cognition* is an unreliable process type, that won't be relevant. Instead, what will be relevant is whether, for some particular proposition *p*, the process type *forming a belief about whether p by trusting the deliverances of intuitive cognition* is reliable. Hence, the poor reliability of intuitive cognition when it comes

to propositions about physics can be screened off from our considerations of the reliability of intuitive cognition regarding religious core propositions. The fact that, say, *forming a belief about whether "objects that fall go straight down" by trusting the deliverances of intuitive cognition* is an unreliable process type will thus have no bearing on the reliability of, say, the type *forming a belief about whether "there is a moral law with a transcendent personal source" by trusting the deliverances of intuitive cognition*. Indeed, if theism is true and hence it is true, for example, that there is a moral law with a transcendent personal source, then the process type *forming a belief about whether "there is a moral law with a transcendent personal source" by trusting the deliverances of intuitive cognition* will be a highly reliable process type, assuming that CSR is correct that maturationally natural cognition results in intuitive belief in that proposition in human beings across all cultures.[40]

Thirdly, even setting aside the previous two points, the following point still stands. If one deems the process type *trusting the deliverances of intuitive cognition* to be the epistemically salient type when it comes to evaluating the reliability of non-inferential beliefs in religious core propositions, then on pain of inconsistency, one should also say that a process type at the equivalent level of generality is the salient type when it comes to evaluating the reliability of other sorts of beliefs. When it comes to beliefs about other minds and the states they are in, beliefs in the reality of the external world and the past, beliefs about what properties various kinds of entity possess, and indeed everyday moral beliefs, insofar as these beliefs are the

[40] See Clark and Winslett (2011).

product of intuitive cognition, one will have to say that the salient process type for evaluating these beliefs is also *trusting the deliverances of intuitive cognition*, and thus, if one claims that that type is unreliable, one will also have to say that none of these beliefs amounts to knowledge either. I take it that such serious and widespread scepticism is a highly unattractive consequence.

So, I contend that short of inviting serious scepticism, we should say that, if true, beliefs of core propositions that are formed in a fully non-inferential manner are the product of a process type that is sufficiently reliable to yield knowledge.

§3.3 Knowing Core Propositions—Way II: Inference from False Non-Core Propositions

Let us now consider what is likely to be the more controversial of my two suggestions as to how core propositions could be known. The thought here is that a person may come to know a true core proposition by way of drawing a deductive inference from a false non-core proposition which she believes as a result of a combination of bio-psychological dispositions and cultural influences.[41] Before going further, I need to clarify the respective causal contributions of the bio-psychological dispositions and the cultural influences in bringing about belief in non-core propositions.

[41] I consider that this claim is still in the spirit of Reformed Epistemology, because the sorts of inferences I have in mind are ones that are obvious, straightforward, and come very naturally to most people.

According to current CSR, the causal story goes roughly as follows.[42] The bio-psychological dispositions outlined in §0.2 dispose a human individual *S* to intuitively believe that there exists a supernatural person (or persons) with the sorts of "core" properties sketched out in §0.2: chiefly, being superknowing, disembodied, very powerful, creating the cosmos and being responsible for designing its apparently goal-oriented features, and being the source of an objective moral law. Unless *S* finds herself in an especially secular culture, religious testimony from her parents, peers, and religious leaders comes into the picture once *S* is old enough to understand what adults are saying, and from it she hears about a supernatural being or beings with many or most of the core properties just mentioned, as well as many other properties besides—for instance, properties of having performed various actions in human history. The fact that cultural testimony tells *S* about a supernatural being or beings with these core properties is itself due to the way that intuitive cognition constrained what *S*'s parents and their ancestors were liable to accept upon reflection and pass onto others. Because this testimony accords with *S*'s intuitions—that is, because the core propositions entailed by cultural testimony are ones that *S* is already disposed to find intuitively plausible—and because *S*'s intuitions don't say much one way or the other about the supernatural being(s)'s possession of various other properties to which her culture testifies, and also because people whom *S* takes to be authoritative are the ones giving her that testimony, *S* is inclined to reflectively believe the propositions being recommended

[42] See Barrett (2011); Boyer (2001).

to her by her culture. Now, one might wonder at this point why someone in this position would *need* to make deductive inferences to core propositions from propositions to which culture testifies, if core propositions are already part of the content of cultural testimony. My thought here is that, typically, when someone receives cultural religious testimony, it doesn't come simply in the existentially generalized form "something is F," where F is (say) the property of being an extremely powerful supernatural person. Rather, cultural religious testimony usually comes in the form "*a* is F," for example, "Vishnu is an almighty supernatural person"; or in the form "something is F and is identical with *a*," for example, "there is an almighty supernatural person who is identical with Shangdi." So, making deductive inferences from these sorts of propositions to core propositions could be a non-redundant way for someone to come to have reflective (rather than merely intuitive) beliefs of true core propositions. The inferences I have mind here are both simple and obvious: for instance, inferring from "*a* is F" that "something is F"; inferring from "something is F and is identical with *a*" that "something is F"; inferring from "*a* = the *x* which is uniquely F" that "there is an *x* which is uniquely F."

My task now is to try to argue that if theism is true, making such inferences from false non-core propositions to true core propositions can sometimes generate knowledge. The worry that immediately confronts this thought concerns whether it is even possible in principle for an inference from a falsehood to result in knowledge. Recall that Edmund Gettier's (1963) counterexamples to the tripartite

account of knowledge involved inference from a falsehood. In one of Gettier's examples, Smith mistakenly but justifiedly believes that "Jones owns a Ford," from which he competently deduces the disjunction "either Jones owns a Ford or Brown is in Barcelona," which is true, but only because the second disjunct happens through sheer luck to be true. In light of Gettier cases, a number of epistemologists have been led to endorse a no-KFF (no knowledge from falsehood) principle, the thought being that any true belief that is based upon a falsehood is going to be subject to the sort of knowledge-preventing accidentality that characterises Gettier's cases.[43] More recently, however, several authors have defended the claim that some cases involving inference to a true proposition from a falsehood do in fact amount to knowledge, and that the cases in question are characterised by certain structural features that render them importantly different from Gettier cases.[44] The way I shall proceed here is by presenting some putative cases of knowledge from falsehood (henceforth, KFF) before outlining a suggestion as to how such cases are structurally different from Gettier cases, and then finally, trying to show that the structural features in question are present in certain cases of inference from false religious propositions to true core propositions.

[43] See Lycan (2006: 156-57, 66); Feldman (2002: 36-37); Harman (1973: 47).

[44] See Hawthorne and Rabinowitz (Forthcoming); Hiller (2013); Turri (2011: 2); Fitelson (2010); Luzzi (2010); Klein (2008); Warfield (2005).

§3.3.1 Knowledge from falsehood vs. Gettierized belief

Here is an alleged case of KFF offered by Ted Warfield:

> *Party*
>
> With hopes of getting him to attend a party in Providence on Saturday night, Jaegwon Kim asks Christopher Hill what he's doing on Saturday. Hill replies "I'm flying to Fayetteville on Saturday night" and the conversation ends. Kim, recalling that Hill taught for many years in Fayetteville, Arkansas, reasons as follows: "Hill will be in Arkansas on Saturday night; so, he won't be in Providence on Saturday night." Kim knows his conclusion, but his premise is false: Hill is flying to Fayetteville, North Carolina. (2005: 407).

The following is a case offered by Peter Klein:

> *Santa Claus*
>
> Mom and Dad tell young Virginia that Santa will put some presents under the tree on Christmas Eve. Believing what her parents told her, she infers that there will be presents under the tree on Christmas morning. She knows that. (2008: 37).

Next, a case presented by Avram Hiller:

> *Rendezvous*
>
> Natasha is a spy in the field. Messages to her from Headquarters often are detected

> by enemy intelligence, and Headquarters is aware of that fact. Today, Headquarters needs to communicate to Natasha that her contact will be at the train station at 4:00 pm, but Headquarters cannot directly tell her so. However, Headquarters knows that Natasha happens to have a justified false belief that the train from Milan is arriving at 4:00 pm. (It really arrives at 8:00 pm; also, there are no signs posted at the station indicating at what time it will arrive, and thus it is very unlikely that Natasha will find out the truth about the train's arrival time.) Headquarters knows that the enemy does not know that she has this false belief. So Headquarters sends a communiqué to Natasha stating that her contact is on the train from Milan. Natasha goes to the station at 4:00 pm and meets her contact there. (2013: 10).

And finally, one of my own:

> *Age of the earth*
>
> Ken watches a documentary about the age of the earth, which asserts that the earth is 5 billion years old. Another documentary that Ken nearly watched asserts that the earth is 4 billion years old. Neither figure is strictly accurate, the true age of the earth being closer to 4.5 billion years. Nonetheless, Ken competently infers from the false proposition that "the earth is 5 billion years old" to the true proposition "the earth is more than 6,000 years old," and he would have deduced the same conclusion had he watched the other documentary. Ken knows his conclusion.

The flat-footed response to these cases is to deny that any of them involve knowledge and hold fast to the thought that one can never get knowledge by way

of inference from a falsehood. Warfield claims that this reaction is contrary to "clear and widely shared intuitions about the cases" (2005: 408). On the other hand, one might respect the intuition that these cases involve knowledge, and try to offer an explanation for how they differ structurally from Gettier-style cases. I am setting aside the first option, aware that if it turns out to be correct then the current proposal for how core propositions can be known fails.

Let's consider the second option. Klein (2008: 47-48) has presented a detailed account of how it is that inference from falsehood can sometimes generate knowledge. According to him, the falsehood that is involved in genuine cases of KFF is indeed causally essential to the production and sustenance of the true inferred belief, which is to say, the false belief is the thing that causes the subject to hold the true inferred belief and there is no causal over-determination such that were the false belief to cease being held the true inferred belief would persist regardless. However, Klein thinks that the falsehood is not *evidentially* essential. He contends that in all cases in which a subject has an item of knowledge that is essentially causally based upon a false belief, the false proposition *f* is always: (i) supported by the subject's evidence, and (ii) is such that it entails a true proposition *t* (also, of course, supported by the subject's evidence), which (iii) evidentially supports the inferred true proposition *h*. Klein claims that this pattern is instantiated in his *Santa Claus* case, since Virginia's false belief that "Santa will put presents under the tree" is supported by her evidence (the evidence of her parents having told her, and of their being reliable truth-tellers), and entails the true

proposition "someone will put presents under the tree," which evidentially supports (indeed entails) Virginia's correct conclusion that "there will be presents under the tree."

The fatal flaw with Klein's characterization of cases of KFF is that it doesn't illuminate the structural difference between cases of KFF and Gettier cases, because some clear Gettier cases also instantiate the three features that Klein describes. Consider the following case:

> *House fire*
>
> Late one night while the inhabitants are sleeping, a fire breaks out in the Smith family home, but is not yet issuing any smoke that can be seen from the street. Jones, passing the house on the street, sees a plume of smoke rising up, which very much appears to be emanating from Smith's house but is in fact emanating from an entirely unrelated bonfire in the next door garden. Jones believes the false proposition that "*that* smoke is emanating from a fire in Smith's house," and infers (inductively) the true proposition that "the inhabitants of Smith's house are in danger."

Given the wholly accidental connection between the apparent indication given by the smoke and the fact that the inhabitants of Smith's house are in danger, it seems that Jones doesn't know his conclusion. Yet, all three of Klein's features are present: (i) Jones's belief of the falsehood *f* (that "*that* smoke is emanating from a fire in Smith's house") is supported by his evidence (the strong appearance of

smoke coming from the house), and (ii) *f* entails a true proposition *t* (that "there is a fire in Smith's house"), which (iii) evidentially supports Jones's true conclusion *h* (that "the inhabitants of Smith's house are in danger"). But I take it to be a chief desideratum of an account of KFF that it be able to illuminate the structural differences between Gettier cases and cases of KFF. After all, a key motivation for adopting a no-KFF principle is that standard Gettier cases involve inferences from falsehoods. The failure of Klein's account to meet this desideratum is thus a serious shortcoming.

§3.3.2 Safely inferring truth from falsehood

Let's consider an alternative characterization of KFF. Warfield writes of his own purported examples of KFF that "the path from falsehood to truth in each example seemed remarkably stable and secure" (2005: 414), though Warfield himself offers no further explication of this thought. Yet it seems to me that a more apt characterisation of KFF must pay close attention to the way that epistemic luck operates in Gettier cases, and by contrast, how the sort of luck at issue in Gettier cases is absent from cases of KFF. To emphasise, the principal motivation for denying that there can be cases of KFF is the thought that the sort of knowledge-undermining luck that characterises standard Gettier cases afflicts *any* case in

which the subject infers a truth from a falsehood.[45] It is worth noting that having the subject draw an inference from a falsehood is not necessary for constructing a Gettier case. Bertrand Russell's (1948: 170) stopped clock case involves no such inference. The subject in this case simply looks at the clock, which correctly reads noon, and non-inferentially believes that the time is noon—unaware that the clock stopped precisely twelve hours ago. A no-KFF principle will be unable to rule out Gettier cases like this one,[46] and yet a truly illuminating diagnosis of Gettier cases should be able to locate some deeper structural feature of *all* such cases, and not only those involving inference from a falsehood. The safety-theoretic diagnosis offered by Pritchard seems to be able to do just that.

According to Pritchard, a Gettier case is one in which "the relationship between the agent's belief and the truth is such that, in a wide range of nearby possible worlds, forming a belief in the way that the agent did in the actual world would lead to false beliefs" (2005: 156). This safety-theoretic characterization is true of both Gettier cases involving inference from falsehood as well as ones involving no such inference. We thus have no need of an independent no-KFF condition on knowledge in order to deal with Gettier cases—a condition that, in any case, fails to exclude non-inferential Gettier cases. What's more, we can see

[45] Warfield (2005: 413) examines various possible motivations for upholding a no-KFF principle, concluding that the only serious motivation is the worry that knowledge-undermining luck, of the sort that characterises Gettier cases, is present whenever a truth is inferred from a falsehood.

[46] More generally, many cases involving so-called environmental epistemic luck, such as fake barn cases, are cases of non-knowledge that involve no inferences, and so cannot be accounted for in terms of a no-KFF principle.

that the feature Pritchard identifies as characterizing Gettier cases is *not* present in the putative cases of KFF that I presented earlier. In Warfield's *Party* case, there is no nearby world in which Kim forms a false belief about whether Hill will be in Providence on Saturday night by way of the same or a similar causal mechanism—namely, asking Hill and making inferences from Hill's reply in conjunction with a false assumption about which town called "Fayetteville" Hill is talking about. In Klein's *Santa Claus* case, in nearby worlds in which there will not be any presents under the tree on Christmas morning, Virginia doesn't believe that there will be any, since her parents don't tell her that her that anybody (Santa or otherwise) will be bringing them. In Hiller's *Rendezvous* case, in nearby worlds in which Natasha's contact is arriving at the station at a time other than 4:00 pm, presumably Natasha has a correspondingly different, true belief about what time her contact is arriving into the station. After all, HQ knows full well what time her contact will arrive into the station, and would find some other way to inform Natasha of that information if her contact were arriving at a time other than 4:00 pm—perhaps by making use of other false beliefs Natasha holds about what times various trains arrive into the station, or perhaps by some other means. Finally, in my *Age of the earth* case, in nearby worlds in which Ken watched other documentaries asserting claims about the age of the earth that are at least approximately true although strictly false, Ken still competently deduces the truth that the earth is more than 6,000 years old. In short, despite the fact that the subjects in these cases all make use of a false belief in drawing their true conclusions, it looks as

though their arrival at the truth is very safe.

Now, besides the fact that the belief in these cases is safe, there doesn't seem to be one single pattern that is common to all of them in virtue of which the belief is safe. But consider one sort of pattern, which is exemplified by the *Age of the earth* case. The subject makes a competent deductive inference from a false proposition to a less fine-grained, true proposition: for instance, from "*a* is F" to "something is F." In all the nearby worlds in which the subject comes to believe a different fine-grained proposition by way of a similar causal mechanism—for instance, the proposition that "*b* is F"—that fine-grained proposition, even if false, is such that the subject is still able to deduce from it a less fine-grained proposition that is true, namely, the proposition that "something is F."

What I want to suggest is that the following sort of case exhibits this pattern:

Olympus

As a result of the CSR-described cognitive dispositions, Achaius is already intuitively disposed to believe in the existence of a supernatural creator of the universe. Achaius's culture testifies to him that "Zeus is the supernatural creator of the cosmos," and finding that the idea of a supernatural creator accords with his intuitions and that his intuitions say nothing about the notion that the supernatural creator is identical with Zeus, Achaius accepts this testimony. Noticing the obvious entailment from "Zeus is the supernatural creator of the cosmos" to "there is a supernatural creator of the cosmos," Achaius comes to reflectively (and not merely intuitively) believe the latter.

Now, why should we think that, if true, Achaius's reflective belief in this core proposition is safe? Well, that belief is safe just in case there is no nearby world in which Achaius comes to have a false belief about the proposition "there is a supernatural creator of the cosmos" by way of a causal history that is sufficiently similar (for the purposes of evaluating safety) to the actual causal history of his belief. Let's grant for the sake of the argument that the class of relevantly nearby worlds includes worlds in which Achaius is located in a substantially different religious-cultural setting. So, the question is: are there nearby such worlds in which Achaius ends up (*ex hypothesi*) falsely believing that "there is *no* supernatural creator of the universe" by way of a mechanism that is relevantly similar to the one he employs in the actual world?[47] Well, in all the nearby worlds in which Achaius is subject to religious teachings that testify about some supernatural creator (or creators) and in which Achaius performs the same sort of inference as he performed in the actual world (i.e. a deductive inference from a non-core proposition about the supernatural creator(s) to an existentially generalized core proposition), he will end up believing that "there is a supernatural creator of the cosmos." And if God exists in the actual world, then given the modal stability of

[47] Of course, it is relevant to the safety of someone's belief that *p* whether or not there are nearby worlds in which *p* is false and yet *S* continues to believe *p*. As I noted in §3.2.1, however, core propositions are, if theism is true, highly modally stable if not outright metaphysically necessary, meaning that the nearest worlds in which such propositions are false are extremely modally distant from the actual world. And what goes on in very far-off worlds is not relevant for safety, so I have ignored here the issue of whether *S* continues to believe that "something is F" in the very far-off worlds in which that proposition is false (assuming there are any such worlds at all).

His existence He also exists in all these nearby worlds, and so, Achaius's belief in these nearby worlds that "there is a supernatural creator of the cosmos" will also be true. Of course, there may also be nearby worlds in which Achaius is subject to irreligious cultural influences that teach him to ignore and override his intuitive religious cognition, so that he ends up believing that "there is no supernatural creator of the cosmos." But I want to suggest that the etiology of his belief in such worlds will be insufficiently similar to that of his actual-world belief to be relevant to the safety of his actual-world belief. Allow me to say something in favour of this latter contention.

The bio-psychological dispositions described by CSR, which are part of the causal history of Achaius's actual-world belief that "there is a supernatural creator of the universe," are presumably not part of the causal history of his belief that "there is no supernatural creator of the cosmos" in the world in which he is raised in a secular cultural setting. That is not to deny that Achaius possesses those dispositions in that world, just that they form no part of the causal history of his belief in the atheistic propositions that he believes in that world. This difference in belief etiology seems to me to be sufficiently great so as to render Achaius's (*ex hypothesi*) mistaken belief that he holds in that world irrelevant to the safety of his actual-world belief that "there is a supernatural creator of the cosmos." In support of this, consider an analogy. Suppose that in actuality Sally believes that "there are other minds besides my own," and that she believes this in part on the basis of the deliverances of intuitive cognition (the theory of mind). But suppose that there is a

very close world (call it w^*) in which Sally was adopted by the lady who famously commended solipsism to Bertrand Russell, and who would have offered Sally arguments for solipsism whilst teaching Sally to distrust the deliverances of intuitive cognition so far as other minds are concerned, so that Sally would have believed that "there are no other minds besides my own." Now, I am not remotely inclined to think that Sally's actual-world knowledge that "there are other minds" is threatened by her having a mistaken belief on the topic in w^*, despite the closeness of w^* to the actual world. And the most promising explanation for why the false belief she holds in w^* is irrelevant to the safety of her actual belief is that the causal history of Sally's belief does not remain sufficiently similar across the two worlds. In particular, the causal history of her belief in the actual world includes the deliverances of intuitive cognition, whereas the causal history of her belief in w^* excludes those deliverances.

§3.3.3 Successful inference from falsehood can be a cognitive achievement

Can a belief in a true core proposition that has been inferred from a false non-core proposition satisfy a cognitive achievement condition on knowledge? I contend that it can.

One might worry, however, that it is a general feature of cases in which a true proposition is inferred from a falsehood that such cases fail to be cognitive achievements. The worry, in short, might be that whenever a false premise is used

in reasoning to a true conclusion, the subject's cognitive success in arriving at that true conclusion cannot count as a cognitive *achievement*, due to the way in which a cognitive failure (i.e., a false belief) was part of the means to that cognitive success. A cognitive failure is not something for which the subject can take credit, and so, the thought might go, the subject's arrival at a true conclusion owes its success in significant part to something for which she can take no credit, and so her success falls short of being an achievement.[48] But I now want to try to show that a cognitive success that relies in part on a prior cognitive failure *can* sometimes be a cognitive achievement.

Recall that the central analogy in Sosa's (2007, 2011) account of knowledge as a cognitive achievement is that of an archer who successfully propels an arrow through his target in a way that manifests his archery ability. The accuracy of the archer's shot is *because of* his archery ability; it is creditable to the archer's ability. His shot is thus *apt*, and an apt performance is equivalent to an achievement. At the risk of slightly straining Sosa's analogy, I'd like to adapt it so as to try to capture the relevant features of the alleged cases of KFF that I presented in §3.3.1. A key feature we want to capture is the way in which the accuracy of the subject's inferential belief is partly dependent upon the accuracy of the belief(s) from which it is inferred.

[48] Indeed, Sosa's remarks about means-ends action—of which inference might be seen as a species—could be read as affirming an objection in this vein. He writes that: "Suppose the means-end belief [i.e., the belief upon which a means-ends action is based] is epistemically competent but not apt... In that case, I submit, the means-end action itself fails to be apt. It falls short in this performance-normative way" (2011: 45).

So, suppose that our archer has been set a challenge in which the accuracy of his second shot will rely in part upon the accuracy of his first shot: specifically, he is to try to land his first arrow in the innermost of three rings that are being projected onto a large wall, but he must shut his eyes as soon as he has released his arrow so as to avoid seeing where it lands. The target rings will then cease being projected onto the wall before the archer is allowed to look again, and so for his second shot, he will have only the placement of his first shot to guide his aim. Now, suppose that the archer's first shot is fairly competent but fails to hit its intended target, the innermost ring, instead landing inside the outermost ring. That first shot is not accurate, but it is not very far from being accurate, and the fact that it is not far from being accurate is due to the archer's competence. (This is supposed to capture the way that the false premise—from which the subject in my alleged religious case of KFF reasons to a true conclusion—is inaccurate, and yet it is still to the credit of the believer's cognitive faculties—specifically, his intuitive religious cognition—that the false non-core beliefs that he is inclined to accept from culture are at least ones that *entail* true core propositions.) Suppose that the archer then competently uses the placement of the first arrow in order to guide his second arrow to a location right beside the first. And suppose further that for the second shot to count as accurate, it need only land within the outermost ring. (This is supposed to capture the way that, in my alleged religious case of KFF, the true core proposition is a less fine-grained proposition than the non-core proposition from which it is inferred, and hence, the inferred core proposi-

tion can be true in more ways than can the non-core proposition from which it is inferred.) Now, I contend that the archer's success in getting his second arrow within the specified target area is indeed an achievement. Perhaps it is *less* of an achievement than it would have been had his first shot—upon which the competence of his second shot partly relied—been fully accurate, but it is still *something* of an achievement, and achievement is intuitively a gradable phenomenon.

What this adaptation of Sosa's analogy is meant to capture are the following crucial features of my alleged religious case of KFF: (i) the second belief (the inferred belief in a core proposition) is partly dependent for its accuracy upon the accuracy of the first belief (the belief in a non-core proposition); (ii) the first belief is inaccurate though in some sense is an approximation of the truth, or at any rate, the fact that the false non-core proposition that the believer has accepted from her culture is one that entails a true core proposition is a fact that is in some measure creditable to the believer's cognitive dispositions—namely, those intuitive dispositions described by CSR; (iii) the accuracy of the second belief is due both to the subject's deductive competence (in the analogy, the archer's competence in placing his second shot with reference to the placement of the first) and to the approximate accuracy of the first belief, or at any rate, to the fact that the first belief contains an important element of the truth, which fact is no accident but is due the subject's cognitive abilities. If the analogy is a good one, and the archer's success in getting the second shot within the target area is indeed an achievement as I contend, then we should say that the religious believer who infers a true core

proposition from a false non-core proposition in the manner outlined attains a cognitive achievement.

§3.3.4 Satisfying the process reliability condition

As I said in §3.2.3, I don't have a proposal for how to resolve the generality problem, and so my strategy here won't be to try to demonstrate that a particular process type is the salient one and then try to show that the process type in question is reliable (i.e. has a high truth-ratio). Again, I shall settle for the somewhat weaker contention that the process type that the religious sceptic will most likely want to claim is the salient one should not, in fact, be deemed unreliable. The religious sceptic who wants to claim that beliefs of core propositions are unreliably formed when arrived at by way of inferences from falsehoods will want to show that the *salient* process type here is such that, either: (i) it yields a high proportion of mutually inconsistent outputs, in which case no assumptions about the truth-values of any particular beliefs are required in order to show that the overall truth-ratio of the type in question is low; or (ii) a significant proportion of the type's outputs are beliefs that, all parties to the debate (in this case, the atheist-theist debate) agree are false beliefs. So, when it comes to core propositions inferred from false non-core propositions, which process type, of the indefinitely many types at various levels of generality that could be picked out, is going to give the religious sceptic the result she wants—i.e., is going to be such that either (i) or (ii) is

true of that type? By far the most promising candidate, I would think, is the process type *drawing inferences from falsehoods*. The religious sceptic will presumably want to claim that (ii) is true of this type.

Now, even supposing there is a good reason for singling out *this* type rather than another of the indefinitely many as the epistemically salient one, there seem to me to be serious problems. For one thing, is it really all that obvious that the type *drawing inferences from falsehoods* has an insufficiently high truth-ratio to yield knowledge? After all, a great many instances of that type are cases in which a person reasons from a premise that is strictly speaking false but which adequately approximates the truth for her purposes, and hence securely leads her to a true conclusion. As John Turri (2011: 8) notes, cases like this include a great deal of instances of reasoning from empirically adequate though false scientific theories, and also, a great many cases of everyday reasoning from so-called "naïve" physics, much of which is strictly speaking false (by the lights of our best physical theories, at any rate), but which very frequently permits us to draw correct inferences.[49]

But secondly, even setting aside the worry just raised, in saying that *drawing inferences from falsehoods* is an unreliable process type and that *drawing inferences from falsehoods* is the epistemically salient type when it comes to evaluating the reliability of core beliefs formed in the manner under consideration, one will be committed to saying the very same thing about all manner of beliefs of the sort just mentioned—ones that are competently inferred from premises that are strictly

[49] For more on naïve physics, see see De Cruz and De Smedt (2007: 353).

speaking false yet that perfectly adequately approximate the truth for the purposes at hand. The religious sceptic will thus be committed to saying that cases of this sort can never amount to knowledge. But that seems unduly sceptical. As Turri points out,

> [F]alsehood in the form of idealization pervades scientific reasoning, much of which is competent and confers knowledge… And for some purposes it doesn't matter if we believe that the gravitational constant is exactly, as opposed to approximately, 6.7 x 10 − 11m^2/ks^2 or that *pi* equals exactly 3.14. We might nevertheless reason from these false premises to reach a true conclusion, which outcome would constitute knowledge. (2011: 8).

Picking out a slightly narrower process type as the salient one, say, *drawing inferences from premises that are false but that approximate the truth or contain an important element of the truth*, will not have the aforementioned unwarranted sceptical consequences. But nor will it enable the religious sceptic to secure the sceptical conclusion against cases in which true belief in a core proposition is inferred from a false non-core proposition, provided the false non-core proposition does indeed contain an important element of the truth.

So, short of inviting unwarranted sceptical consequences, I contend that we should deem true beliefs of core propositions, arrived at in the manner under consideration in this section, as satisfying the process reliability condition.

§3.4 Conclusion

In summary, I have suggested that there are two ways in which core propositions can be known if God exists, one wholly non-inferential and the other involving very simple and obvious deductive inferences. I have tried to show that if God exists, beliefs formed in either of these two ways are likely to satisfy the safety condition, a virtue-theoretic cognitive achievement condition, and a process reliability condition. The significance of all this will become clear in the next chapter, where I shall argue that knowledge of core propositions is a sufficient basis for relating properly to God—a claim that has considerable import for the ongoing debate about divine hiddenness.

Chapter Four

A Defence against the Demographic Divine Hiddenness Challenge

> It is because some vestige of the capacity to be aware of God is present in all or most people that there is such a thing as religion… An awareness of God does not necessarily seem to be an awareness of God to the person who has it: an awareness of a distant mountain range may seem to the person who had it to be an awareness of a bank of clouds.
>
> Peter van Inwagen

§4.0 Introduction

I have argued that even supposing that religious beliefs that are partly the product of contingent cultural influences fail to amount to knowledge (even if true), there is nevertheless a body of core religious content that can be known if theism is true, and which knowledge is widely available to people across various cultures. The picture I have painted raises some troubling questions, though. Why would a perfectly loving God be content to permit a situation in which the ebbs and flows of cultural tides play such a significant role in determining the specific contours of an individual's religious beliefs? Why, indeed, would God create a world that is

religiously ambiguous to the extent that different human cultures can end up with such diverse ideas about the nature of divine reality? Surely, if there were a God of the sort depicted by the major monotheistic traditions, He would want to make things considerably clearer to us; He wouldn't permit entire regions of the world to persist in significant error about His nature and attributes. These are the questions, or challenges, that this final chapter will take up. My aim here will be to present a kind of defence for theism in the face of these worries. That is, I will try to make a case for thinking that despite the initial appearances to the contrary it shouldn't be surprising, given God's existence, that we observe the sort of world we do: a world in which human beliefs about many of life's most important matters, including religious matters, are influenced to a significant degree by the norms and consensuses of the communities in which we find ourselves. To begin with, I shall outline a potent version of the challenge, which I term *the demographic divine hiddenness challenge*. I explore its contours and try to formulate it as precisely as possible before going onto present my defence for the remainder of the chapter.

§4.1 The Demographic Divine Hiddenness Challenge

§4.1.1 Schellenberg's divine hiddenness argument

J. L. Schellenberg (1993; 2007: Chs. 9, 10) is the foremost defender of the atheistic argument from divine hiddenness. His argument runs roughly as follows. A

perfectly loving God would want a personal relationship with each one of His human creatures who is capable of such a relationship and is not actively resisting such a relationship, and so God would at all times supply those who are willing with whatever is necessary for them to be in a position to enter into a relationship with God just by trying. One thing that is necessary for being in a position to enter into a personal relationship with God is having the belief that God exists. And yet, there are in fact individuals—and many of them—who, through no fault of their own, fail to believe that God exists, and thus, fail to be in a position to enter into a personal relationship with God. Schellenberg terms such individuals "non-resistant non-believers" (2007: 195-218). According to Schellenberg, the existence of non-resistant non-believers is inconsistent with the existence of a loving God. His most up-to-date formulation of this argument goes 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). [premise].

(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. [premise].

(3) 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. [(1) & (2)].

(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. [premise].

(5) God does not exist. [(3) & (4)]. (2007: 204-06)

Various theistic philosophers have responded to this argument. The primary target for such responses is premise (1).[1] Responses to (1) have sought to show that it is possible, in the broadly logical sense, that God justifiably permits there to be individuals who are not culpably resisting God and yet who fail at some times in their lives to be in a position to enter into relationship with God just by trying. The way such responses typically proceed is by identifying some putatively outweighing good (which may be the prevention of a worse harm) for the sake of which God could justifiably, for a time at least, permit some human being to be in state of non-resistant non-belief. Examples of the putatively outweighing goods that have been put forward include the exercise of morally significant free will by humans, which, so authors such as Michael Murray (1993) and Robert McKim (2001) have suggested, would be curbed in some individuals were God to reveal Himself to a degree sufficient for all non-resistant non-belief to be eliminated. Another putative outweighing good suggested by Daniel Howard-Snyder (1996) is that of our relating to God out of wholesome motivations; on the basis of certain non-propositional attitudes towards God, such as humility and gratitude. In a similar vein, Paul Moser has suggested that "God cares mainly about what and

[1] As we shall see shortly, my own response is in effect to challenge premise (2), or at any rate, an interpretation of premise (2) according to which personal relationship with God requires having *de dicto* beliefs about God rather than merely *de re* beliefs of God.

how we *love*, not just what we believe... So, contrary to a typical philosophical attitude, knowledge of God is not a spectator sport" (2004: 46, 48). According to Moser, God would have good reason to fail to bring about theistic belief in a person if that belief would not likely be accompanied by the sorts of attitudes to which Moser alludes. And Andrew Cullison (2010) has proposed that radical self-sacrifice is a great good, and—drawing on a suggestion by Erik Wielenberg (2005: 91-96) to the effect that radical self-sacrifice is possible only if one fails to believe in the existence of a God who will reward such sacrifice in an afterlife—Cullison proposes that God would have to permit instances of non-resistant non-belief in order to make such radical self-sacrifice possible.

§4.1.2 Maitzen's demographic divine hiddenness argument

Stephen Maitzen (2006) has offered a novel version of the hiddenness argument, which he claims can circumvent all of the aforementioned sorts of responses to Schellenberg's argument. Maitzen's version of the hiddenness challenge places the emphasis squarely upon the apparently highly uneven manner in which monotheistic belief seems to be geographically distributed across the globe—something he claims is highly improbable given the existence of a perfectly loving God, even granting that God would have to permit *some* instances of non-resistant non-belief

in order to bring about the various goods just mentioned.[2] Maitzen writes that,

> Because it abstracts from individual cases and considers the large-scale distribution of non-belief, [this version] need not ask "whether any particular candidate for inculpable nonbelief possesses or fails to possess those motivations, attitudes, and dispositions that putatively explain their inculpable nonbelief." Instead, the demographic objection assumes, plausibly, that the observed large-scale patterns iron out individual differences with regard to any of the "motivations, attitudes, and dispositions" invoked by [various responses to Schellenberg's version of the hiddenness argument], such as stupidity, insensibility, presumptuousness, susceptibility to fear of punishment, lack of contrition or humility, and even one's predestination as non-elect. The objection assumes that individuals with these characteristics do not cluster by country or culture so as to show up twenty times more often in Thailand than in Saudi Arabia. These dispositions thus resemble other fundamental human characteristics, such as the ability to hear; despite marked differences among individuals, we don't find entire countries whose citizens are nearly all deaf. (2006: 184).

I take it that this way of putting the divine hiddenness challenge is best construed as a probabilistic rather than deductive argument against the existence of God.

[2] In one of the earliest papers responding to Schellenberg's argument, Daniel Howard-Snyder acknowledges that "It is an interesting question why God would reveal himself clearly to the members of some races but not at all to others. Perhaps another argument from divine hiddenness could be constructed which appeals, not to individual nonbelief but to what we might call *cultural* nonbelief" (1996: 452). Maitzen appears to have constructed just such an argument.

Maitzen doesn't seem to be claiming that the uneven geographical distribution of belief in God is logically inconsistent with theism, but rather, that such a distribution is not all what we would expect to see given theism.[3] That is, it is not at all what we would expect to see if theism were true, given that a loving God would not be at all likely to allow people in entire geographical regions of the world—and we might add, whole eras during human history—to forego the opportunity to have an accurate picture of ultimate reality.

Probabilistic arguments from evil, which provide an apt analogue for probabilistic arguments from hiddenness, have been formulated in two basic ways. William Rowe's (1979) argument from apparently pointless (or gratuitous) evils instantiates the following argument form:

(P1) It appears that the state of affairs *A* obtains

(P2) If it appears that *A* obtains, then probably *A* obtains

(C1) Probably, *A* obtains

(P3) Necessarily, if God exists then *A* does not obtain

(C2) Probably, God does not exist

It doesn't look like this is the right way to formulate Maitzen's argument, though,

[3] Maitzen writes in a subsequent paper, in which he offers some clarifications on the earlier one, that "I've never alleged the logical incompatibility of theism and the lopsided distribution of theistic belief we observe in the world… All I claim is that the lopsided distribution of theistic belief is less surprising on naturalistic explanations than on theistic ones" (2008: 475, 77).

since Maitzen doesn't treat the state of affairs at issue (i.e. the state of affairs in which theistic belief and non-belief are geographically distributed as they actually appear to be) as one whose actuality needs to be argued for. Instead, it looks as though Maitzen's argument is apt to be formulated along the same lines as Paul Draper's (1989) probabilistic argument against theism from the distribution of pleasure and pain.[4] Rather than proceeding by way of a deduction from a set of premises one of which involves a probabilistic claim, Draper instead seeks to establish the following comparative likelihoods claim (where "O" stands for our observations about the distribution of pleasure and pain and "HI" for the hypothesis of indifference, which is effectively the atheistic hypothesis):

(C) Pr(O/HI) is much greater than Pr(O/theism)

Put simply, Draper's claim is that the observed distribution of pleasure and pain "is much less surprising" on HI than it is on theism (1989: 333). The significance of (C), according to Draper, is this:

> Since the denial of theism is obviously entailed by HI and so is at least as probable as HI, the truth of (C) is a prima facie good reason to believe that theism is less probable than not. And since it is epistemically irrational to believe both that theism is true and that it is less probable than not, the truth of (C) is also a prima facie

[4] For more on the differences between these two ways of formulating a probabilistic atheological argument, see Murray (2008: 31-33).

> good reason to reject (i.e., to cease or refrain from believing) theism. (1989: 333).

A probabilistic argument from demographic divine hiddenness of the sort Maitzen offers is best formulated in a manner that straightforwardly parallels Draper's argument. For the purposes of such an argument, we can compare the theistic hypothesis with its negation, namely, atheism (with which Draper's hypothesis of indifference is basically equivalent). And we can let "E" stand for a proposition describing our observations about the geographical distribution of theistic belief and non-belief, not only at the present time, but across all times during human history. Now, as we've seen already, current CSR suggests that some sort of supernaturalistic belief is ubiquitous throughout the world. So it seems to me most charitable to understand Maitzen as being concerned with the uneven distribution of specifically *monotheistic* belief,[5] which does indeed appear to be rather patchy, both across history and across the globe at any given point in history. From henceforth, monotheism is what I shall have in mind when talking about "theistic belief" and "theism." Accordingly, the crucial comparative likelihoods claim is this:

(C*) Pr(E/atheism) is much greater than Pr(E/theism)

It is this claim that will preoccupy me for the remainder of this chapter.

[5] I understand monotheistic belief to be the belief that there exists a single personal deity, who is morally perfect, omnipotent, and omniscient.

§4.1.3 The defence sketched

When an argument against theism is cast in probabilistic rather than deductive terms, the appropriate way for the theist to respond is also in probabilistic terms. Alvin Plantinga's (1974a; 1974b: Ch. 9) free will defence was aimed at disarming the claim that it is logically impossible for God to co-exist with evil, and so, appropriately enough, Plantinga sought to show that it is in fact possible in the broadly logical sense that the two co-exist. But this sort of response won't be sufficient to undermine a claim like (C*). It is completely consistent with (C*) that there is some logically possible world in which God exists and permits a distribution of theistic belief akin to what we observe in our world. Peter van Inwagen suggests that when the theist is confronted with a probabilistic challenge like (C*), she has basically three options:

> 1. The theist may argue that S [the nefarious phenomenon] is much more surprising, given HI [the atheistic hypothesis], than one might suppose.
> 2. The theist may argue that S is much less surprising, given theism, than one might suppose.
> 3. The theist may argue that there are reasons for preferring theism to HI that outweigh the *prima facie* reason for preferring HI to theism that we have provided. (1991: 138).

I am going to take the second option. To take this option is to provide a defence.

The nefarious phenomenon that is much less surprising on theism than initially appears, so I will argue, is E: the uneven distribution of theistic belief.

Notably, the defence concerning E that I am going to offer differs quite markedly from one that has been presented by certain Molinist philosophers of religion such as William Lane Craig (1989) and Jason Marsh (2008). According to Craig and Marsh, God (if He exists) providentially settles the locations and times at which each human being lives, and does so with reference to His exhaustive "middle knowledge" of counterfactual truths about how each individual *would have* responded to Christian revelation had he or she been presented with it. According to this Molinist defence, there are some—indeed, many—individuals of whom it is true that if God were to have brought him or her into existence at a time and place in which Christian theism is a culturally live option, he or she would freely have rejected it. God, knowing this, groups together these counterfactually stubborn individuals[6] in regions of the world that Christian revelation does not penetrate. Now, I won't attempt to critique Craig's and Marsh's defence, as that has been done fairly extensively elsewhere,[7] but I do want to emphasise that my

[6] Of such individuals, Craig writes that "on the analogy of transworld depravity [that is, the property, allegedly possessed by human persons, of going wrong with respect to at least one moral choice in every possible world in which God could create that person], we may accordingly speak of the property of *transworld damnation*, which is possessed by any person who freely does not respond to God's grace and so is lost in any world feasible for God in which that person exists" (1989: 50).

[7] Maitzen (2008) has offered a detailed critique of this Molinist defence. A key objection of Maitzen's is that even supposing that some of the possible individuals God could create are stubborn, such that they would respond unfavourably to God in any world in which He creates

own defence will *not* depend upon invoking the sorts of brute, ungrounded counterfactual truths that play so central a role in their story.

The defence I shall present has somewhat more in common with Tim Mawson's (2012) defence, though I think my own defence can also be seen to be distinct from Mawson's. According to Mawson, God wants people to freely choose whether or not they come to believe in Him, and whether people freely choose to accept God is something that is affected by whether those around them also do so, and so, given that God wants people to freely come to believe in Him, one would expect those who do so to be somewhat "clustered" around the surface of the globe. Mawson contends that it is plausible that God's desire for people to come to believe in Him freely—which brings with it the likelihood of geographical unevenness of distribution—is not overbalanced by His desire for people to come to believe in Him by any means necessary. The only reason for thinking it would be thus overbalanced, Mawson holds, is the thought that God would want to ensure that everyone has the opportunity to accept or reject Him within this lifetime, but Mawson argues that this thought is mistaken and is not taught by the Bible.

Here, then, is a sketch of my own defence:

(i) To begin with, in §4.2, I will argue that E is highly probable given an actu-

---

them, there is no reason at all for thinking that God finds Himself stuck with a particular set of human individuals each of whom He must create. Neither Marsh nor Craig claim that all or even most possible individuals God could create are stubborn in this way, and so it is wholly unclear why God would choose to actualize millions or even billions of such individuals.

ally obtaining state of affairs, S, in which humans are cognitively constituted in such a way that we are forced to rely upon testimony for much of what we know about the world and in such a way, moreover, that we are liable to be significantly influenced by what those around us believe, particularly when it comes to matters that aren't readily susceptible to empirical investigation, including religious matters.

(ii) I shall then argue that we should not be surprised were God to bring about S, even though it has E as a highly probable consequence. My case for this claim comes in two parts:

(iii) Firstly, in §4.3, I shall try to undermine a potentially powerful reason for thinking that God would be unlikely to bring about S given that it has E as a highly probable consequence: namely, that an uneven distribution of theistic belief means that significant swathes of world's population are cut off from the possibility of entering into personal relationship with God. It is at this point that the claims for which I argued in Chapter 3 become highly relevant. There I argued that if theism is true then knowledge of a body of core religious content is available even to those who have not been exposed to specifically monotheistic cultural influences. What I shall argue here is that knowing core propositions, as I called them in Chapter 3, entails having a cognitive grasp of God that is sufficient for relating personally to God. Hence, even despite an uneven geographic distribution of theistic belief, personal relationship with God—even if only *de re* person-

al relationship—is very widely available.

(iv) Secondly, in §4.4, I shall argue that creatures with our cognitive constitution—our combination of cognitive powers, dispositions, and limitations—are among a range of creatures whom it would be good for God to create. I shall suggest that there plausibly isn't a uniquely best cognitive constitution for God to bestow upon a race of intelligent, morally sensitive, free creatures. Plausibly there are other possible such creatures with somewhat different cognitive constitutions than us, whom God would have had equally good reason to create. But I shall argue that our cognitive constitution clearly lies between two undesirable extremes, and that as such, it would not be surprising were God to create beings like us; it would not be surprising were God to bring about S. Furthermore, I try to dispel the worry that if God did bring about S, He would act to prevent S from leading to E, even if S would otherwise very like lead to E.

All in all, then, my aim is to show that when all things are considered, it shouldn't in fact be surprising were God to create a world in which E obtains; a world in which theistic belief is unevenly geographically distributed.[8]

[8] Now, it might be that E is fairly expectable on atheism—that isn't something I'm going to contest, at any rate. If E is indeed expectable on atheism, then even supposing it can be successfully shown that E is unsurprising on theism, there may remain something of an inequality between Pr(E/atheism) and Pr(E/theism). I would make two points about this. Firstly, it seems to me that even the most promising theodicies with respect to the problem of evil don't achieve more than this (*mutatis mutandis*). Currently prominent theodicies, such as those presented by Adams and Adams

§4.2 The Human Cognitive Constitution and the Uneven Distribution of Theistic Beliefs

I take it to be a contingent fact that humans are cognitively constituted in such a way that we have no choice but to rely upon testimony for much of what we know about the world and in such a way, moreover, that we are liable to be significantly influenced by what those around us believe, particularly when it comes to matters that aren't readily susceptible to empirical investigation, including religious matters. It is this state of affairs that I am labelling "S," and which I shall try to characterize in more detail now. On the matter of our cognitive constitution, the early-modern Scottish philosopher Thomas Reid is well worth quoting. Of all the philosophers of the early-modern period, a number of whom actively sought to

---

(1990), Swinburne (1998), van Inwagen (2006), and Stump (2010), all seek to show that God would have good reasons for permitting the evils that we observe in our world. But from this it doesn't follow that a world like ours is *highly* expectable on theism—after all, none of these theodicists try to show that God would *have* to create a world, let alone a world like ours—and nor does anything follow from it regarding the expectability of a world like ours given atheism. Rather, these theodicists seek to show just that our world, with all its evils, is in fact among the good worlds that God could create, and a parallel claim is what I will be seeking to establish: namely, that our world, with its uneven distribution of theistic belief, is among the good worlds that God could create. Secondly, showing that this is so is far from a trivial or insignificant result. If the situation initially appears to be that "E is not at all what we'd expect if there's a God, but E is pretty much what we'd expect if there are just blind, brute physical forces and chance," then the theist will have made important headway if she can move a fair-minded audience to alter that initial judgment to the following: "Though it still seems that E is pretty much what we'd expect if there are just blind, brute natural forces and chance, it actually looks as though E *wouldn't* be surprising if there's a God."

engender a deep suspicion about testimony as a source of knowledge, it was arguably Reid who exhibited by far the keenest appreciation of our mutual epistemic dependence. He wrote that,

> The wise and beneficent Author of Nature, who intended that we should be social creatures, and that we should receive the greatest and most important part of our knowledge by the information of others, hath, for these purposes, implanted in our natures two principles that tally with each other… The first of these principles is a propensity to speak and to use the signs of language so as to convey our real sentiments… Another original principle implanted in us by the Supreme Being, is a disposition to confide in the veracity of others, and to believe what they tell us. This is the counterpart to the former; and as that may be called the *principle of veracity*, we shall, for want of a more proper name, call this the *principle of credulity*. (1997 [1764]: 194).

Reid here identifies two principles, or really, *tendencies*, that he recognizes as being native to human beings: the tendency to *tell* (what we take to be) truth, and the tendency to *trust* what others tell us. Reid goes onto point out that, given the constraints upon what each one of us can come to know just through our own senses and faculties of reasoning, we would be in an epistemically forlorn situation indeed were it not for the fact that we possess the two tendencies just mentioned:

> It is evident that, in the matter of testimony, the balance of human judgment is by

> nature inclined to the side of belief; and turns to that side of itself when there is nothing put into the opposite scale. If it was not so, no proposition that is uttered in discourse would be believed, until it was examined and tried by reason; and most men would be unable to find reasons for believing the thousandth part of what is told them. Such distrust and incredulity would…place us in a worse condition than that of savages. (1997 [1764]: 194).

As Reid astutely notes, our pursuit of epistemic goods is a deeply communal enterprise. With the exception of knowledge about our own mental states, immediate surroundings, memories of our past experiences, and a handful of self-evident truths, humans are cognitively constituted in such a way as to be significantly dependent for our acquisition of a wide range of truths upon a massive division of epistemic labour amongst ourselves. The task of gathering evidence for vastly many of the propositions we believe is a thoroughly cooperative affair. Perhaps the paradigm examples of cooperative evidence gathering occur in the sciences. The task of testing and confirming a large-scale theoretical framework such as quantum mechanics, Big Bang cosmology, or evolutionary theory in biology, is such that it is a practical necessity that the task be massively distributed across vast communities of scientists. Indeed, it is quite plausible that some—perhaps very many—scientific hypotheses are such that no individual scientist ever has in his or her possession all of the evidence for that hypothesis. Rather, any given individual scientist must trust the testimony of her colleagues with respect to those portions of the evidence that are beyond her capability to assess

properly.[9] As for the laypeople—most of us, we are entirely dependent upon the testimony of experts in our acquisition of scientific knowledge. And I take it that it isn't by any means only in the scientific realm where this division of labour is a practical necessity, given our actual cognitive powers. Many of us have a very considerable stock of knowledge about what went on in the distant past. It goes without saying that the transmission of recorded history is a hugely cooperative activity, involving oral testimony from one generation to the next, the gathering, preservation, interpretation, and publicization of all manner of documents and physical evidence by experts from across the world, and now, of course, the creation and transmission of vast quantities of digital media. And most of us know a good deal about present occurrences of which we have no first hand experiences. The acquisition of such knowledge is, again, the result of mutual sharing of first-hand experience via testimony chains—chains that are in some cases extremely long and complex.[10] What's more, most of us know a fair amount about the inner lives of other people with whom we are acquainted. We know largely (though not entirely) as a result of their testimony various truths about the thoughts, feelings, beliefs, and desires of those around us, and what we know here far outstrips that which we could infer just from observations of people's facial expressions, tones of voice, and bodily postures. John Hardwig aptly summarizes our epistemic situation: "If the metaphor of foundation is still useful, the trustworthiness of members of epistemic communities is the ultimate foundation for much

[9] For more on this, see Hardwig (1991); de Ridder (2014).

[10] For an excellent discussion of the nature of testimonial dependence, see Goldberg (2011).

of our knowledge" (1991: 694).

Another feature of our mutual epistemic dependence to which I wish to draw attention is the manner in which we are liable to be influenced by the doxastic norms of the communities in which we find ourselves. That is, the degree to which a proposition appears plausible or worthy of being taken seriously by an individual is partly—and when it comes to certain subject matters, *significantly*—influenced by what is taken for granted in the community of which she is a part.[11] It's worth stressing that we appear to be more or less susceptible to such influence depending on the subject matter of the proposition in question (among other factors). Those subject matters that don't admit of straightforward empirical verification and which are such that shared convictions about them are a key part of what makes for community formation are the subject matters concerning which we are most liable to be influenced by the prevailing assumptions of our communities. The spheres of politics, ethics, metaphysics, and religion all appear to be of this character. William James noticed that a religious proposition that appears utterly incredible to a person in one culture or community can at the same time seem to be a very real possibility to someone trained in a different set of assumptions.

[11] Sociologist Peter Berger has coined the term "plausibility structures" for the mechanisms by which a hypothesis comes to seem plausible or implausible in a given cultural setting. He writes that "The reality of the Christian world depends upon the presence of social structures within which this reality is taken for granted and within which successive generations of individuals are socialized in such a way that this world will be real to them. When this plausibility structure loses its intactness or continuity, the Christian world begins to totter and its reality ceases to impose itself as self-evident truth" (1990: 47).

James writes,

> Let us give the name of *hypothesis* to anything that may be proposed to our belief. And just as the electricians speak of live and dead wires, let us speak of any hypothesis as either *live* or *dead*. A live hypothesis is one which appeals as a real possibility to him to whom it is proposed. If I ask you to believe in the Mahdi, the notion makes no electrical connection with your nature,—it refuses to scintillate with any credibility at all. As an hypothesis it is completely dead. To an Arab, however (even if he be not one of the Mahdi's followers), the hypothesis is among the mind's possibilities: it is alive. This shows that deadness and liveness in an hypothesis are not intrinsic properties, but relations to the individual thinker… A living option is one in which both hypotheses are live ones. If I say to you: "Be a theosophist or be a Mohammedan," it is probably a dead option, because for you neither option is likely to be alive. But if I say: "Be an agnostic or be a Christian," it is otherwise: trained as you are, each hypothesis makes some appeal, however small, to your belief. (1979 [1896]: 2-3).

Building upon the foregoing observations, then, we might characterize the metaphysically contingent state of affairs that I am labelling "S" in the following manner. There exists a race of creatures, namely, humans, who are considerably mutually epistemically dependent, that is, who possess a cognitive constitution which is such that,

(i) Without relying upon the testimony of others, no individual creature pos-

sessing cognitive abilities within the range that is normal would be capable of acquiring anywhere near as extensive a range of knowledge as is actually possessed by most adult humans.

(ii) Individuals are disposed to testify to others when doing so seems likely to transmit useful or interesting truths.

(iii) Individuals are disposed to trust the testimony of others in the absence of obvious defeaters for that testimony, especially when the testifier is someone personally known and trusted.

(iv) Given the perceptual and reasoning abilities that are normal for such creatures, it is frequently far from obvious to them how to adjudicate metaphysical, religious, political, and moral questions with reference to their perceptual experiences and their native inferential abilities.

(v) Individuals' judgments about whether a given hypothesis is "live" or is plausible are prone to be influenced by the prevailing views of their peers, especially when a hypothesis is not susceptible of straightforward empirical investigation.

(vi) The task of evidence-gathering on a wide range of topics tends to be widely shared and delegated.

It should not come as a surprise that mutually epistemically dependent creatures, such as we are, tend to some extent to believe things *en masse*, so to speak. More specifically, it should not come as a surprise that a given proposition con-

cerning the nature of ultimate reality would be believed by many such creatures in some communities and by hardly any of them in other communities[12]—at least, insofar as the proposition in question is not one of those core propositions in which all such creatures are naturally disposed to believe. After all, such creatures have bodies, and given that they have bodies and hence must be located in particular physical locations alongside some of their fellow creatures and not others, and given that they form communities that are mutually isolated to some degree, the extent to which some creature A is liable to be influenced by the testimony of another creature B depends greatly upon the physical proximity of A to B. And of course, insofar as such creatures are social in nature, the degree to which they trust the testimony of others will be influenced by the degree to which they personally know and trust others. What's more, propositions concerning such matters as metaphysics, religion, politics, and morality tend to be such that, for creatures who are subject to such cognitive limitations as we are, it is frequently far from obvious how to adjudicate the truth or falsity of these sorts of propositions with reference to perceptual experience and our native inferential abilities. So, even despite the fact that all such creatures share broadly the same range of perceptual experiences, that fact significantly underdetermines *which* non-core religious propositions such creatures will end up believing. Therefore, whereas one would not expect much divergence among such creatures with respect to their

[12] Included in this: it's not surprising that they form mutually insulated communities, and, for instance, not surprising that they exist in small groups spatially isolated from one another, for a while. (Thanks to Brian Leftow for pointing this out.)

beliefs about the physical properties of publicly observable, medium-sized physical objects,[13] it would not be at all surprising to find that they diverge considerably with respect to their non-core religious beliefs. And since they are liable to be influenced most heavily by the testimony of those who are most physically proximate, and perhaps in particular, by those whom they personally trust and respect, we would expect to see that the non-core religious beliefs of such creatures tend to cluster up geographically and culturally in something a lot like the way that we in fact do observe that human non-core religious beliefs cluster up. For these reasons, Pr(E/S) is very high.

§4.3 *De Re* Relationship with God

My task in this section will be to try to remove a potentially powerful obstacle to my claim that it wouldn't be surprising were God to bring about S. The potential obstacle in question is the thought that God would not bring about S given that it has E as such a likely consequence, and given that an uneven distribution of theistic belief (i.e. E) would mean that whole regions of the world are cut off from the possibility of personal relationship with God. If it is indeed true that lacking explicitly monotheistic beliefs means that one cannot personally relate to God, then Pr(S/theism) is indeed very low. But what I shall now try to show is that mere

[13] I don't mean here to include their theoretical, scientific beliefs about the properties of physical objects, but rather, just their "everyday" beliefs about such matters.

knowledge of core propositions, which (as I argued in Chapter 3) is widely available even to those who are not subject to specifically monotheistic cultural influences, is sufficient for having the cognitive contact with God that is required for engaging in some kind of meaningful personal relationship with God, even if only *de re* relationship with God.

§4.3.1 Personal relationship

Before proceeding further it will be helpful to get somewhat clearer on the nature of personal relationship. I don't propose to attempt here the quite possibly futile task of trying to locate a set of non-circular necessary and sufficient conditions for personal relationship. But I do want to draw attention to several crucial aspects of personal relationship, which will be highly pertinent to the discussion that follows.

1. Reciprocity. Schellenberg has rightly emphasised that personal relationship is necessarily a reciprocal phenomenon.[14] A personal relationship, by its nature, involves mutually recognized, two-way contact between two people. As Schellenberg puts it, "The common meaning here, or as we might say, the relevant content of personal relationship is a pattern of reciprocal activity: a giving and receiving directed towards (among other things) a deeper knowing and being known, in which both terms of the relation participate and which both find

[14] Note that as I am using the term "reciprocity," reciprocal interaction between two people does not entail that the two parties make equal contributions to the interaction.

valuable" (2002: 42). Schellenberg appears to be correct, in that it seems wrong to say that one has a personal relationship with someone if there has only ever been one-way contact between oneself and the individual in question. In the film *The Truman Show*, the main character Truman is (unbeknownst to him) living in an enormous studio set in which he is being secretly filmed at every moment of his life and having the film of his life broadcast to the outside world in the format of a daily TV show. Despite that some lifelong viewers of the show are intimately familiar with Truman's personality traits, his life history, and so on, none of those viewers has a personal relationship with Truman since none of them has ever engaged in any kind of two-way interaction with him.

2. Expression of non-propositional attitudes. Relatedly, personal relationship involves the mutual expression of certain non-propositional, emotional attitudes.[15] As Robert Roberts (2009) has noted, different non-propositional attitudes are characteristic of different types of relationships. Concern for the good of the other is an attitude whose mutual expression is characteristic of friendship; sexual desire is characteristic of romantic relationship; hatred is characteristic of enmity. Roberts emphasises that "actions are not constitutive of a friendship proper unless they express the emotions characteristic of friendship" (2009: 286). I take it the same point goes, *mutatis mutandis*, for other types of relationships besides friendship. Episodes of thought, action, or feeling presumably cannot be constitutive of a

[15] This feature is at least characteristic of typical human-to-human relationships. Chris Kyle suggested to me that that Vulcans (the species in *Star Trek* who claim to have no emotions) may perhaps be able to have personal relationships—though I am doubtful.

relationship of enmity unless those episodes express a wish for the other to be worse off in some way. To the extent that their thoughts, feelings, and actions towards one another cease to express that emotional attitude, two people will cease to be enemies.

3. Second-personal knowledge. Having a personal relationship with someone entails knowing him or her, in a sense of "knowing" that is distinct from propositional knowledge.[16] Eleonore Stump (2010: Ch. 3) has recently written about second-personal knowledge—"Franciscan knowledge" as she calls it[17]—which, she argues, is irreducible to knowledge of propositions about other persons.[18] In support of its irreducibility to propositional knowledge, Stump offers an ingenious adaptation of Frank Jackson's "Mary's room" thought experiment.[19] In Stump's adaptation, Mary has grown up in solitary confinement, never interacting personally with another human being, yet having access to an exhaustive array of propositional knowledge about other people: "In short, Mary has been kept from

[16] Presumably personal relationship with someone does also involve knowing some propositions about him or her.

[17] In fact, for Stump, Franciscan knowledge includes second-personal knowledge but more else besides. It seems that Franciscan knowledge, for her, is any kind of objectual knowledge (otherwise known as acquaintance) that is not reducible to knowledge that some proposition is true.

[18] Stump (2010: Ch. 3) gives the designation "Franciscan knowledge" to this kind of knowledge in honour of St Francis of Assisi, whom Stump regards as having lived a life which prioritised lived experience of the truths of Christianity over what she calls "Dominican knowledge"—that is, propositional knowledge, which Stump thus designates in honour of St Dominic, whose life is said to have been characterised by the quest to convince others of the propositions central to the Christian faith.

[19] See Jackson (1982).

anything that could count as a second-person experience, in which one can say "you" to another person. And then suppose that Mary is finally rescued from her imprisonment and united for the first time with her mother, who loves her deeply" (2010: 52). Stump suggests that upon having a personal encounter with her mother for the first time in her life, Mary will gain some new knowledge, but what she comes to know won't be the sort of thing that can be reduced to a proposition (2010: 52-53). The kind of second-personal knowledge to which Stump calls attention seems to be entailed by having a personal relationship with someone. Consider those lifelong viewers of the TV show of Truman's life. For all their knowledge of propositions about Truman and even acquaintance knowledge of what Truman is like, both of which they have acquired by way of watching his life unfold on the TV every day, the kind of second-personal knowledge to which Stump calls attention is exactly what those viewers lack.[20]

4. Final value. Personal relationships can be more or less valuable. It seems plausible, moreover, that personal relationships can have *final* value; that is, can be valuable for their own sake and not merely for the sake of other ends to which

[20] Matthew Benton (Unpublished) has proposed some necessary conditions for second-personal knowledge. For one thing, he suggests that if *S* personally knows *R*, then *S* has had first-hand causal contact with *R*, which is mutually recognized by *R*. Benton suggests that this requirement gets the right result in the case of the biographer, who knows all manner of propositions about the person she is researching, yet would recognize that there is an important kind of knowledge of him that she lacks and wishes she had, namely, second-personal knowledge. Moreover, Benton suggests a symmetry condition for second-personal knowledge, according to which *S* personally knows *R* only if *R* personally knows *S*. He argues that this get the right results, for instance, in cases of death, where upon learning of the death of their loved one, relatives and friends will immediately switch to the past tense: they will say they personally *knew* her but not that they still *know* her.

they may be a means.[21] And I would suggest that it is sufficient for a token personal relationship to have final value that the parties to that relationship express concern for one another, where that concern for the other is not merely concern for how one may benefit from being related to the other.

§4.3.2 Personally relating to God

Theistic religions, especially Judeo-Christian theism, claim that humans are designed to have personal relationship with God.[22] Yet if God exists and has indeed designed humans for relationship with Himself, the mode of interaction between God and a human being prior to receiving the beatific vision seems likely to differ in some fairly significant ways from the mode of interaction in typical human-to-human cases of personal relationship. Consider the following features:

(i) *Asymmetry*. For any given human being, God's propositional knowledge about that human being vastly outstrips that human being's propositional knowledge about God.

(ii) *Non-locality*. Setting the incarnation aside, God has no particular physical location that a human being can address and act upon, as in the typical

[21] Finnis (2001), notably, holds that personal relationship is among the basic goods of human life, which are valuable in their own right.

[22] Augustine put this thought best: "[T]o praise you is the desire of man, a little piece of your creation. You stir man to take pleasure in praising you, because you have made us for yourself, and our heart is restless until it rests in you" (1991: 5).

human-to-human case.

(iii) Uncertainty. The manner in which God seeks to communicate with a human being is such that often (though not necessarily or always) it is uncertain for the human being whether it is really God who is trying to communicate with her, and uncertain exactly what, if anything, is being communicated.

Lest one worry that personal relationship between two people is impossible when one or other of these features is present, it will be helpful to try to find a human-to-human case that exhibits these features to a significant degree and yet seems intuitively to be a case of personal relationship: [23]

The Tapping Case

Since the death of his father, Jones has fallen in with a bad crowd—drug running, money laundering, armed robbery, and all the rest—and the strong arm of the law has finally caught up with him. He is now locked away in solitary confinement in a small, dark prison cell for what will be an indefinite duration, as far as he knows. In the depths of his solitude and despair, Jones recalls that he has an uncle, Smith, whom Jones scarcely remembers, not having seen him for decades. But something Jones does remember is that his uncle Smith long ago promised Jones that if he (Jones) ever got into real serious trouble, he (Smith) would come to his aid. One

[23] This case is inspired by one originally offered by Ted Poston and Trent Dougherty (2007: 190-91), though it differs in a number of respects. Poston and Dougherty were drawing attention to the uncertainty aspect, but the following case exemplifies the other two features as well.

day, Jones hears a faint tapping sound in his cell, which seems as though it might be another person attempting to communicate with him. Jones can't discern from which direction the tapping is coming, but he responds in the only way he can think to do: by tapping back on the wall of his cell. Jones' taps seem to be met with more taps in response, and Jones gradually begins to find that with considerable effort he seems to discern specific messages in the sequences of taps: messages of encouragement—though sometimes he isn't able to make out any message given his previous patterns of translation. As their tapping correspondence develops, Jones takes himself to be learning some things about his companion: about his hobbies, his likes and dislikes, some of his life history. Jones suspects—indeed very much hopes—that it really is his uncle Smith with whom he is communicating, but he can't be sure, and at times he seriously wonders whether he is simply imagining patterns and messages where there are in reality none. Nevertheless, his apparent companion is the only thing keeping Jones sane; indeed, the only thing keeping him from attempting suicide. What's more, Jones has begun to feel towards his tapping companion the kind of deep affection and concern that he had rarely felt toward anyone in his drug-running days. In fact, the tapper on the other end is indeed Jones' uncle Smith. Smith, good to his word, has infiltrated the prison and gained employment as a guard, and furthermore, has acquired the responsibility for monitoring Jones day and night by way of the hidden cameras and microphones that are installed in Jones' cell. Smith knows all about Jones' wayward escapades and about the troubled circumstances that precipitated his downward spiral in life, having read the extremely detailed police file on him, and now Smith intends to find a way to get Jones free. In the meantime, however, he has resolved to befriend and en-

> courage Jones in the only way available to him: by communicating with him via tapping on the pipes that run through Jones' cell. Eventually, Smith finds his opportunity to engineer a jailbreak, and the two men finally meet face-to-face.

Now, clearly, the manner in which Jones and Smith interact with one another is rather unlike the usual manner in which human friendships are conducted. Indeed, it looks as though all three of the aforementioned features—asymmetry, non-locality, and uncertainty—are present in this case to a significant degree. Prior to their meeting face-to-face, Smith knows hugely many more propositions about Jones than Jones knows about Smith. For Jones, moreover, there is no particular physical location that can be associated with the presence of Smith. Finally, Jones is subject to a substantial degree of uncertainty about whether Smith is really there at all, and about what, if anything, Smith is trying to communicate to him. And yet, intuitively, Smith and Jones have a valuable personal relationship with one another even prior to their face-to-face meeting. Their interactions exemplify the four aspects of personal relationship that I outlined in §4.3.1. They are engaging in reciprocal interactions that express various non-propositional attitudes characteristic of friendship, and they can be correctly said to be gaining some second-personal knowledge of one another. Their relationship seems to instantiate the final value that I suggested relationships have when their members care about one another not merely for the sake of what they can get from one another. Indeed, these episodes in their relationship are such that when Smith and Jones eventually to meet face-to-face, they will regard these earlier episodes as

something to be deeply cherished. I submit, then, that something analogous is indeed possible between God and a human being.

§4.3.3 De re relationship with God

Having tried to motivate the thought that personal relationship between God and a human being is possible despite the asymmetry, the uncertainty on the human side, and the fact of God's non-locality, I now want to make a stronger claim: that if God exists, a human being can have a valuable personal relationship with God *without realizing that it is God with whom she is in relationship*,[24] provided that she picks out God using at least some important descriptions. The sorts of descriptions I have in mind, in fact, are those that feature in the core religious content that in Chapter 3 I argued could be known by way of the cognitive mechanisms described by CSR: descriptions such as "the moral lawgiver," "the personal creator of the universe," "my ultimate benefactor," and so on.

For a person to relate to God without realizing that it is God with whom she is

[24] There are, broadly speaking, two major views as to what sort of term "God" is. On one view, "God" is a titular term; a definite description, whose unabbreviated form is something like "the necessarily existent, omnipotent, omniscient, morally perfect, personal creator and sustainer of the universe" (see, e.g., Rachels 1971: 333). On this view, to relate to a person *a* without realizing that *a* = *God* is to be ignorant of the fact that the description "the necessarily existent, omnipotent, omniscient, morally perfect, personal creator and sustainer of the universe" is true of *a*. On another view, "God" is a proper name for the divine being (see, e.g., de Ridder and van Woudenberg 2014: 56). On this view, to relate to a person *a* without realizing that *a* = *God* is to be ignorant of the fact that *a* is the referent of the proper name "God."

in relationship would be for her to relate *de re* to God. That is, she wouldn't self-describe her situation in terms of "being in relationship with *God*," since she wouldn't realize that it is God to whom she is relating, even though a relationship with God would in fact be what is occurring.[25] Again, though, since it may be less than obvious that it is even possible to engage in personal relationship with X without realizing that one is in relationship with X, let's look at what seems to be a human-to-human case of *de re* relationship:

The Case of the Anonymous Helper

Molly and Ted work in the same office building, but have never spoken to one another. One day, whilst staring out across the office in idle moment, Ted notices that Molly appears to be extremely stressed—on top of all her other responsibilities, she has just been assigned the burdensome task of stuffing hundreds of envelopes with invitations for an upcoming promotional event. When everyone else including Molly has gone home, Ted decides to stay late and stuff the envelopes for Molly, leaving only an anonymous note: "I saw you looking a bit overworked—I hope this helps." When Molly comes in the next morning and finds the note with the completed pile of envelopes she feels deep gratitude to the person who did this kind thing for her. She goes out and buys an expensive box of chocolates, and leaves them in the break room with a note: "These are for the lovely person who helped me out yesterday—thank you so much, whoever you are!" Once everyone else is

[25] As Benjamin Cordry (2008: 4) notes, to *de re* ascribe an attitude or action to *S* concerning an object X is to make an ascription that doesn't attempt to capture how *S* herself might characterize her attitude or action and its object.

> out of sight, Ted gratefully takes the gift, and replaces it with a note of his own: "Thank you for these! I'm glad to have been able to help." We can imagine that they carry on in this fashion for a while until Molly finally learns that her anonymous helper was Ted, whereupon the two of them begin to engage in friendship in the more usual sorts of ways.

Molly in this case starts out by personally relating *de re* to Ted.[26] She wouldn't have described her situation in terms of "relating to *Ted*," since she didn't recognize the person who helped her out as Ted, but her relating to Ted was nonetheless what was really going on. Importantly, it looks as though all of the aspects of personal relationship that we identified in §4.3.1 are present in this case. Molly and Ted are engaging in mutually recognized two-way contact with one another, and in doing so are expressing a variety of non-propositional attitudes towards one another—sympathy, concern for the wellbeing of the other, gratitude, and so on. They seem, indeed, to gain some significant second-personal

[26] The term "*de re*" has tended to be used in philosophical discussions (e.g. Chisholm 1976) in a way that implies that if one has a *de re* attitude towards an object then one is able at least to single out the object in question. For example, if one believes that "Mark Twain is a great author" whilst being ignorant that Mark Twain is identical with Samuel Clemens, then one has a *de re* belief of Samuel Clemens and is able to single out the *re* that one's belief is about. However, in *The Case of the Anonymous Helper*, it seems that Molly can't single out the object of her beliefs and attitudes in a similarly precise way—after all, for all she knows the pile of envelopes was stuffed on her behalf by several people rather than just one person. But I shall be using the term "*de re*" in a way that does not have the implication that if one has a *de re* attitude towards X then one can single out X. Rather, when I speak of someone's "having *de re* relationship with X," what I shall mean is that relationship with X is what is in fact occurring even if the subject would not self-describe her situation in terms of being in relationship with X.

knowledge of one another through these episodes, and they appear to care about one another not merely for the sake of what they can gain from one another. And I take it that when Molly finally comes to learn that Ted was the person who so kindly helped her out and thus subsequently engages in a friendship with Ted in the more usual manner, they will both look back fondly upon those episodes and regard them as having contributed importantly to their personally knowing one another. I contend, then, that it is possible to have a valuable personal relationship with someone even though one is merely relating *de re* to that person.

Now, I want to suggest that if God exists, something analogous to *The Case of the Anonymous Helper* can and likely does occur between God and many human beings who recognize God only under a limited description.[27] Clearly enough, though, it matters a good deal *which* descriptions someone uses to pick out God. When Molly is relating *de re* to Ted, the description she uses in thinking about Ted is something like "the person who helped me out yesterday at a really stressful time." It is in virtue of her thinking about Ted under this description that she is liable to feel such gratitude towards him and to want to express her gratitude towards him in the sorts of ways that she does. If instead Ted had not helped out Molly in this

[27] Something like this claim has had some noteworthy defenders, including Karl Rahner and possibly C.S. Lewis. Rahner (1976: 283) held that it is possible for a non-Christian to be an "anonymous Christian"; to possess saving faith in the God of Christianity through an implicit awareness of God which is ubiquitous. Lewis may have at least had toyed with a similar idea. In the final instalment of *The Chronicles of Narnia*, *The Last Battle*, followers of the imposter Tashlan appear to have their faithful service to Tashlan counted as service to the true king Aslan, perhaps hinting that faithful devotees of non-Christian religions may be implicitly following Christ, though under another name.

way, but rather, had (say) merely disposed of the used coffee cup on her desk, and if Molly had thought about "the person who disposed of my used coffee cup," she would presumably have felt a far weaker and more fleeting sort of gratitude towards Ted than she in fact did feel towards him and she presumably would not have been motivated to engage in the sorts of expressions of gratitude that she did.

A similar point seems to be true of the divine-human case. Only some of the infinitely many definite descriptions that God satisfies, if He exists, are such that if one thinks about God under those descriptions, then one will be able and liable to express the appropriate sorts of emotional attitudes towards God.[28] Suppose, for instance, that Jill has the belief that "the person who put the TV remote back in its place is thoughtful," and that Jill feels very mild and fleeting gratitude towards "the person who put the TV remote back in its place," and suppose, moreover, that in fact God is the person who put the TV remote back in its place (through an instance of special divine action). It seems pretty clear that if Jill only thinks about God under the description "the person who put the TV remote back in its place," she won't be liable to feel the appropriate sorts of emotional attitudes towards God, nor to attempt to express gratitude or worship. The sort of gratitude that is fitting for Jill to have towards God (if God exists) is of a far deeper sort—gratitude

[28] Obviously, one can have a personal relationship with someone that involves inadequate attitudes towards that person. But it would be better, all else equal, if the sorts of descriptions under which humans are able to know God are such they render humans liable to have the appropriate sorts of attitudes towards God.

for her very existence and the good things it contains.[29]

Other descriptions that God satisfies, however, seem to be such that thinking about God under those descriptions will render a person liable to feel and express the appropriate sorts of attitude. The definite descriptions I have in mind include "the moral lawgiver," "the personal ground of being," "the creator of the universe (or some specific part thereof)," "The Good." And as I argued in Chapter 3, various propositions involving such descriptions, and involving existential generalizations of the form "there is an *x* such that *x* is the ___ (where the blank is filled in with one or other of the foregoing descriptions)," can be known by people across various cultures as a result of the operation of the maturationally natural cognitive mechanisms described by CSR. A person who knows that "there exists a creator of the universe" and who feels emotions of awe and wonder towards "the creator of the universe" will thereby be feeling those attitudes *de re* towards God. The same is true of the person who knows that "there exists a personal ground of my being," and who feels gratitude towards "the personal ground of my being." The sort of gratitude that is involved here is plausibly the sort of gratitude that is fitting for a person to have towards God, if He exists—that is, gratitude for one's very life and the good things it contains. Again, if someone strives to obey "the moral lawgiver," then her striving will be directed *de re* towards God.[30]

---

[29] After all, if God exists, He is, as Swinburne (1989: Ch. 8) puts it, our "ultimate benefactor."

[30] Unfortunately, I don't have the space here to explore how an outright atheist—someone who thinks that such descriptions as "the personal creator of the universe" don't refer—might relate *de re* to God. My suspicion is that it is possible, though. Atheists often do have feelings of gratitude

§4.3.4 Relating to God under mistaken descriptions

One might very reasonably wonder how *de re* relationship with God is supposed to work in cases where the human party to the would-be relationship has false beliefs about who it is that satisfies descriptions like "the moral lawgiver" and "the creator of the universe" and so on, or where these descriptions are included within more detailed descriptions that are overall false of God, for instance, "the moral lawgiver and the one who placed the earth on the back of an elephant." I want to attempt to develop a couple of thoughts about how this worry might be addressed.

One thought is this. There may be some such cases—though perhaps not all will be like this—in which a subject uses descriptions that God does not satisfy and yet the subject nonetheless succeeds in picking out God by way of employing those descriptions in a purely referential manner. This thought rests on a view about a certain use of definite descriptions that Keith Donnellan (1966) has defended. In short, Donnellan has suggested that there is a purely referential use of definite descriptions wherein a description may successfully refer to an individual even though the individual in question doesn't fulfil the description:

> [I]n the referential use the definite description is merely one tool for doing a certain

---

towards something that is extensionally equivalent with God if God exists (e.g., the ground of being, ultimate reality), and moreover, many atheists strive to live in accordance with a moral law, which striving may well amount to *de re* obedience towards God if God exists. Also, Brian Leftow has suggested to me that outright atheists are part of the overall population, and their outright atheism is a matter for them individually, so this particular problem can perhaps be overlooked in dealing with the specifically demographic problem.

> job—calling attention to a person or thing—and in general any other device for doing the same job, another description or a name, would do as well. In the attributive use, the attribute of being the so-and-so is all important, while it is not in the referential use. (1966: 285).

Donnellan adduces various cases in support of this claim. In one such case (1966: 286), Jones is on trial accused of the murder of Smith. Jones' erratic behaviour is being discussed by the court, and it emerges that Jones is not mentally well. In fact Jones is innocent of the murder, but someone in the gallery, who mistakenly assumes Jones' guilt, says that "Smith's murderer is insane." Despite the fact that the description "Smith's murderer" is actually false of Jones, it seems plausible that the subject here succeeds in saying something true of Jones: namely, that he is insane.

There appear to be at least a couple of mechanisms by which a subject might succeed in picking out a certain object with a description, despite the fact that the description is false of that object. For one, it might be that the subject who is using the description has some kind of direct cognitive contact with the object and is merely using the definite description as a means of ostension. By "direct cognitive contact," I mean that the object is in some way present to the consciousness of the subject, perhaps as an object of her perception or as a proximate cause of some salient aspect of her perceptual experience. In a case in which the subject has

direct cognitive contact with God,[31] the subject might use a description that is false of God, and yet, because the description is simply a means of "pointing to" the object of her experience, she nonetheless succeeds in referring to God. Analogously, suppose that in *The Tapping Case* discussed in §4.3.2, Jones holds a number of false beliefs about the person with whom he is in contact—perhaps due to the considerable difficulty of discerning the messages that are coming to him by way of the tapping sounds. Suppose that he thinks about his tapping companion under such mistaken descriptions as "the person who likes to eat cold porridge for breakfast on Wednesdays" or "the person who won the Tour de France in 2002." Still, insofar as Jones is simply using such descriptions as a device for gesturing at whoever it is that is behind the tapping of which he is consciously aware, he succeeds in picking out the right person, namely, Smith.

Another mechanism might be as follows. Suppose the subject employs a description involving several properties—for instance, "the person who is F, G, and H"—some of which are possessed by the object and some which are not. The mechanism by which the subject succeeds in referring to the object with that description might be the *greater salience*, in the context at issue, of the properties that are possessed by the object as compared with the properties that aren't.[32] What I mean is this. Recall *The Case of the Anonymous Helper* that I presented in §4.3.3.

[31] Someone might have direct cognitive contact with God by way of a perceptual religious experience of the sort William Alston (1991) discusses, or an experience of a striking and unusual event that is proximately caused by God's special, local activity.

[32] Thanks to Trent Dougherty for raising this suggestion in conversation.

Suppose that Molly is gratefully reflecting upon her pleasant discovery of the pile of stuffed envelopes with the anonymous note, and suppose that she wrongly assumes that the person who is responsible is a woman, but that that assumption isn't playing much of a role at all in causing her to feel grateful for the kind deed. The description Molly is using in thinking about her anonymous helper is "the *x* who is a woman and who kindly stuffed the envelopes for me." But it isn't too implausible, I think, that Molly's gratitude is still directed towards Ted, even though Ted is not a woman. And the reason is that the property of *being the x who stuffed the envelopes for Molly* is significantly more salient than the property of *being a woman* in the context at issue: Molly's belief that someone possesses the former property plays a much greater role in causing her feeling of gratitude than does her belief that someone possesses the latter property. Similarly, it might be that in some cases in which a person uses a description involving some properties that God possesses and some that God doesn't possess—for instance, "the *x* who created the universe and who caused the earth to be flat"—those properties that God does possess are more salient than those that God doesn't possess, in the sense that the person's beliefs about the properties that God does possesses are playing a much greater role in generating her feelings of gratitude, reverence, awe, or whatever.

Perhaps, though, there remain some recalcitrant cases, in which neither of these two mechanisms is plausibly exemplified. Perhaps, that is, there are some cases in which the human subject's descriptions fail altogether to pick out God. Still, I

want to suggest that even if her descriptions are such that she fails to refer to God, there may yet be a kind of quasi-relationship with God that is available to her.[33] Consider the following case:

The Case of Santa

Gracie, like many children, has been told that Santa Claus is the person who brings her Christmas presents each year. She has been told all about Santa's Lapland dwelling, his portly appearance and his red and white outfit, and his impressive ability to distribute Christmas presents across the entire globe in a single evening. Of course, the person who really brings the Christmas presents each year is Gracie's father, Bill. But let's grant that in thinking about Santa, Gracie doesn't even implicitly have her father in mind; she doesn't refer to him by the name "Santa." Now, Gracie knows that "the person who brings the Christmas presents is to be thanked" (this is the analogue of a core religious proposition such as "the person who created the universe is to be thanked"). But since she also believes that "the person who brings the Christmas presents = Santa Claus," let's grant that her feelings of gratitude are directed towards the non-existent Santa rather than towards her father, Bill. Nonetheless, Bill, knowing his daughter's character intimately, knows that were Gracie to be apprised of the true identity of the person who brings the Christmas presents, she would direct her gratitude towards him (i.e. Bill). Bill thus chooses to "count" Gracie's gratitude *as though* it were rightly directed towards

[33] Jeroen de Ridder and René van Woudenberg (2014) express scepticism about the notion that adherents of non-Abrahamic religions could refer to the God of Abrahamic monotheism if He exists. This final proposal of mine is supposed to work even granting that such scepticism is well-founded.

> him. Seeing Gracie's delight and thankfulness towards Santa when she opens her presents, Bill is deeply gladdened. And whenever Gracie sends a thank you note to Santa, Bill replies to it warmly in what Gracie takes to be Santa's hand. Eventually Gracie learns of the non-existence of Santa, and of the fact that her father was really the person who brought the Christmas presents each year and replied to her thank you notes. Once she has got over the initial shock, Gracie comes to see those episodes in which she had thought she was interacting with Santa as something to be deeply cherished for what they show of her father's love for her.

Similarly, it seems to me that God might very well "count" the expressions of gratitude, reverence, and obedience that a human being directs towards some non-existent deity *as though* they were directed towards God, so long as God knows that she *would* direct those same attitudes towards Him were she apprised of the true identity of "the moral lawgiver," "the creator of the universe," and so on.[34] Now, even if one insists that in the case just described Gracie was not *in any way* relating to her father during those episodes when she thought she was relating to Santa, it seems hard to deny that those episodes *retrospectively* take on an im-

[34] To be sure, this involves God's having knowledge of a counterfactual truth about the human being in question. But the counterfactual truth here need not be ungrounded in categorical facts in the way that the Molinist's counterfactuals are. Rather, in line with the usual Stalknaker-Lewis semantics for counterfactuals (Stalnaker 1975; Lewis 2001), we may say that what makes it true that "*S* would direct her gratitude or reverence or obedience towards God if she were apprised of the truth" is the fact that by keeping fixed her actual beliefs, character traits, and dispositions to as great an extent as is possible whilst altering her beliefs about the identity of "the moral lawgiver," "the creator of the universe" and so on, we get the result that she directs those attitudes towards God.

portance within her relationship with her father, once she becomes apprised of what was really going on. In that sense, those episodes make an important contribution to the value of her relationship with her father. Likewise, episodes during which a person takes herself to be engaging in relationship with some non-existent deity (without thereby referring to God) could come to have a similar retrospective significance—if indeed it was really God who was causally responsible for that which she took the non-existent deity to be responsible for—once she meets God face-to-face.[35]

§4.3.5 The eschatological dimension

Even supposing *de re* relationship with God is (if God exists) as widely available as I have claimed, wouldn't it still be quite *unfair* of God to permit a situation in which the level of awareness of God that is available to people in various parts of the world is subject to significant variation, and thus still quite unlikely that God would bring about S given its propensity to lead to an uneven geographical distribution of explicitly theistic belief (i.e. E)?

---

[35] In the last of the Narnia Chronicles, *The Last Battle*, C.S. Lewis depicts an episode in which something like this "counting as" appears to be going on. Emeth, an earnest follower of the counterfeit king Tash, finally meets the true king Aslan and is told by Aslan that "I take to me the services which thou hast done to him [Tash]. For I and he [Tash] are of such different kinds that no service which is vile can be done to me, and none which is not vile can be done to him. Therefore if any man swear by Tash and keep his oath for the oath's sake, it is by me that he has truly sworn, though he knows it not, and it is I who reward him" (200).

To begin with, I should note that it is quite commonly held that there can be no successful theodicy with respect to the problem of evil unless the theodicist has recourse to the claim that God would provide an afterlife in which those who have been gravely harmed in this lifetime are compensated or given the opportunity to somehow overcome the evils done to them.[36] I concur with this sentiment, and I suspect something similar may well be true of the problem of divine hiddenness. Without appeal to a post-mortem encounter with God in which each person will have the opportunity to encounter God "face-to-face," the theist's challenge with respect to the problem of hiddenness may well be insurmountable. I should emphasize that appeal to such a post-mortem encounter is by no means *ad hoc*. If there exists a perfectly loving God, then God would be compelled by virtue of His perfectly loving character to seek out as deep a loving relationship with each of His creatures as is possible,[37] and so if such depths of relationship are not made available to a creature during this lifetime, they will be during the next.

Here is how I would make use of appeal to such a postmortem encounter in trying to dispel the worry about the unfairness inhering in the fact that the level of

[36] Dougherty rightly notes that "It is no part of theism that such enlightenment comes only or mostly in this life. Rather, it is part of theism that there will be an afterlife to continue the process. And the afterlife is not an auxiliary hypothesis [to theism]. For there is no possible world [in which God exists] in which creatures suffer and are not given full opportunity for compensation and recovery" (2014: 112-13).

[37] As Schellenberg puts it, "Love—real love—is one of the most awesome qualities of personal being we know of (and even that may be an understatement). How could God exist but lack it? ... the best human love…seeks a kind of *closeness* between itself and the beloved. The lover clearly wants this for its own sake" (2002: 41).

awareness of God that is available to people in various parts of the world is subject to significant variation. Recall again *The Tapping Case* discussed in §4.3.2. Now consider a pair of such cases. In the first case, Jones is communicating with his uncle Smith by tapping on the wall of his cell, and Jones recognizes his uncle under mostly accurate but nonetheless limited descriptions. In the second case, Jones is communicating with his uncle Smith as before, but this time many of the descriptions Jones employs are false of Smith. Now compare both of these cases with a third case, in which Smith has finally set Jones free and the two men meet face-to-face, hug one another, speak with one another, look one another in the eye, share a meal together, and so on. For all the differences between the first two cases as regards the level of Jones' awareness of Smith, there appears to be a world of difference between both of these cases and the third case. If God exists and will extend to everyone the opportunity to meet Him "face-to-face" in an afterlife, then it seems that something closely analogous to this contrast is true when it comes to different people's differing levels of awareness of God in this lifetime as compared with their level of awareness of God in the beatific vision.[38]

Perhaps there is another worry lurking, though: namely, that people who in this lifetime only get the opportunity for *de re* relationship with God are thereby disadvantaged with respect to attaining salvation in comparison with those who

[38] Indeed, as the Apostle Paul wrote in his first epistle to the Corinthians, "For now we see through a glass, darkly; but then face-to-face: now I know in part; but then shall I know even as also I am known" (1 Cor. 13:12).

have the opportunity to relate to God more explicitly.[39] But even if it is true that a person's salvation is dependent upon her response to God in this lifetime,[40] from God's perfect justice it follows that no one will be denied salvation in virtue of factors over which she had no control.

Now, of course, one might respond to all of this that God could simply have given the beatific vision to everyone from the start, just as God could have created a suffering free paradise from the start. That is true. But were God to do that, He would have lost out on the opportunity to realize various goods that I am about to explore in the following section, which goods presuppose human uncertainty, cognitive limitedness in a number of respects, and a tendency to cluster up in our views about the nature of ultimate reality.

§4.4 The Human Cognitive Constitution and Balancing Competing Goods

So far, then, I've tried to show that what seemed initially to be a strong reason for thinking that Pr(S/theism) is very low is not in fact a strong reason for thinking so. But now, more positively, I want to try to show that S is a state of affairs that God would have good reasons to bring about. That is, God would have good reasons for creating beings who are cognitively constituted roughly as we are: beings with

[39] This seems to be part of Maitzen's (2006) worry, at least.

[40] For arguments that salvation is not dependent upon one's response to God in this lifetime, see Mawson (2012: §§8-9; 2005: Ch. 5).

roughly our need and propensity to rely on what others tell us and roughly our inclination to be influenced by what those around us think on matter of deep importance in life, religious matters included. S, as I argued earlier, has E as a highly probable consequence. At this point it will be helpful to recapitulate my earlier characterization of S. S is the actually obtaining state of affairs in which there exists a race of intelligent, morally sensitive, free creatures—namely, humans—who are considerably mutually epistemically dependent; who possess a cognitive constitution which is such that,

(i) Without relying upon the testimony of others, no individual possessing cognitive abilities within the range that is actually normal would be capable of acquiring anywhere near as extensive a range of knowledge as is actually possessed by most adult humans.

(ii) Individuals are disposed to testify to others when doing so seems likely to transmit useful or interesting truths.

(iii) Individuals are disposed to trust the testimony of others in the absence of obvious defeaters for that testimony, especially when the testifier is someone personally known and trusted.

(iv) Given the perceptual and reasoning abilities that are normal for such creatures, it is frequently far from obvious to them how to adjudicate metaphysical, religious, political, and moral questions with reference to their perceptual experiences and their native inferential abilities.

(v) Individuals' judgments about whether a given hypothesis is "live" or is plausible are prone to be influenced by the prevailing views of their peers, especially when a hypothesis is not susceptible of straightforward empirical investigation.

(vi) The task of evidence-gathering on a wide range of topics tends to be widely shared and delegated.

§4.4.1 Competing goods

If God exists and chooses to create intelligent, morally sensitive, free creatures, He has open to Him a range of possible types of such creatures, of which we are just one type. There are possible races of creatures who have significantly greater cognitive powers than us—for instance, who possess the ability to read minds, to perceive spatially remote occurrences, and so on. At the same time, there are possible races of intelligent, morally sensitive, free creatures who have significantly lesser cognitive powers than us—for instance, who lack even our ability to make various competent inferences about other people's beliefs, desires, and intentions from their facial expressions, tones of voice, and body language. If it could be shown that our cognitive constitution is the best one for God to bestow upon a race of intelligent, morally sensitive, free creatures, then the probability that God would create beings with such a cognitive constitution (i.e., that God would bring about S), given that He creates intelligent, morally sensitive, free creatures at all,

would be 1. But I can't think of a way to show that our cognitive constitution is the best one for a race of intelligent, morally sensitive, free creatures to have. Indeed, I can't think of a way to show that there *is* a best cognitive constitution for such creatures to have.

What I'm going to suggest is that there are certain trade-offs between competing goods that would have to be made were God choosing what kind of cognitive constitution to bestow upon a race of intelligent, morally sensitive, free creatures. (Goods here are to be understood as including preventions of harms or evils.) Whilst there are clearly undesirable extremes wherein one good is realized to the excessive detriment of another good with which it is in competition, in between the extremes lies a range of possible configurations none of which is clearly better than another, but any of which is clearly better than the extremes.

By way of analogy, imagine that Jill wishes to encourage her teenage children to appreciate the great importance of sharing valuable commodities among themselves in a fair manner. Jill faces a trade-off, in the sense that the more responsibility she gives them for sharing among themselves, the less she is able to prevent them from being unjust to one another. Suppose that money is the commodity with which Jill is considering entrusting her teenage children. Plausibly, given her goals, there is no uniquely best sum of money with which Jill should entrust them. There are extremes that she has good reason to avoid. Entrusting them with a mere handful of pennies would certainly ensure that they can't really hurt one another in their sharing of what Jill gives them, but it wouldn't afford

them a remotely serious opportunity to appreciate the value of taking responsibility for fairly sharing amongst themselves. At the other extreme, entrusting them with thousands of pounds or more to share among themselves would give them a very considerable degree of responsibility, but it would result in too grave an injustice were the more assertive of the children to dominate the others. In between these extremes lies a range of values such that no single value is obviously better than another, but such that any value in that range is better than any value that lies at the extremes. Supposing that Jill is ideally rational, outstandingly morally virtuous, and has a full grasp of the situation, we should be unsurprised if Jill were to pick any value that lies clearly in between the extremes, and very surprised indeed if she were to pick a value that clearly lies at the extremes.[41] Or consider another analogy. Suppose an architect wants to design a house whose rooms feel light and spacious. Achieving the desired feeling of light and space requires that the architect design rooms that have relatively high ceilings and large windows. But presumably the architect's desire for light and spacious rooms doesn't settle exactly how high the ceilings and how large the windows should be. For sure, if the ceilings are too high and the windows too large, the rooms will feel oddly proportioned and lacking privacy; if the ceilings are too low and the windows too small, the rooms won't achieve the desired feeling of light and space. But plausibly, in between the extremes lies a range of values—perhaps a moderately wide range—none of which is obviously better for the architect's purposes than

[41] I take it that the boundaries between the extreme ranges and the non-extreme range lying in between are going to be vague boundaries.

any other within that range, but any of which is better than values at the extremes. Again, it would be surprising to find that the architect has chosen a value that lies clearly at one of the two extremes, but unsurprising to find that he has chosen any value that lies clearly in between the extremes.

I now want to suggest that S, the state of affairs in which humans have the cognitive constitution that we actually do—in which we are mutually epistemically dependent to roughly the degree that we are, is such as to strike a balance between competing goods that clearly lies between two undesirable extremes, so that it would be unsurprising were God to bring about a race of creatures who are so constituted. It would be unsurprising, that is, were God to bring about S. To that end, I shall discuss three pairs of competing goods, and in each case will try to show that our actual cognitive constitution very plausibly makes for a balance between them that is favourable, even if not uniquely optimal.

§4.4.1.1 Exercising interpersonal trust vs. avoiding risk of deception

By "interpersonal trust," I mean the following: a person *S* places trust in another person *R* just in case *S* is less than certain that *R* has the property of being benevolently disposed towards *S* in some respect—for instance, the property of being a faithful spouse to *S*, the property of being willing and able to catch *S*'s fall, the property of intending to make good on a promise to *S*, the property of loving *S*, the property of being honest towards *S*, and so on—and yet, despite her uncer-

tainty, *S* undertakes action(s) that will have a favourable outcome for *S* only if *R* does indeed possess the property in question.[42] Now, I don't intend this as a theory about ordinary language usage of the term "trust," which may well have a plurality of interconnected meanings.[43] Rather, I contend that the activity that I have stipulated herein as the meaning of "interpersonal trust" is one in which it is very valuable for humans to engage. Why is it valuable? Well, for one thing, insofar as *S* trusts *R* with something important to her, *R*'s morally significant freedom to affect the life of *S* for ill or for good is thereby increased.[44] Obviously there are other ways in which we can freely make deep differences to one another's lives that don't involve interpersonal trust. We can harm or do good to people who are not even acquainted with us, and thus who don't trust us. But the degree to which *R* is free to affect *S*'s life for ill or good is typically greater if *S* trusts *R* than if *S* doesn't trust *R* and all else is equal, especially if *S* trusts *R* with something very important to her. A king can freely make deep differences to the lives of his subjects just in virtue of his wielding coercive political power over them. But a king whose subjects trust him with their wellbeing thereby has a greater level of freedom to make deep differences to their lives for good or ill than does a king

[42] Lara Buchak's (2014) decision-theoretic analysis of faith is very much along these lines.

[43] As Tom Simpson writes, "There is a strong prima facie case for supposing that there is no single phenomenon that 'trust' refers to, nor that our folk concept has determinate rules of use," though he goes onto suggest that "in living socially, people must rely on others to act cooperatively. So it is highly desirable for any community of language users, who necessarily live socially, to possess a simple root concept to refer to this fact. Trust initially answers to this" (2012: 551).

[44] For an excellent account of the good of bearing responsibility for affecting one another's lives in deeply significant ways, see Swinburne (1998: Ch. 8).

with the same level of coercive political power but whose subjects distrust him. The king whose subjects trust him is freely able to bring about a very good thing that the king whose subjects distrust him is not able to bring about: namely, fulfilment of their trust. In addition to this, there seems to be another reason to think that God, if He exists, would wish for humans to have significant opportunities to trust one another. According to theistic religions, perhaps especially Judeo-Christian theism, God wants human beings ultimately to place trust in Him,[45] and so by constituting us in such a way that we are frequently afforded opportunities to place trust in one another, God thereby provides us with a very tangible analogy for how He desires for us to relate to Him. Why would God want humans to trust Him? For one thing, if God invites humans to depend upon Himself whilst not making fully plain to them either His existence or His exact purposes, He thereby enables humans to demonstrate the seriousness of their commitment to Him in a way they wouldn't be able to do under conditions of full certainty.[46] For another, God is likely to want humans to relate personally to Himself out of the appropriate sorts of motivations, and not just because (say) they fear Him or want Him to perform miracles on their behalf.[47] Poston and Dougherty suggest that "the kind

[45] See, e.g., Psalm 84:12; Proverbs 3:5; John 14:1; Romans 15:13.

[46] The author of the New Testament book of Hebrews highly commends Abraham for the way that he followed God whilst being uncertain where it would take him: "By faith Abraham, when he was called, obeyed by going out to a place which he was to receive for an inheritance; and he went out, not knowing where he was going" (Hebrews 11:8).

[47] For an extended discussion of various ways in which human motivations for relating to God could be deficient, see Howard-Snyder (1996).

of relationship God wants is one in which the agent longs for God in a way that is best accomplished in many individuals via a period of doubt" (2007: 184). If God permits a situation in which humans have to trust Him due to their being uncertain about His intentions or even His very existence, then He thereby enables humans to cultivate just this sort of longing for Him.

Clearly enough, having the opportunity to engage in interpersonal trust, as I have defined it, requires having some degree of uncertainty about the relevant properties of the person in whom one places one's trust.[48] And this in turn requires that we are not so constituted as to be able very easily to acquire certainty on a wide range of propositions concerning one another's thoughts, intentions, feelings, capabilities, and so on, independently of what others tell us about themselves.[49] In short, having significant opportunities to place interpersonal trust in one another requires our being cognitively limited in roughly the ways in which we in fact are: in particular, it requires that we cannot simply read one another's minds, and more generally, that others can do things in private about which we cannot learn independently of their testimony.

---

[48] Perhaps there is a different activity that could be called "interpersonal trust," such that one may trust in God in this sense even in the beatific vision—i.e., even whilst being certain of God's existence and benevolence towards oneself. I don't wish to dispute that thought. To reiterate, all I am claiming is that the activity I have stipulatively labelled "interpersonal trust," which does involve uncertainty of the sort outlined, is one that, intuitively, it is valuable for humans to engage in.

[49] Of course, insofar as we learn about one another via one another's testimony and act upon that testimony, and insofar as we are less than certain about the veracity of that testimony, we thereby exercise interpersonal trust in one another.

Now, the good of having opportunities to trust one another is in tension with the good of not being vulnerable to deception or betrayal by one another. The more we are cognitively constituted so as to have to trust one another, the more we will be at risk of being duped by one another. But the degree to which humans are in fact dependent upon one another's testimony for learning about one another and the world is such as to permit a favourable balance between having significant opportunities to place trust in one another, on the one hand, and yet still being capable of discerning enough about one another via non-testimonial means so as to prevent us from being excessively at risk of deception by one another, on the other hand. Consider, for example, the degree to which we are able, independently of testimony, to learn about one another's inner lives and about what is going on elsewhere in the physical world. If we were a lot more capable in these regards than we actually are—if, say, we could read one another's minds, had X-ray vision, and could immediately perceive events in distant spatial locations—we would rarely have any opportunity to place trust in others, and what's more, it would be very difficult indeed for us to keep things to ourselves, which in turn would make it very difficult to engage in the sorts of free actions for good or ill that require that our thoughts and intentions are not immediately obvious to everyone around us. On the other hand, if we were much less able in these regards—if, say, we were unable to make the sorts of competent inferences that we often make concerning people's thoughts, intentions, and feelings from observing their facial expressions, body language, and tones of voice—it would be

far harder to make competent initial judgments about whether to trust a person, which in turn would render us excessively vulnerable to other's evil intentions.

§4.4.1.2 Sharing responsibility for one another's acquisition of epistemic goods vs. practicing epistemic self-reliance

By epistemic goods, I have in mind such goods as true belief, justified belief, knowledge, understanding, wisdom, and so on.[50] Sharing responsibility for one another's acquisition of such goods adds substantially to the range of morally significant free actions that are available to humans, and hence, to the range of morally good states of affairs that humans are able to realize.[51] Specifically, it affords a human person the morally significant choice between depriving others of their knowledge or sharing it with them, between willfully distorting or suppressing the truth and trying their best to relay it faithfully, between co-operating with others in the search for truth or trying to thwart others' in that search, and so on. As Robert Roberts and Jay Wood rightly note, "It is possible to be generous in the conduct of intellectual practices, or stingy and greedy, just as in the handling of material and other goods. The reason is that the intellectual life has its own set of goods, and these can be shared with others and given to others, or grasped and hoarded for oneself" (2007: 292).

[50] For a comprehensive list of epistemic goods, see Roberts and Wood (2007: Ch. 2).

[51] Richard Swinburne goes so far as to suggest that "We cannot do good or harm to each other in many ways without humans already cooperating in the acquisition of knowledge" (1998: 147).

Now, the good of our sharing responsibility for one another's acquisition of epistemic goods is obviously in tension with what some epistemologists (e.g. Fricker 2006) also regard as a significant good, namely, epistemic self-reliance. To practice epistemic self-reliance is to try to figure things out by oneself. Roughly speaking, the more our cognitive constitution is such as to force us to be epistemically reliant upon one another, the fewer opportunities we will have to practice epistemic self-reliance. Ryan Byerly (2014) has recently argued that the value of epistemic self-reliance lies in the way in which cognitive successes that are attained solely through one's *own* cognitive competences typically count as greater cognitive achievements than they would do were those same successes the result of relying upon others.[52] I don't wish to dispute this. But the way in which humans are actually cognitively constituted is evidently somewhere in between the twin extremes of being so cognitively weak that we are forced to rely upon one another almost continually and being so cognitively well-endowed that we seldom need help from others in order to get at the truth. Our actual cognitive constitution seems to be such as to permit a rather favourable balance between the competing goods of sharing responsibility for one another's intellectual wellbeing and practic-

[52] In support of this judgment, Byerly (2014: 60-61) offers a pair of cases, one in which an agent, Edward, works hard at solving a Sudoku puzzle and arrives at the solution all by himself, and another in which Edward's wife Jen, who is more adept at solving such puzzles, simply shares the solution with him. Byerly suggests that intuitively, despite the fact that Edward ends up acquiring the same item of knowledge in both cases, the former case obviously instantiates something valuable that is missing in the latter.

ing epistemic self-reliance.[53] What's more, it looks as though the sorts of self-reliant cognitive achievements that we find valuable are frequently such that they manifest intellectual virtues that have been acquired and cultivated through a host of intellectual practices which involve reliance upon others: reliance, for instance, upon one's schoolteachers and parents, and later on, one's colleagues and interlocutors.[54] Indeed, *if* successful, this very doctoral thesis will constitute something of a cognitive achievement, and one that will to a fairly significant extent have been achieved through epistemic self-reliance, but clearly enough, the intellectual virtues whose exercise has been so essential to the completion of the thesis have been wrought over the years through all manner of relationships involving mutual epistemic dependence. Our being mutually epistemically dependent to roughly the degree that we in fact are seems to afford us a wide range of opportunities for cultivating the intellectual virtues that are required for exercising epistemic self-reliance. This leads me to one last pair of competing goods.

§4.4.1.3 Opportunities to acquire, practice, and perfect the intellectual virtues vs. freedom from intellectual obstacles and challenges

Just as there are virtues of character pertaining to moral conduct, so too there are

[53] Thanks to Ryan Byerly for encouraging me to take into account the value of epistemic self-reliance.

[54] See Pritchard's (2013) account of the role of communal epistemic practices in education in cultivating the cognitive abilities that are required for the highest-grade cognitive achievements. See also Roberts and Wood (2007: Ch. 5).

virtues pertaining to our intellectual conduct; that is, pertaining to the manner in which we go about acquiring and managing our beliefs about the world and transmitting the contents of those beliefs to others. In Chapter 3 we already encountered the thought that certain habits of intellectual character are intellectually virtuous. Jason Baehr defines an intellectual virtue as "a character trait that contributes to its possessor's personal intellectual worth on account of its involving a positive psychological orientation toward epistemic goods" (2011: 102).[55] And Roberts and Wood have suggested that "An intellectual virtue is an acquired base of excellent epistemic functioning" (2007: 95). Lists of intellectual virtues presented by Linda Zagzebski (1996), Roberts and Wood (2007), and Baehr (2011: Ch. 2) include such character traits as intellectual justice, fair-mindedness, impartiality, open-mindedness, intellectual integrity, honesty, intellectual humility, transparency, self-awareness, self-scrutiny, intellectual generosity, and intellectual courage.

Just as our having the opportunity to acquire, practice, and perfect a wide range of moral virtues requires that human life unfold against a backdrop of

[55] I noted in Chapter 3 that virtue epistemologists are sometimes presented as being divided into two camps, some of whom see virtues as merely powers or dispositions (virtue reliabilists), while others view them as person-level habits of intellectual character (virtue responsibilists). Following Greco (2000), my suggestion there was that a fully adequate virtue epistemology must acknowledge that both powers *and* habits are included in the class of intellectual virtues, and that whilst acquiring knowledge in some domains calls only for the exercise of cognitive powers, in other, more challenging domains, traits or habits of intellectual character also are called for.

struggle and adversity,[56] similarly, our having the opportunity to acquire, practice, and perfect a wide range of intellectual character virtues requires that our intellectual lives—our strivings to acquire intellectual goods—take place against a backdrop of intellectual adversity: a backdrop which includes such things as persistent disagreements over deeply important matters; withholding or suppression of information by those who occupy positions of power; uncertainty about the best means to acquire important truths about the nature of reality; temptation to deceive ourselves about our own motivations, including our motivations for believing various things; pressure to conform our opinions to a prevailing consensus; pressure to be untruthful in our dealings with others in order to gain a competitive advantage; temptation to malign a person rather than engage her arguments; and so on. What's more, the existence of such obstacles, challenges, and temptations requires that our cognitive capacities be limited in the following sorts of ways: in particular, that we are significantly dependent upon one another for acquiring a great deal of the truths that are needed for living a worthwhile life; that our perceptual experience of the world is generally insufficient to decisively settle a wide range of deeply important questions, such as those pertaining to politics, morality, metaphysics, and religion; that our opinions and plausibility judgments—especially on questions that cannot be decisively settled through empirical investigation—are prone to be influenced by the norms and consensuses

[56] This thought plays a central role in so-called "soul-making" theodicies, notable examples of which include Hick (1979), and more recently, Stump (2010). Dougherty (2014) has notably applied a similar line of thought in developing a theodicy for the suffering of non-human animals.

that prevail within the intellectual communities in which we find ourselves; that our reasoning processes are prone to be influenced by unconscious, emotional biases, which must be effortfully overcome.[57]

Having a significant range of opportunities to acquire, cultivate, and perfect the intellectual virtues seems to require that the degree to which we are subject to the various aforementioned obstacles lies somewhere in between two extremes. To illustrate what I mean, consider the following analogy. Having the opportunity to become skilled at driving a car requires that one gets some practice at driving in challenging conditions: in bad weather; in built up areas; in heavy traffic; whilst carrying noisy passengers. Part of the skill of driving is having the ability to respond appropriately to such challenges. But if these challenges are too severe—if, say, one's passengers continually try to distract one; the weather is so dreadful that visibility is almost nil; pedestrians are frequently running out into the road at random; the traffic is such that other vehicles are continually making hazardous lane changes all around—then it will be extremely hard to acquire in the first place the skills involved in being a good driver. One will not even be able to get off the ground, so to speak. On the other hand, if the challenges are too insubstantial—if, for instance, one only ever drives around an empty practice circuit inside a disused aircraft hangar—then one won't ever be called upon to respond to adverse circumstances, the skillful handling of which is involved in being a good driver. The same seems to be true of acquiring virtues of intellectual character.

[57] For more on this latter, see Kahneman (2003).

Acquiring them requires that one face intellectual challenges and obstacles that are neither too severe nor too insubstantial. And it looks plausible that the way in which humans are actually cognitively constituted is such as to present us with an array of intellectual challenges and obstacles that is not so severe that we can scarcely make any headway in acquiring virtues of intellectual character, nor is so insubstantial that we seldom encounter any situations that call for the exercise of such virtues.

Let's take the virtue of intellectual courage as an example. Baehr suggests that "intellectual courage is (at least roughly and generally) a matter of (1) pursuing an intellectual good (2) despite the fact that doing so involves a certain threat or potential harm to oneself" (2011: 169).[58] Consider a race of creatures who are cognitively much stronger than us: for instance, who are constituted in such a way that each one on their own can arrive at knowledge of almost all of the truths they desire or need to know, without relying upon testimony. Such creatures would seldom have cause to exercise intellectual courage. We humans, in virtue of our modest epistemic capabilities, fairly often encounter opportunities to risk our reputation or wellbeing in some way (even if only in a small way) in order to benefit the epistemic situation of others: for instance, by communicating to a friend some truth that is hard for her to hear; by holding firm to a view that one's peers reject and think contemptible or laughable; by engaging in a risky investigation to uncover the truth; or more dramatically, by blowing the whistle on

[58] For an extended discussion of intellectual courage, see Roberts and Wood (2007: Ch. 8).

institutional malpractice. Presumably, the extremely cognitively strong creatures I mentioned would seldom have opportunities to benefit the epistemic situations of their fellows in these ways. On the other hand, consider a race of creatures who are much cognitively weaker than us: who, for instance, have such weak perceptual faculties that they can decisively establish very little of significance by way of them, or are so prone to conform that for any one of them to think independently of their peers requires monumental effort, or who are subject to all sorts of overwhelmingly powerful, error-prone biases in their reasoning processes. Though these creatures might well face an even greater array of situations than we do in which they could benefit the epistemic situations of their peers by risking their own wellbeing, their extreme cognitive frailties will tend overwhelmingly to prevent them from usefully doing very much to alleviate the epistemic impoverishment of their fellows.[59]

Space does not permit me to go into detail about other virtues of intellectual character in order to make a similar point. But I submit that it is very plausible with respect to a range of other such virtues—intellectual humility, impartiality, self-awareness, charity, honesty, open-mindedness—that the human cognitive constitution is such as to afford us a significant range of opportunities to acquire, cultivate, and perfect those virtues that we would not have we were either much

[59] Possessing an intellectual virtue does, after all, require that one fairly frequently *succeeds* in one's exercise of that virtue. A virtue is a kind of ability, and as Greco (2010: 77) notes, having an ability with respect to some range of outcomes R in a certain range of conditions C entails attaining a high rate of success with respect to R in C across a set of relevantly close worlds.

cognitively weaker or much cognitively stronger than we in fact are.

In sum, I have argued that the actual human cognitive constitution makes for a favourable balance between various competing goods; that it clearly lies in between various undesirable extremes. In the analogies I presented in §4.4.1, in which Jill was trying to decide how much money to entrust to her teenage children to share among themselves and in which an architect was trying to decide how tall to make the ceilings and windows in the rooms of a house, I said that we should expect a value that lies clearly in between the extremes to be chosen, and that we should be surprised to find that a value lying at the extremes had been chosen. So too, we should not be surprised were God to bring about S, the state of affairs in which there exists a race of creatures with roughly our cognitive constitution.

§4.4.2 God's failure to prevent uneven distribution

Finally, I want to try to address the following worry: even supposing God has very good reasons to bring about S, would God not intervene in various ways to ensure that theistic belief (by which I mean monotheistic belief, or perhaps some specific variant of it) permeates human populations more rapidly and thoroughly than it in fact has done?

One response here would be to point out that for all anyone knows, God in fact *has* engaged in much miraculous activity so as to facilitate the spread of theistic

belief throughout the world's populations.[60] That is, for all we know, the testimonial transmission of theistic belief has been directly aided by God and would have been much less effective than it has been were it not for God's engaging in acts of local, special divine action.

But there is more else to say besides this. I strongly suspect that were we to try to sharpen up the worry here, it would end up being exactly parallel to a complaint that is sometimes raised in debates about the problem of evil: namely, that there is a minimum quantity of horrific occurrences that is consistent with God's obtaining the great goods that He seeks to realize, and that God would be unjustified in permitting any more than that minimum, and moreover, that the actual quantity of horrific occurrences in the world is plausibly greater than that minimum. The proper response to this thought, it seems to me, is one that van Inwagen offers: that there just *isn't* any non-arbitrary line to be drawn. As he puts it,

> There is no minimum number of horrors consistent with God's plan of reconciliation, for the prevention of any one particular horror could not possibly have any effect on God's plan. For any *n*, if the existence of at most *n* horrors is consistent with God's plan, the existence of at most *n*-1 horrors will be equally consistent with

[60] This parallels a response that van Inwagen offers to the worry that God, if He exists, would engage in miraculous activity so as to prevent many horrific occurrences: "God perhaps does act to prevent any number of horrors. For all we know, he reduces the number of horrors in our world to some very small fraction of what it would have been if not for his specific and local miraculous action" (2006: 104).

God's plan. (2006: 106).

By the same token, where *n* represents the number of people whom the testimonial transmission of theistic belief fails to reach (or at any rate, for whom such testimony is not successful in rendering theistic belief a "live" option), if the existence of *n* such individuals is consistent with God's realizing the goods that He wishes to realize through permitting humans to be significantly mutually epistemically dependent, then the existence of *n*-1 such individuals will be equally consistent with God's realizing those goods. There is no minimum number of such individuals, and therefore, no non-arbitrary line to be drawn.

§4.5 Conclusion

This final chapter set out to address a cluster of worries that were prompted by the picture I painted in the previous chapter, of a situation in which human beings can acquire some very basic "core" supernaturalistic knowledge in virtue of our natural cognitive equipment, whilst at the same time being subject to cultural influences concerning the specific attributes of the divine reality—influences that are often misleading. The worry, in a nutshell, was that a perfectly loving God would be very unlikely to permit such a situation. I have endeavoured to show that upon reflection, however, we should not in fact be surprised were a perfectly loving God to create a world in which human beliefs on a wide range of matters,

including religious matters, are subject to cultural pushes and pulls. The reason we should not be surprised is basically twofold. On the one hand, the core supernaturalistic knowledge, which I argued in the previous chapter is very widely available if God exists, is such that having it enables one to personally relate *de re* to God. On the other hand, the human cognitive constitution is such that a patchy distribution of specifically monotheistic beliefs is highly likely to obtain among creatures who are so constituted, and what's more, our having such a cognitive constitution permits a favourable balance between various competing goods that God is likely to want to realize. A world containing the sort of distribution of theistic belief and non-belief that we observe is, after all, among the worlds that a perfectly loving God would have good reason to create.

Bibliography

Adams, Marilyn McCord, and Robert Merrihew Adams. 1990. *The Problem of Evil*. Oxford: Oxford University Press.

Alston, William P. 1991. *Perceiving God: The Epistemology of Religious Experience*. Ithaca, NY: Cornell University Press.

Atran, Scott, and Ara Norenzayan. 2004. "Religion's Evolutionary Landscape: Counterintuition, Commitment, Compassion, Communion". *Behavioral and Brain Sciences* 27: 713-70.

Augustine. 1991. *Confessions*. Translated by Henry Chadwick. Oxford: Oxford University Press.

Baehr, Jason S. 2011. *The Inquiring Mind: On Intellectual Virtues and Virtue Epistemology*. Oxford: Oxford University Press.

Barrett, Justin L. 2004. *Why Would Anyone Believe in God?* Oxford: Altamira Press.

———. 2007. "Is the Spell Really Broken? Bio-Psychological Explanations of Religion and Theistic Belief". *Theology and Science* 5 (1): 57-72.

———. 2008. "Coding and Quantifying Counterintuitiveness in Religious Concepts: Theoretical and Methodological Reflections". *Method and Theory in the Study of Religion* 20: 308-38.

———. 2011. *Cognitive Science, Religion, and Theology: From Human Minds to Divine Minds*. West Conshohocken, PA: Templeton Press.

Barrett, Justin L., and Jonathan A. Lanman. 2008. "The Science of Religious Beliefs". *Religion* 38: 109-24.

Barrett, Justin L., and Rebekah A. Richert. 2003. "Anthropomorphism or Preparedness? Exploring Children's God Concepts". *Review of Religious*

Research 44 (3): 300-12.

Becker, Kelly. 2007. *Epistemology Modalized*. London: Routledge.

———. 2008. "Epistemic Luck and the Generality Problem". *Philosophical Studies* 139: 353-66.

Beilby, James, ed. 2002. *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument against Naturalism*. Ithaca, NY: Cornell University Press.

Benton, Matthew. Unpublished. "Epistemology Personalized".

Berger, Peter S. 1990. *The Sacred Canopy: Elements of a Sociological Theory of Religion*. Garden City, NY: Anchor Books.

Bergmann, Michael. 2006. *Justification without Awareness: A Defense of Epistemic Externalism*. Oxford: Oxford University Press.

Bloom, Paul. 2010. "Religious Belief as an Evolutionary Accident". In *The Believing Primate: Scientific, Philosophical, and Theological Reflections on the Origin of Religion*, edited by Jeffrey Schloss and Michael J. Murray. 118-27. Oxford: Oxford University Press.

Bogardus, Tomas. 2013. "The Problem of Contingency for Religious Belief". *Faith and Philosophy* 30 (4): 371-92.

Boyer, Pascal. 2001. *Religion Explained: The Human Instincts That Fashion Gods, Spirits and Ancestors*. London: Heinemann.

Brogaard, Berit, and Joe Salerno. 2007. "Why Counter-Possibles Are Not Trivial". *The Reasoner* 1 (1): 1-7.

Brueckner, Anthony, and Christopher Buford. 2013. "Becker on Epistemic Luck". *Philosophical Studies* 163: 171-75.

Buchak, Lara. 2014. "Rational Faith and Justified Belief". In *Religious Faith and Intellectual Virtue*, edited by Timothy O'Connor and Laura Frances Callahan. 49-73. Oxford: Oxford University Press.

Byerly, T. Ryan. 2014. "The Special Value of Epistemic Self-Reliance". *Ratio* 27 (1): 53-67.

Calvin, John. 1960 [1536]. *Institutes of the Christian Religion*. Translated by Ford Lewis Battles. Edited by John T. McNeill. Philadelphia, Pennsylvania: Westminister Press.

Chisholm, Roderick M. 1973. *The Problem of the Criterion*. Milwaukee, WI: Marquette University Press.

———. 1976. "Knowledge and Belief: 'De Dicto' and 'De Re'". *Philosophical Studies* 29: 1-20.

Churchland, Patricia. 1987. "Epistemology in the Age of Neuroscience". *Journal of Philosophy* 84 (10): 544-53.

Clark, Kelly James. 1990. *Return to Reason: A Critique of Enlightenment Evidentialism, and a Defense of Reason and Belief in God*. Grand Rapids, MI: Wm. B. Eerdmans Publishing Company.

———. 2010. "Reformed Epistemology and the Cognitive Science of Religion". In *Science and Religion in Dialogue*, edited by Melville Y. Stewart. 500-13. Oxford: Wiley-Blackwell.

Clark, Kelly James, and Justin L. Barrett. 2011. "Reidian Religious Epistemology and the Cognitive Science of Religion". *Journal of the American Academy of Religion* 79 (2): 1-37.

Clark, Kelly James, and Dani Rabinowitz. 2011. "Knowledge and the Objection to Religious Belief from Cognitive Science". *European Journal for the Philosophy of Religion* 2 (3): 67-81.

Clark, Kelly James, and Justin T. Winslett. 2011. "The Evolutionary Psychology of Chinese Religion: Pre-Qin High Gods as Punishers and Rewarders". *Journal of the American Academy of Religion* 79 (4): 928-60.

Clarke-Doane, Justin. 2012. "Morality and Mathematics: The Evolutionary

Challenge". *Ethics* 122 (2): 313-40.

Cordry, Benjamin S. 2008. "Divine Hiddenness and Belief *De Re*". *Religious Studies* 45: 1-19.

Craig, William Lane 1989. "No Other Name: A Middle Knowledge Perspective on the Exclusivity of Salvation through Christ". *Faith and Philosophy* 6: 172-88.

Crisp, Thomas M. 2011. "An Evolutionary Objection to the Argument from Evil". In *Evidence and Religious Belief*, edited by Kelly James Clark and Raymond J. van Arragon. 115-32. Oxford: Oxford University Press.

Cullison, Andrew. 2010. "Two Solutions to the Problem of Divine Hiddenness". *American Philosophical Quarterly* 47 (2): 119-34.

De Cruz, Helen, and Johan De Smedt. 2007. "The Role of Intuitive Ontologies in Scientific Understanding—the Case of Human Evolution". *Biology and Philosophy* 22: 351-68.

de Ridder, Jeroen. 2014. "Epistemic Dependence and Collective Scientific Knowledge". *Synthese* 191 (1): 37-53.

de Ridder, Jeroen, and René van Woudenberg. 2014. "Referring to, Believing in, and Worshipping the Same God: A Reformed View". *Faith and Philosophy* 31 (1): 46-67.

De Smedt, Johan, and Helen De Cruz. 2011. "The Cognitive Appeal of the Cosmological Argument". *Method and Theory in the Study of Religion* 23: 103-22.

Donnellan, Keith S. 1966. "Reference and Definite Descriptions". *The Philosophical Review* 75 (3): 281-304.

Dougherty, Trent. 2014. *The Problem of Animal Pain: A Theodicy for All Creatures Great and Small*. Basingstoke: Palgrave Macmillan.

Draper, Paul. 1989. "Pain and Pleasure: An Evidential Problem for Theists". *Noûs* 23 (3): 331-50.

Drees, Willem B. 1996. *Religion, Science, and Naturalism*. Cambridge: Cambridge University Press.

Dretske, Fred. 1971. "Conclusive Reasons". *Australasian Journal of Philosophy* 49 (1): 1-22.

Enoch, David. 2010. "The Epistemological Challenge to Metanormative Realism: How Best to Understand It, and How to Cope with It". *Philosophical Studies* 148: 413-38.

Evans, C. Stephen. 2010. *Natural Signs and Knowledge of God: A New Look at Theistic Arguments*. Oxford: Oxford University Press.

Feldman, Richard. 1995. "Reliability and Justification". *The Monist* 68: 159-74.

———. 2002. *Epistemology*. Upper Saddle River, NJ: Pearson.

Feuerbach, Ludwig. 1989 [1841]. *The Essence of Christianity*. Translated by George Eliot. Amherst, NY: Prometheus Books.

Finnis, John. 2001. *Natural Law and Natural Rights*. 2nd ed. Oxford: Clarendon Press.

Fitelson, Branden. 2010. "Strengthening the Case for Knowledge from Falsehood". *Analysis* 70 (4): 666-69.

Flint, Thomas P. 2009. "Divine Providence". In *The Oxford Handbook of Philosophical Theology*, edited by Michael C. Rea and Thomas P. Flint. 262-85. Oxford: Oxford University Press.

Fodor, Jerry. 2002. "Is Science Biologically Possible?". In *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument against Naturalism*, edited by James Beilby. 30-42. Ithaca, NY: Cornell University Press.

Freud, Sigmund. 1973 [1927]. *The Future of an Illusion*. Translated by W.D.

Robson-Scott. London: The Hogarth Press and the Institute of Psycho-Analysis.

Fricker, Elizabeth. 2006. "Testimony and Epistemic Autonomy". In *The Epistemology of Testimony*, edited by Jennifer Lackey and Ernest Sosa. Oxford: Oxford University Press.

Gellman, Jerome. 1993. "Religious Diversity and the Epistemic Justification of Religious Belief". *Faith and Philosophy* 10 (3): 345-64.

Gettier, Edmund. 1963. "Is Justified True Belief Knowledge?". *Analysis* 23 (6): 121-23.

Goldberg, Sanford. 2011. "The Division of Epistemic Labour". *Episteme* 8: 112-25.

Goldman, Alvin. 1967. "A Causal Theory of Knowing". *The Journal of Philosophy* 64 (12): 357-72.

———. 1976. "Discrimination and Perceptual Knowledge". *The Journal of Philosophy* 73 (20): 771-91.

———. 1979. "What Is Justified Belief?". In *Justification and Knowledge*, edited by George S. Pappas. 1-23. Dordrecht: D. Reidel.

Greco, John. 2000. "Two Kinds of Intellectual Virtue". *Philosophy and Phenomenological Research* 60 (1): 179-84.

———. 2010. *Achieving Knowledge: A Virtue-Theoretic Account of Epistemic Normativity*. Cambridge: Cambridge University Press.

Griffiths, Paul E., and John S. Wilkins. 2010. "When Do Evolutionary Explanations of Belief Debunk Belief?". In *Darwin in the 21st Century: Nature, Humanity, and God*, edited by P. R. Sloan. Notre Dame, IN: Notre Dame University Press.

Gutting, Gary. 1983. *Religious Belief and Religious Scepticism*. Notre Dame, IN: University of Notre Dame Press.

Hanson, Philip P. 1978. "Prospects for a Causal Theory of Knowledge". *Canadian Journal of Philosophy* 8 (3): 457-73.

Hardwig, John. 1991. "The Role of Trust in Knowledge". *Journal of Philosophy* 88 (12): 693-708.

Harman, Gilbert. 1973. *Thought*. Princeton, NJ: Princeton University Press.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hawthorne, John, and Dani Rabinowitz. Forthcoming. "Knowledge and False Belief". In *Gettier at Fifty*, edited by Peter D. Klein. Oxford: Oxford University Press.

Hick, John. 1979. *Evil and the God of Love*. Glasgow: Collins.

———. 1997. "The Epistemological Challenge of Religious Pluralism". *Faith and Philosophy* 14 (3): 277-86.

Hiller, Avram. 2013. "Knowledge Essentially Based Upon False Belief". *Logos and Episteme* 4 (1): 7-19.

Howard-Snyder, Daniel. 1996. "The Argument from Divine Hiddenness". *Canadian Journal of Philosophy* 26 (3): 433-53.

Hume, David. 1956 [1757]. *The Natural History of Religion*. Edited by H.E. Root. London: William Clowes and Sons.

Jackson, Frank. 1982. "Epiphenomenal Qualia". *Philosophical Quarterly* 32: 127-36.

James, William. 1979 [1896]. *The Will to Believe and Other Essays in Popular Philosophy*. Cambridge, MA: Harvard University Press.

Joyce, Richard. 2006. *The Evolution of Morality*. London: MIT Press.

Kahane, Guy. 2011. "Evolutionary Debunking Arguments". *Noûs* 45 (1): 103-25.

Kahneman, Daniel. 2003. "A Perspective on Judgment and Choice: Mapping Bounded Rationality". *American Psychologist* 58 (9): 697-720.

———. 2013. *Thinking, Fast and Slow*. New York, NY: Farrar, Strauss, and Giroux.

Kallestrup, Jesper, and Duncan Pritchard. 2012a. "Robust Virtue Epistemology and Epistemic Anti-Individualism". *Pacific Philosophical Quarterly* 93 (1): 84-103.

———. 2012b. "Virtue Epistemology and Epistemic Twin Earth". *European Journal of Philosophy* 21 (2).

Kelemen, Deborah. 2004. "Are Children 'Intuitive Theists'? Reasoning About Purpose and Design in Nature". *Psychological Science* 15 (5): 295-301.

Kelemen, Deborah, and Cara DiYanni. 2005. "Intuitions About Origins: Purpose and Intelligent Design in Children's Reasoning About Nature". *Journal of Cognition and Development* 6 (1): 3-31.

Kelemen, Deborah, and Evelyn Rosset. 2009. "The Human Function Compunction: Teleological Explanation in Adults". *Cognition* 111: 138-43.

Kitcher, Philip. 2011. "Challenges for Secularism". In *The Joy of Secularism: 11 Essays for How We Live Now*, edited by George Levine. 24-56. Princeton, NJ: Princeton University Press.

Klein, Peter D. 2008. "Useful False Beliefs". In *Epistemology: New Essays*, edited by Quentin Smith. 25-61. Oxford: Oxford University Press.

Kuhn, Thomas S. 1996. *The Structure of Scientific Revolutions*. 3rd ed. Chicago: University of Chicago Press.

Lackey, Jennifer. 2007. "Why We Don't Deserve Credit for Everything We Know". *Synthese* 158: 345-61.

Lasonen-Aarnio, Maria. 2010. "Unreasonable Knowledge". *Noûs Supplement: Philosophical Perspectives* 24 (1): 1-21.

Laudan, Larry. 1981. "A Confutation of Convergent Realism". *Philosophy of Science* 48 (1): 19-49.

Law, Stephen. 2011. "Plantinga's Belief-Cum-Desire Argument Refuted". *Religious Studies* 47: 245-56.

Leech, David, and Aku Visala. 2011. "The Cognitive Science of Religion: A Modified Theist Response". *Religious Studies* 47: 301-16.

Leftow, Brian. 1991. *Time and Eternity*. Ithaca, NY: Cornell University Press.

———. 2012. *God and Necessity*. Oxford: Oxford University Press.

Lewis, David K. 1973. "Counterfactuals and Comparative Possibility". *Journal of Philosophical Logic* 2 (4): 418-446.

———. 2001. *Counterfactuals*. Rev. ed. Malden, MA: Blackwell Publishers.

Lowe, E. J. 2002. *A Survey of Metaphysics*. Oxford: Oxford University Press.

Luzzi, Federico. 2010. "Counter-Closure". *Australasian Journal of Philosophy* 88 (4): 673-83.

Lycan, William. 2006. "On the Gettier Problem Problem". In *Epistemology Futures*, edited by Stephen Hetherington. 148-68. New York, NY: Oxford University Press.

Maitzen, Stephen. 2006. "Divine Hiddenness and the Demographics of Theism". *Religious Studies* 42: 177-91.

———. 2008. "Does Molinism Explain the Demographics of Theism?". *Religious Studies* 44 (4): 473-77.

Marsh, Jason. 2008. "Do the Demographics of Theistic Belief Disconfirm Theism? A Reply to Maitzen". *Religious Studies* 44 (4): 465-71.

Marx, Karl, and Friedrich Engels. 1964 [1843]. "Contribution to the Critique of Hegel's Philosophy of Right, Introduction". In *On Religion*. Chico, CA: Scholar's Press.

Mason, Kelby. 2010. "Debunking Arguments and the Genealogy of Religion and Morality". *Philosophy Compass* 5 (9): 770-78.

Mawson, T. J. 2005. *Belief in God: An Introduction to the Philosophy of Religion*. Oxford: Oxford University Press.

———. 2009. "Mill's Argument against Religious Knowledge". *Religious Studies* 45 (4): 417-34.

———. 2012. "The Rationality of Classical Theism and Its Demographics". In *Scientific Approaches to the Philosophy of Religion*, edited by Yujin Nagasawa. 184-204. London: Palgrave Macmillan.

McCauley, Robert. 2011. *Why Religion Is Natural and Science Is Not*. New York, NY: Oxford University Press.

McKim, Robert. 2001. *Religious Ambiguity and Religious Diversity*. Oxford: Oxford University Press.

Merricks, Trenton. 1995. "Warrant Entails Truth". *Philosophy and Phenomenological Research* 55 (4): 841-55.

Mill, John Stuart. 1991 [1859]. "On Liberty". In *The Collected Works of John Stuart Mill*, edited by J.M. Robson. Toronto, ON: Toronto University Press.

Moser, Paul K. 2004. "Divine Hiddenness Does Not Justify Atheism". In *Contemporary Debates in the Philosophy of Religion*, edited by Michael L. Peterson and Raymond J. Van Arragon. 42-54. Oxford: Blackwell.

Mumford, Stephen. 2009. "Causal Powers and Capacities". In *The Oxford Handbook of Causation*, edited by Peter Menzies, Christopher Hitchcock and Helen Beebee. 265-78. Oxford: Oxford University Press.

Murray, Michael J. 1993. "Coercion and the Hiddenness of God". *American Philosophical Quarterly* 30 (1): 27-38.

———. 2007. "Four Arguments That the Cognitive Psychology of Religion Undermines the Justification of Religious Belief". In *The Evolution of Religion: Studies, Theories, and Critiques*, edited by Joseph Bulbulia et al. 393-98. Santa Margarita, CA: Collins Foundation Press.

———. 2008. *Nature Red in Tooth and Claw: Theism and the Problem of Animal Suffering*. Oxford: Oxford University Press.

———. 2010. "Scientific Explanations of Religion and the Justification of Religious Belief". In *The Believing Primate: Scientific, Philosophical, and Theological Reflections on the Origin of Religion*, edited by Jeffrey Schloss and Michael J. Murray. 168-78. Oxford: Oxford University Press.

Murray, Michael J., and Andrew Goldberg. 2010. "Evolutionary Accounts of Religion: Explaining and Explaining Away". In *The Believing Primate: Scientific, Philosophical, and Theological Reflections on the Origin of Religion*, edited by Jeffrey Schloss and Michael J. Murray. 179-99. Oxford: Oxford University Press.

Nolan, Daniel. 1997. "Impossible Worlds: A Modest Approach". *Notre Dame Journal of Formal Logic* 38 (4): 535-72.

Nozick, Robert. 1981. *Philosophical Explanations*. Cambridge, MA: Harvard University Press.

Pasnau, Robert. 2013. "Epistemology Idealized". *Mind* 122 (488): 987-1021.

Pereboom, Derk. 2011. "Theological Determinism and Divine Providence". In *Molinism: The Contemporary Debate*, edited by Ken Perszyk. 262-79. Oxford: Oxford University Press.

Perszyk, Ken. 2013. "Recent Work on Molinism". *Philosophy Compass* 8 (8): 755-70.

Philipse, Herman. 2012. *God in the Age of Science?: A Critique of Religious Reason*. Oxford: Oxford University Press.

Plantinga, Alvin. 1974a. *God, Freedom, and Evil*. Grand Rapids, MI: William B. Eerdmans Publishing Company.

———. 1974b. *The Nature of Necessity*. Oxford: Clarendon Press.

———. 1983a. "On Existentialism". *Philosophical Studies* 44 (1): 1-20.

———. 1983b. "Reason and Belief in God". In *Faith and Rationality: Reason and Belief in God*, edited by Alvin Plantinga and Nicholas Wolterstorff. 16-93. Notre Dame, IN: University of Notre Dame Press.

———. 1993a. *Warrant and Proper Function*. Oxford: Oxford University Press.

———. 1993b. *Warrant: The Current Debate*. Oxford: Oxford University Press.

———. 1995. "Pluralism: A Defence of Religious Exclusivism". In *The Rationality of Belief and the Plurality of Faith: Essays in Honor of William P. Alston* edited by Thomas D. Senor. 191-215. London: Cornell University Press.

———. 1997. "Ad Hick". *Faith and Philosophy* 14 (3): 295-98.

———. 2000. *Warranted Christian Belief*. Oxford: Oxford University Press.

———. 2002. "Introduction: The Evolutionary Argument against Naturalism". In *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument against Naturalism*, edited by James Beilby. 1-12. Ithaca, NY: Cornell University Press.

———. 2011. *Where the Conflict Really Lies: Science, Religion, and Naturalism*. New York, NY: Oxford University Press.

Pollock, John L. 1986. *Contemporary Theories of Knowledge*. Totowa, NJ: Rowman & Littlefield.

Poston, Ted, and Trent Dougherty. 2007. "Divine Hiddenness and the Nature of Belief". *Religious Studies* 43 (2): 183-98.

Pritchard, Duncan. 2005. *Epistemic Luck*. Oxford: Oxford University Press.

———. 2009. "Safety-Based Epistemology: Wither Now?". *Journal of Philosophical Research* 34: 33-45.

———. 2012. "Anti-Luck Virtue Epistemology". *Journal of Philosophy* 109 (3): 247-79.

———. 2013. "Epistemic Virtue and the Epistemology of Education". *Journal of*

Philosophy of Education 47 (2): 236-47.

———. 2014. "Knowledge Cannot Be Lucky". In *Contemporary Debates in Epistemology*, edited by Matthias Steup, John Turri and Ernest Sosa. 152-63. Oxford: Wiley Blackwell.

Pritchard, Duncan, Alan Millar, and Adrian Haddock. 2010. *The Nature and Value of Knowledge: Three Investigations*. Oxford: Oxford University Press.

Quine, W.V. 1969. *Ontological Relativity, and Other Essays*. New York, NY: Columbia University Press.

Rachels, James. 1971. "God and Human Attitudes". *Religious Studies* 7 (4): 325-37.

Rahner, Karl. 1976. *Theological Investigations*. Vol. 14. London: Darton, Longman & Todd.

Ramsey, William. 2002. "Naturalism Defended". In *Naturalism Defeated? Essays on Plantinga's Evolutionary Argument against Naturalism*, edited by James Beilby. 15-29. Ithaca, NY: Cornell University Press.

Rea, Michael C. 2002. *World without Design: The Ontological Consequences of Naturalism*. Oxford: Oxford University Press.

Reid, Thomas. 1997 [1764]. *An Inquiry into the Human Mind on the Principles of Common Sense*. Edited by Derek R. Brookes. Edinburgh: Edinburgh University Press.

Roberts, Robert C. 2009. "Emotional Consciousness and Personal Relationships". *Emotion Review* 1: 281-88.

Roberts, Robert C., and W. Jay Wood. 2007. *Intellectual Virtues: An Essay in Regulative Epistemology*. Oxford: Oxford University Press.

Rowe, William L. 1979. "The Problem of Evil and Some Varieties of Atheism". *American Philosophical Quarterly* 16 (4): 335-41.

Ruse, Michael. 1986. *Taking Darwin Seriously*. Oxford: Blackwell.

———. 2006. "Is Darwinian Metaethics Possible (and If It Is, Is It Well Taken?)". In *Evolutionary Ethics and Contemporary Biology*, edited by Giovanni Boniolo and Gabriele De Anna. 13-26. Cambridge: Cambridge University Press.

Russell, Bertrand. 1948. *Human Knowledge: Its Scope and Its Limits*. London: Allen & Unwin.

Schellenberg, J.L. 1993. *Divine Hiddenness and Human Reason*. Ithaca, NY: Cornell University Press.

———. 2002. "What the Hiddenness of God Reveals: A Collaborative Discussion". In *Divine Hiddenness: New Essays*, edited by Daniel Howard-Snyder and Paul K. Moser. 33-61. Cambridge: Cambridge University Press.

———. 2007. *The Wisdom to Doubt: A Justification of Religious Skepticism*. Ithaca, NY: Cornell University Press.

Searle, John R. 1995. *The Construction of Social Reality*. London: Penguin.

Shope, Robert K. 2005. "Conditions and Analyses of Knowing". In *The Oxford Handbook of Epistemology*, edited by Paul K. Moser. 25-70. Oxford: Oxford University Press.

Simpson, Thomas W. 2012. "What Is Trust?". *Pacific Philosophical Quarterly* 93: 550-69.

Sosa, Ernest. 1991. *Knowledge in Perspective*. Cambridge: Cambridge University Press.

———. 1999. "How to Defeat Opposition to Moore". *Noûs Supplement: Philosophical Perspectives* 33 (13): 141-53.

———. 2007. *A Virtue Epistemology*. Vol. 1, *Apt Belief and Reflective Knowledge*. Oxford: Oxford University Press.

———. 2009. *Reflective Knowledge*. Vol. 2, *Apt Belief and Reflective Knowledge*. Oxford:

Oxford University Press.

———. 2011. *Knowing Full Well*. Princeton, NJ: Princeton University Press.

Speaks, Jeff. 2012. "On Possibly Nonexistent Propositions". *Philosophy and Phenomenological Research* 85 (3): 528-62.

Stalnaker, Robert C. 1975. "Causation and Conditionals". In *Oxford Readings in Philosophy*, edited by Ernest Sosa. 165-79. New York, NY: Oxford University Press.

Steup, Matthias, and Jonathan Jenkins Ichikawa. 2012. "The Analysis of Knowledge". In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Available at: http://plato.stanford.edu/entries/knowledge-analysis/

Street, Sharon. 2006. "A Darwinian Dilemma for Realist Theories of Value". *Philosophical Studies* 127 (1): 109-66.

Stump, Eleonore. 2010. *Wandering in Darkness: Narrative and the Problem of Suffering*. Oxford: Oxford University Press.

Swinburne, Richard. 1989. *Responsibility and Atonement*. Oxford: Oxford University Press.

———. 1997. *Simplicity as Evidence of Truth*, *The Aquinas Lecture*. Milwaukee, WI: Marquette University Press.

———. 1998. *Providence and the Problem of Evil*. Oxford: Oxford University Press.

———. 2001. *Epistemic Justification*. Oxford: Clarendon Press.

———. 2004. *The Existence of God*. 2nd ed. Oxford: Oxford University Press.

Thurow, Joshua. 2013. "Does Cognitive Science of Religion Show Belief in God to Be Irrational? The Epistemic Consequences of the Cognitive Science of Religion". *International Journal for Philosophy of Religion* 74 (1): 77-98.

Turri, John. 2011. "Manifest Failure: The Gettier Problem Solved". *Philosophers' Imprint* 11 (8): 1-11.

van Inwagen, Peter. 1988. "The Place of Chance in a World Sustained by God". In *Divine and Human Action: Essays in the Metaphysics of Theism*, edited by Thomas V. Morris. 211-35. Ithica, NY: Cornell University Press.

———. 1991. "The Problem of Evil, the Problem of Air, and the Problem of Silence". *Noûs Supplement: Philosophical Perspectives* 5: 135-65.

———. 1995. "Non Est Hick". In *The Rationality of Belief and the Plurality of Faith: Essays in Honor of William P. Alston* edited by Thomas D. Senor. 216-41. London: Cornell University Press.

———. 2006. *The Problem of Evil: The Gifford Lectures Delivered in the University of St Andrews in 2003*. Oxford: Oxford University Press.

Visala, Aku. 2011. *Naturalism, Theism and the Cognitive Study of Religion: Religion Explained?* Farnham: Ashgate.

Warfield, Ted A. 2005. "Knowledge from Falsehood". *Noûs Supplement: Philosophical Perspectives* 19 (1): 405–16.

White, Roger. 2010. "You Just Believe That Because...". *Noûs Supplement: Philosophical Perspectives* 24 (1): 573-615.

Whitehouse, Harvey. 2008. "Cognitive Evolution and Religion: Cognition and Religious Evolution". In *The Evolution of Religion: Studies, Theories, and Critiques*, edited by Joseph Bulbulia, Richard Sosis, Erica Harris, Russell Genet, Cheryl Genet and Karen Wyman. 19-29. Santa Margarita, CA: Collins Foundation Press.

Wielenberg, Erik J. 2010. "On the Evolutionary Debunking of Morality". *Ethics* 120 (3): 441-64.

Wierenga, Edward. 1998. "Theism and Counterpossibles". *Philosophical Studies* 89: 87-103.

Wilkins, John S., and Paul E. Griffiths. 2013. "Evolutionary Debunking Arguments in Three Domains: Fact, Value, and Religion". In *A New Science*

of Religion, edited by Gregory W. Dawes and James Maclaurin. London: Routledge.

Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.

———. 2009. "Reply to John Hawthorne and Maria Lasonen-Aarnio". In *Williamson on Knowledge*, edited by Patrick Greenough and Duncan Pritchard. 313-29. Oxford: Oxford University Press.

———. 2014. "Knowledge First". In *Contemporary Debates in Epistemology*, edited by Matthias Steup, John Turri and Ernest Sosa. 1-10. Oxford: Wiley Blackwell.

Wolterstorff, Nicholas. 1983. "Can Belief in God Be Rational If It Has No Foundations?". In *Faith and Rationality: Reason and Belief in God*, edited by Alvin Plantinga and Nicholas Wolterstorff. 135-86. Notre Dame, IN: University of Notre Dame Press.

———. 2009. "How Philosophical Theology Became Possible within the Analytic Tradition of Philosophy". In *Analytic Theology: New Essays in the Philosophy of Theology*, edited by Oliver D. Crisp and Michael C. Rea. 155-68. Oxford: Oxford University Press.

Zagzebski, Linda. 1994. "The Inescapability of Gettier Problems". *Philosophical Quarterly* 44 (174): 65-73.

———. 1996. *Virtues of the Mind: An Inquiry into the Nature of Virtue and the Ethical Foundations of Knowledge*. New York, NY: Cambridge University Press.

Zimmerman, Dean. 2009. "Yet Another Anti-Molinist Argument". In *Metaphysics and the Good: Themes from the Philosophy of Robert M. Adams*, edited by Samuel Newlands and Larry M. Jorgensen. 33-94. Oxford: Oxford University Press.