

Oxford Studies in Philosophy of Religion

Volume V

Jonathan L. Kvanvig, Editor

CLARENDON PRESS • OXFORD

2012

CONTENTS

1	Divine action beyond the personal omniGod	1
	<i>John Bishop and Ken Perszyk</i>	
2	God Knows the Future by Ordering the Times	18
	<i>Ryan Byerly</i>	
3	Evil and The Problem of Anomaly	40
	<i>Trent Dougherty and Alexander R. Pruss</i>	
4	Anselm on Freedom and Grace	71
	<i>James A. Gibson</i>	
5	God and Moral Perfection	98
	<i>Shawn Graves</i>	
6	The Father of Lies?	117
	<i>Hud Hudson</i>	
7	God Acts in the Quantum World	133
	<i>Bradley Monton</i>	
8	No (New) Troubles with Ockhamism	147
	<i>Garrett Pendergraft and D. Justin Coates</i>	
9	Social Evil*	166
	<i>Ted Poston</i>	
10	The Arbitrariness of the Primal Sin	186
	<i>Timpe</i>	

DIVINE ACTION BEYOND THE PERSONAL OMNIGOD

John Bishop and Ken Perszyk

The dominant conception of God—within analytical philosophy of religion, at least—has been that of a person, without a body, who is creator and sustainer of all else that exists and who is supremely powerful, knowing and good (omnipotent, omniscient and omnibenevolent).¹ We call this the conception of God as *e the personal omniGod*.² In our view, there is a version of the Argument from Evil that shows that God cannot be the personal omniGod, given the ethical commitments theists typically have for assessing the moral perfection of persons. A person who first causes or sustains horrific suffering and then—as sophisticated theodicies maintain—brings the participants in horrors into eternal relationship with himself could not *overall* have the most perfectly loving kind of relationship with them. Or so we have argued.³ Others reject an omnipotent personal God as a cosmic tyrant,⁴ or on the grounds that a God who is ‘a’ person, however supreme, is improperly anthropomorphic,⁵ or because the idea of a supernatural being supposedly introduces idolatry.⁶

Such reasons for dissatisfaction with the personal omniGod conception may not be decisive, and some will think they are able to defeat them. Yet we think

¹See, for example, Swinburne (1977, pp. 1 and 101), Alston (1989, p. 198) and Plantinga (2000, p. vii).

²There are different personal omniGod conceptions, depending, for example, on how one understands the omniproperties or their scope.

³Bishop and Perszyk (2011). We think that this version of the Argument from Evil may also succeed against moderate revisions of the omniGod conception—for example, those that hold that God’s power falls short of omnipotence Nagasawa (2008).

⁴Lincoln said, “As I would not be a slave, so I would not be a master.” Is Lincoln to be considered nobler than God? Would God be a master, in the sense some have given this term, a cosmic sovereign? Tyrannical people may worship a tyrant God, but why should the rest of us do so’ (Hartshorne 1984, pp. 58–59).

Compare Leibniz: ‘Our end is to banish from men the false ideas that represent God to them as an absolute prince employing a despotic power, unfitted to be loved and unworthy of being loved. These notions are the more evil in relation to God inasmuch as the essence of piety is not only to fear him but also to love him above all things: and that cannot come about unless there be knowledge of his perfections capable of arousing the love which he deserves, and which makes the felicity of those that love him’ (Leibniz 1951, Part I, section 6). (Thanks to Jeremy Reid for drawing our attention to this quotation.)

⁵Swinburne (1977, p. 101) attributes this sort of view to some (unnamed) Protestant theologians. Levine (1994, pp. 148–149 and 194) attributes it to pantheists.

⁶As recently claimed by Mark Johnston: ‘the very ideas of religion as essentially supernaturalist, and of God as essentially a supernatural being, are idolatrous conceptions’ (Johnston 2009, p. 39). Also see, for example, Phillips (1965, Chap. 6).

that these reasons are weighty enough to warrant investigation of potential alternatives. Those who disagree may wish to consider the matter hypothetically: if there were weighty reasons for regarding the personal omniGod conception of the divine as unsatisfactory, might there be an adequate alternative conception?

There *are* alternatives—including one, at least, that belongs to traditional Christian orthodoxy. We refer to classical theism, for which the divine is atemporal, immutable, impassible and simple (not composite), and so arguably not ‘a’ being or ‘a’ person (even if Trinitarianism requires accepting personhood as belonging to the Godhead analogically).⁷ And of course there has been a range of pantheistic and panentheistic alternatives in the Indian, Chinese and Western traditions both prior and subsequent to the development of Christianity.⁸ Amongst contemporary philosophers, a number of alternatives to the personal omniGod conception have been proposed, including John Leslie’s (1979, 1989) extreme axiarchism, Peter Forrest’s (2007) developmental theism, and Mark Johnston’s Heideggerian notion of God as ‘the outpouring of Existence Itself by way of its exemplification in ordinary existents for the sake of the self-disclosure of Existence Itself.’ (Johnston 2009, p. 116).

Concern about the adequacy of the personal omniGod conception may not, then, simply be dismissed as clearly misconceived or peripheral. Nevertheless, the personal omniGod conception might still be the best we can do. After all, given the mysteriousness of the divine, it is hardly surprising if our best effort at a positive conception has limitations and inadequacies. The question is, then, whether the claimed inadequacies of the personal omniGod conception indicate that it represents a false direction in the search for understanding, or, rather, merely arise from its falling short of full comprehensibility, as even our best positive conception must do.⁹

To answer this question, we need a better sense of what the potential alternatives are. In this paper, we undertake some comparison of personal-omniGod-theism with a certain kind of non-personal alternative (which we call the ‘euteleological’ conception) by focussing on the topic of divine action.

Divine action is fundamental to the conception of the theistic God. The major Western religious traditions understand that God acts—both in creating

⁷See, for example, *Summa Theologica*, Part I, Question 3, Article 6, ad 2, and *Summa Contra Gentiles*, Book I, Chap. 25. Also see Burrell (1993, p. 70).

⁸See, for example, Levine (1994, 2011), Cooper (2006).

⁹We here set aside the tradition of apophaticism, or negative theology—according to which we may grasp only what God is not. Pure apophaticism seems unsatisfactory—both metaphysically (from the perspective of *fides quaerens intellectum*) and for our religious psychology. That God relates ‘himself’ to us is at the core of theistic religion, and this is trivialised if all we can say is that we are related to that-we-know-not-what. Besides, our understanding requires some positive conception of the divine if we are to resist the accusation that there is—and could be—nothing to fill ‘the God-role’ in our system of religious beliefs and practices, our theistic ‘form of life’. It may well be true, however, that we can be fully comprehending only in our understanding of what God is not, with our positive conceptions of God necessarily falling short of complete comprehension.

and sustaining the Universe, and in bringing it to fulfilment, which involves God performing ‘saving acts’ within human history. No conception of God could be adequate to these religious traditions without accommodating these ideas of divine agency.¹⁰

I

On the face of it, personal omniGod theism is well placed to accommodate divine action. That the divine acts may not, perhaps, *entail* that the divine is a person, but taking God to be a person does provide a natural account of how God can be an agent—and a personal conception of God is reinforced in religious traditions that emphasise human personal relationship with God. Indeed, any departure from God as personal agent may appear to rule out the resultant conception of the divine. But appearances here are deceptive. In this Section we argue that the standard picture of divine action is far from straightforward.

Personal omniGod theism takes God to be essentially a supernatural and immaterial person. That status is ensured by the doctrine of God as creator *ex nihilo*, which introduces a radical ontological dualism between God and the natural, material, universe. This doctrine guarantees that God cannot be identified as an item within the universe—certainly a desideratum for any adequate conception of God.¹¹

But the notion of a purely immaterial personal agent is a strange one.¹² We experience personal agency only as embodied. Agency emerges within the physical—how, then, could it somehow precede it? The standard reply to this question is so familiar that we have become immune to its strangeness—namely that understanding our own agency requires that the acting self is, in itself, immaterial, though intimately linked (and not merely as a pilot in a ship) with the body in and through which its physical actions are realised. Modulo *that* philosophical understanding of human agency, divine agency is comprehensible

¹⁰How should we go about assessing candidate conceptions of the divine for adequacy? We need to grasp the role that belief in God plays in the ‘form of life’ of the theistic religious traditions. An essentially anthropological, Wittgensteinian, appreciation of the lived context of theistic belief and practice is thus necessary to philosophical discussion of how, substantively, to conceive of the divine: only such an appreciation can set the proper constraints on what ‘God’ can mean (see Bishop (1998)).

¹¹All we mean here by Gods being supernatural is God’s ontological distinctness from the natural universe and anything that belongs to it. Supernaturalness in this sense belongs uniquely to God: angels, if they exist, will be just as much part of the created order as suns, planets, plants, animals and humans. Angels may be supernatural in the different sense that their existence and activity is essentially beyond the grasp of natural scientific explanation and understanding. Care is therefore needed to distinguish ‘supernatural1’ = ‘beyond the (natural) created order’ from ‘supernatural2’ = ‘in principle outside the scope of natural scientific understanding’. That something is supernatural1 entails that it is supernatural2, but not conversely.

¹²So too is the notion of an atemporal (timeless) personal agent, though the majority of personal omniGod theorists these days reject this way of understanding Gods eternity. Helm (1980) contends that arguments against divine timelessness equally cast doubt on arguments against divine spacelessness: arguments in favour of Gods being in time are equally arguments for God’s being in space.

enough as the agency of a self that has and needs no material body. God says ‘let it be’ (however one does that immaterially), and it is. What is thus caused, not by antecedent events in accordance with natural law, but by God as agent: God’s originating and sustaining all that exists is thus a matter of spontaneous divine *agent-causation*, where agent-causation is the name for a causal relation whose first term (‘the cause’) is essentially an agent.¹³

Understanding human agency in terms of agent-causation is, however, controversial. Our concept of ourselves as personal agents plausibly does employ a primitive *concept* of agent-causation. But it is controversial whether agent-causation is *ontologically* irreducible, and, if it is, whether a ‘naturalist’ ontology can accommodate basic agent-causal relations. ‘Causal theorists’ of action, for example, hold that actions belong to the natural physical causal order through *being constituted by* event-causal sequences of a particular type. According to causalists, psychological events bearing the right kind of event-causal relation to the behaviour intrinsic to a given intentional action provide the sufficient and necessary conditions for the occurrence of the agent-causation the action involves. I raise my arm, for example, when and only when my having the reasons I have for doing so causes—in the right, ‘non-deviant’, way—my arm to go up. (That is, of course, just an *ontological* analysis of what it is for me to raise my arm; it should not be read as a conceptual analysis of what it *means* for me to raise my arm.)¹⁴

The analogical basis for understanding divine agency (on the personal omniGod model) is thus not human agency as such, but a metaphysically controversial, ‘agent-causationist’, understanding of human agency as resting on ontologically irreducible agent-causal relations. Hartshorne (1984, p. 58) objects that creation *ex nihilo* is ‘the phenomenon or supposed phenomenon of magic’, and adds ‘what [this] comes to is that for the creation-out-of-nothing idea there was no *noncontroversial* analogous phenomenon whatsoever.’ We think this hits the mark. Certainly, God’s creation of the Universe *ex nihilo* can be understood as agent-causation by a (supreme) immaterial agent, though (traditionally) it cannot be understood as any kind of change given that there is no underlying substrate. And—as William Alston takes pains to argue—there is nothing intrinsic to our concept of basic action which requires that what is brought about be a movement of the agent’s body: ‘we could think of the coming into being of light or of the parting of the sea of reeds as directly under God’s voluntary control’ (Alston 1989, p. 61). But this is little comfort when it is contestable whether

¹³We here use the term ‘agent-causation’ in a relatively naive way—and, in particular, without Roderick Chisholm’s (1966) implication that whatever is agent-caused must be outside the ordinary chain of event-causality. (Chisholm’s notion of agent-causation thus builds something contestable into the terminology needed to mark the important fact that our *concept* of personal, potentially responsible, agency operates with a different conception of causing from the Humean notion of ‘event-causation’ in which causes are events or states of affairs—different because the cause is an agent. ‘Agent-causation’ seems just the term to mark this difference.)

¹⁴The classic defence of this Causal Theory of Action is Davidson (1963). For further discussion of the commitments of such a Theory, see Bishop (1989, Chapter 3).

our own basic action must be understood as realised in irreducible relations of agent-causation. It is little comfort when it is arguable that this understanding ‘supernaturalises’ finite agents, depicting them as ‘operating on’ the natural causal order from outside it, just as the supernatural personal omniGod is held to do.¹⁵ Making ontological sense of ‘the creation-out-of-nothing idea’ by appeal to agent-causation is making sense of it on the basis of a contested theory of our own action, and one that is doubtfully coherent within a naturalist, properly ‘creaturely’, understanding of ourselves. The intended analogue is, as Hartshorne says, by no means a ‘*noncontroversial* analogous phenomenon’. Indeed, the contestable libertarianism that sees the *finite* agent as an irreducible agent-cause arguably results from importing into the finite case *a priori* commitment to immaterial divine agent-causation.¹⁶ In which case, theists offer no *independent* validation for their account of divine agency by applying an agent-causationist theory of finite personal agency.¹⁷

A further—and more radical—concern with understanding Creation *ex nihilo* as the action of a supernatural person arises from generalising the requirement that, whatever God is, ‘he’ not be identical with any item in the world. This desideratum is elementary; its generalisation more subtle, but arguably as compelling. As David Burrell puts it in explicating the Thomist notion of divine simplicity, ‘to picture God as an additional being over against or parallel to the universe itself will be to treat God similarly to objects within the universe, related to the universe itself as objects within the world are related to each other’ (Burrell 1998, p. 72).¹⁸ Personal-omniGod-theists may suppose they capture the uniqueness of the divine in God’s *supernaturalness*. Yet, as Burrell makes clear, a uniquely supernatural person still shares something with finite persons, since God’s—agent-causal—relation to the universe is just the same type of relation as finite agent-causes (supposedly) have to the events intrinsic to their actions. The personal omniGod conception, then, arguably fails to capture the fullness

¹⁵Some defenders of irreducible agent-causation make heroic attempts to resist this supernaturalising dualism, with its notorious interactionist difficulties, but it is doubtful whether they succeed. A notable attempt of this kind is the work of Timothy O’Connor in his O’Connor (2000). For a critical assessment of O’Connor’s defence of a naturalist agent-causationism, see Bishop (2003).

¹⁶Compare, for example, Flint (1998, pp. 29–30) third line of argument for preferring libertarianism to compatibilism.

¹⁷Hartshorne says that the concept of divine agent-causation (‘let there be light’ and there was light) is ‘a human concept, or supposed concept, *with no basis in well-attested human experience*.’ (Hartshorne 1984, p. 58, our emphasis) But that claim may seem mistaken: surely, it *is* well-attested that we think of our agency as a matter of *ourselves* bringing things about? What is questionable is *whether agent-causation is ontologically irreducible*: our experience of agency cannot, as such, assure us that our bringing things about is ontologically foundational. So Hartshorne’s main point is correct: the basis for the analogy is by no means ‘noncontroversial’.

¹⁸Burrell often laments what he thinks is prevalent in contemporary discussions of divine and human freedom, namely, the idea that God and humans are involved in a zero-sum game, where a gain for one is at the expense of the other(s). See, for example, his Burrell (1993, p. 2).

of divine uniqueness. So long as God counts as an item—albeit highly exalted—God is still one item amongst many, and that is inconsistent with God’s having the ultimate status ‘he’ must have to be God.¹⁹

II

We think it especially interesting to focus on alternative conceptions that give up the idea that God is ‘a’ person.²⁰ ‘Non-personal’ conceptions of God are, of course, only non-personal at the level of *understanding*. There may be nothing improper—even perhaps something inevitable—in the religious psychology that treats God as a person, as ‘Father’, ‘the Eternal Thou’, and so forth. But the ways theists find it *psychologically* useful or compelling to relate to God need not define their properly reflective understanding of who or what God is. Advocates of non-personal conceptions of God, then, have no need to challenge the aptness of relating to God as a person in prayer and worship.

There are several approaches to a non-personal conception of the divine. Our attention here will be focussed on considering the prospects for conceiving of God as the Universe’s End or Goal, and as its Source, not in the sense of being its ultimate producer, but in the different sense that the Universe’s actual existence is explicable only in the light of its End.²¹

Return again to God as personal Creator *ex nihilo*, and the idea that this concept may be understood by appeal to divine agent-causation. Agent-causation is a kind of productive, efficient, causation (agents *produce*, *effect*, or *bring about* the events intrinsic to their actions). But how much explanatory force does the claim that the Universe has a supernatural person as its productive cause possess? Not much! That claim merely asserts that *there is* a certain sort of explanation of the Universe’s existence: it does not say *what the explanation is*. Substantive explanatory content requires a claim about *the purpose for which* God creates the Universe, or, equivalently, a description of the intentional action God performs in creating.

¹⁹Alston (1994, pp. 52–53) cites Schleiermacher and Tillich as advocates of this sort of argument. Alston himself is particularly keen to reject it, mainly on the grounds that understanding God as a personal agent does not bring God down to our level in the sense of making God just another item in the world. But arguably this reply does not get to the heart of the complaint. Of course the tradition never thinks of God as just another item in the world, but personal omniGod theists do understand God as a particular being, and one might legitimately wonder whether this makes ‘him’ sufficiently God-like (regardless of the uniquely distinctive characteristics ‘he’ possesses).

²⁰We do not wish to imply, of course, that there is no interest in alternatives that retain God as a personal agent, but adjust the omniGod conception in one way or another—for example, by holding that God’s power falls short of omnipotence, or by understanding God as evolving within creation, as emptying himself in the process of creation, and so on. If rejecting God’s personhood turns out to be a false direction, then those with misgivings about omniGod theism will need to attend to just these sorts of alternatives.

²¹It is, of course, entirely orthodox for theists to hold that God is ‘our End’ and well as ‘our Source’. The suggestion that a notion of the Divine as the ultimate *telos* of the Universe might yield a ‘naturalist’ and non-personal conception of God is canvassed in Bishop (2007) and Bishop (2009), and our aim here is to examine it more closely.

So, what is (the supposedly personal) God's aim in creating? Answer: the supreme good. But what is that? One answer—drawn from Christian claims to revelation—is that it is the obtaining of perfectly good relationships, between creatures and God, and amongst creatures themselves. Such a goal seems fit to serve as a 'most final end': something for the sake of which other things may be sought, but itself able to be sought only for its own sake.²² The name for this goal is 'love', though much needs to be said about just what this conception of love involves.²³ In creating, the (supposedly personal) God performs the intentional action of bringing love into existence—or, if (as for Trinitarians) love belongs to God's nature eternally, God is 'spreading the love', sharing amongst an indefinitely wide community of diverse beings the fulfilment of perfectly good relationship.

There is, then, no substantive explanation for the Universe's existence (on the assumption that it is produced by a supernatural personal Agent) until that Agent's goal in creating is specified. This follows, of course, from a general feature of intentional explanation: what carries explanatory force is the making intelligible of the behaviour of agents (in the light of their aims and their view of the world, and on the assumption of their being at least minimally practically rational). Though there is no explanation unless the agent's having certain intentions and beliefs causes the behaviour they make reasonable, it is this 'making reasonable' relation that conveys understanding.

Might we, then, have just as much explanation of the Universe's existence if we posit love as the Universe's goal, but omit the personal Agent whose goal it is? Positing a supremely good *telos* for the Universe may be a way of making its existence overall intelligible—but it will no doubt be objected that this does not *explain* the Universe, unless there is a personal Agent who *causally connects* the Universe's *telos* with its existence. Reference to its *telos* may indeed be required in explaining the Universe's existence, but—so one might claim—appeal to an ultimate *efficient* cause is needed as well. Classical support might be cited here: Aquinas holds that God is *both* the final *and* the efficient cause of the Universe.²⁴

In reply, we answer that the doctrine of creation *ex nihilo* may be open to a more radical interpretation than is familiar. On the standard view, creation *ex nihilo* does not involve the Universe's coming into existence from *absolutely* nothing, since the personal omniGod eternally exists, and thus, as its producer, has ontological (if not temporal) priority over the Universe as product. But perhaps the claim that the Universe has the status of a creation *ex nihilo* may be understood as more than the negative claim that the Universe exists other than

²²This is the role of what Aristotle calls to teliotaton, and which he thought filled by *eudaimonia* (*Nicomachean Ethics*, Book I)

²³Love in this sense does not imply any lack, for example; and, though its existence is an ultimate end, it is not, so to speak, a 'completable' end—love, in this sense, does not admit of there being *enough* love; the project of bringing about love is not a project from which one could move on, thinking 'Right! That's done, now what's next?'

²⁴Summa Theologica, Part I, Question 44, Articles 1 and 4.

by the ordering or refashioning of something pre-existing?²⁵ Perhaps this claim may further imply the lack of any temporally or ontologically prior productive cause *altogether*? It may be that *nothing*, not even a supernatural supreme Person, is the producer of the Universe as a whole! Yet, supposing that to be so, the theist will not conclude—as the atheist may—that the Universe’s existence is therefore just ‘brute’, inexplicable, fact. Though not explicable in terms of an overall productive cause, the Universe’s existence may still be rendered explicable through a *euteleological explanation* of its existence. The Universe, albeit *ex nihilo*, will then count for the theist as a creation in the sense that there is an ultimate supremely good purpose (*telos*) for the sake of which it exists—and a Christian version of this euteleological explanation takes that ultimate good to be (agapeistic) Love, perfectly good relationship.

Such an explanation of the Universe’s existence will, admittedly, count as a unique form of teleological explanation. A euteleological explanation of the Universe’s overall existence will be a form of teleological explanation that is not in principle reducible to a causal explanation.²⁶ But *any* teleological explanation of existence will have unique features—even though its explanatory force will no doubt depend on its bearing important analogies to familiar forms of explanation.²⁷ What we here want to consider, then, is *another option* for an account of the unique teleological explicability that the theist believes the Universe has. And this euteleological option may have advantages over the standard personal omniGod view.

III

If the Universe’s existence as a divine creation is understood euteleologically, without reference to a supernatural Person as its overall producer, what conception of God will then ensue? God is Creator, ultimate Source of all that is. Now, since, on the euteleological account, what ultimately explains why the Universe exists is that it has the Supreme Good (Love) as its *telos*, it looks as if, on that

²⁵In other words, in Aristotelian terms, creation *ex nihilo* may be more than creation without any *material* cause.

²⁶There are two standard ways of reducing teleological explanations to causal explanations. The first rests on the Causal Theory of Action (already mentioned, see note 14 above), according to which teleological explanations of an agent’s action explain the agent’s behaviour as caused (‘in the right sort of way’) by those psychological states that constitute his reasons for acting. The second reductive pattern applies to teleological explanations in biology, where a ‘selected effect’ or ‘etiologial’ account can be given, according to which an organ’s or organism’s proper functions are, roughly, effects for which traits were selected by natural selection (see Wright (1973)).

²⁷It would therefore not suffice to object to the personal omniGod explanation of the Universes existence just by complaining that it is an explanation of a unique type—irreducible agent-causation by an immaterial supernatural agent. But that was not, in fact, our earlier objection: following Burrell, we complained, rather, that God’s being in the agent-causal relation to the Universe gave him an *insufficiently unique* status, since (on the metaphysics which deals in irreducible agent-causation) finite agents stand in that same relation to natural events.

account, God *just is* the Love that is the Universe's supremely good *telos*. Could such a conception of God prove to be adequate?

Taking God to be the Universe's supremely good *telos* has the advantage that God is no longer a supernatural item standing over against the Universe, bearing to it a productive causal relation of the same kind as items within it bear to each other. That very feature, however, generates an obvious worry. God is either distinct or not distinct from the Universe. God's being distinct from the Universe may have problems; but God's *not* being distinct has problems also, and arguably greater ones! The euteleological proposal that God's relation to the Universe is as its *telos* seems committed to God's being at least not wholly distinct from the Universe. That—it might be complained—makes the proposal too naturalistic and too pantheistic. How, if at all, can the euteleological view make room for God's *transcendence*—something which is clearly secured by God's supernaturalness on standard personal models?

In response, we first note that the euteleological account is not committed to pantheism in the strict sense that God and the Universe are identical. If God is identified with the supremely good *telos* of the Universe, then God is not thereby identified with the Universe itself, since the Universe and its *telos* (if it has one) are not the same—indeed, they are in quite different ontological categories.²⁸

A second point in response is to consider *in general* what the relation is between the end upon which some system is directed and the system itself—for example, in the case of a machine or a biological organism. Insisting that a system's *telos* must either count as distinct or not distinct from the system itself seems the wrong kind of fork to be wielding. A *telos* surely does, *in some sense*, transcend the system whose *telos* it is, does it not? But perhaps that is because a system's *telos* is an *abstract ideal* that the system may achieve only to some degree of imperfect approximation?

But the notion that God is an abstract ideal, however noble, would give insufficient weight to God's ontological perfection. God must surely be real 'non-abstractly', or 'concretely'. If, as on the suggestion we are exploring, God is the Universe's *telos*, and a *telos* is properly understood as an ideal, there is then an important question about how the fullness of God's reality can be accommodated.

Can there be concrete, non-abstract, ideals? Platonists will say that there are 'fully real' ideals—universals more real, indeed, than the particulars that instantiate them. That is a controversial claim, whose precise meaning is not easy to fathom. But everyone will agree that there can be *concrete realisations* of ideals. Might the full reality due to the divine on the euteleological account be ascribed, then, by taking God to be the *concretely realised telos* of the Universe?

²⁸We will not pause here to consider the interesting and important questions raised by an accusation of pantheism—namely, how pantheism should be understood if it is to be recognised as a significant religious option, and what the relationship then is between pantheism and theism. For discussion see Levine (1994, 2011).

One objection would be—from the Platonists, of course—that no concrete realisation of an ideal can be *perfectly* real. But, of course, whatever counts as God must possess perfect reality. This way of understanding the euteleological view, then, would have to reject the Platonic ontological prejudice in favour of ideals/universals. Rejecting this aspect of Platonism, at least for the case of ethical universals, would be no problem for Christianity, however. Christians maintain that perfect *realisations* of agapeistic love (‘in Christ’) can and do exist in full concrete particularity.

The realised *telos* of the Universe—in virtue of its concreteness—must belong to the Universe, however. As to its status as *telos* it may be transcendent; but as to its concreteness it must be immanent. So the fork does apply, and it turns out (on our current suggestion) that God is *not* distinct from the Universe. How much of a scandal is this?

It certainly is a scandal if God counts as an item amongst other items in the Universe. And that implication does appear to threaten: if God is the realised *telos* of the Universe, and that *telos* is love, then will it not follow that God is the sum of the perfectly loving relationships achieved in the Universe? As items in the Universe go, this might be something pretty amazing, but it will surely fall short of ‘that than which a greater cannot be thought’.

Concretely existing loving relationships, on the relevant euteleological account, will, however, count straightforwardly as *exemplifications* or *manifestations* of the supremely good *telos* which is the ultimate explainer of the Universe’s existence. Concretely existing loving relationships will therefore, on this view, possess a unique status. But the objector may remain disinclined to allow that this status could be *divine*. After all, if actual loving relationships constitute the divine, then it will follow that the divine emerges within the Universe, and is therefore contingent and dependent. But any viable conception of God must surely have this the other way round: God, as that than which a greater cannot be thought, is necessary, and all else depends on Him.

It may be important, however, to question whether the greatness at issue in the Anselmian formula should count contingency and dependence as inferior to necessity and independence. Divine greatness is *onto-ethical*. It is greatness that should not be assessed against *merely* metaphysical criteria of greatness ‘*qua* being’: *ethical* criteria of greatness must also be met. Still, ontological greatness must certainly be part of the mix—but it is important to challenge the assumption that ontological greatness has to be greatness with respect to a being’s degree of dependence or independence *along the dimension of productive causality*. We warn against assuming that God must be that than which a greater producer cannot be thought—an Unproduced Producer of all else. What is *ethically* supremely perfect might necessarily not be supreme in *this* dimension of causality.²⁹ The supremely perfect One must be that than which a more ulti-

²⁹Lord Acton (1907) famous dictum may apply: ‘Power tends to corrupt, and absolute power corrupts absolutely.’

mate realisation of goodness cannot be conceived. For the perfect One also to be *ontologically* supreme, more is required—but that is supplied on the euteleological account by the existence of the Universe being explicable ultimately by its having that One as its concretely realised goal.

The classical idea that creation is a relation of dependence on God may thus be accommodated on the euteleological account. This relation is not, however, dependence on One who is the ultimate Producer of the Universe, but on One who is its ultimate Point. Now, if we employ a *broader* conception of efficient causation than the notion of ‘causing as producer’, we can agree with Aquinas that God is both final and efficient cause of the Universe. If, as final cause, God is that which ultimately explains its actuality, then God is *thereby* also the Universe’s efficient cause, in the broader sense of explaining its ‘transition’ from mere potentiality to actuality. But God need not thereby be the Grand Producer of the Universe—so, anyway, the euteleological account maintains.

A puzzle remains, however, as to how best to understand precisely what the divine is on the euteleological account. That the Universe is God’s Creation, on this account, is indeed a matter of the Universe having Love as its supremely good *telos*: but it seems problematic to identify God directly either with the *telos* itself (as an ideal) or purely with realisations of the *telos* (the sum of loving relationships). Nevertheless, on this account, loving relationships clearly are manifestations of the authentic divine: *ubi caritas et amor, Deus ibi est.*³⁰

IV

On a euteleological account of God as Love, the realised *telos* of the Universe, how much scope would there be for real divine agency? *Personal* divine agency would seem on such an account to belong only to metaphorical descriptions of the divine, apt and psychologically compelling as these may be. Does it follow, though, that, on a euteleological account, *the whole idea of divine action* would have to be regarded as purely metaphorical?

That consequence would be troubling: but it might be avoided if *there can be real agency that is not personal agency*. And it seems clear that there can be. Correlative with the broader sense of ‘efficient cause’ noted towards the end of the previous Section, there is a broader notion of agency in which any efficient cause is the agent of its effect, and its having the effect counts as action, again in a broad sense. Thus, for example, we may speak of the ‘action’ of an acid on a metal, or of the wind in eroding the rocks.

Appealing to this broader notion of agency in order to accommodate divine action may, however, seem a false move, since it may seem to render divine action ‘sub-personal’—on a par with the action of forces like the wind and the sea. That complaint overlooks, however, the possibility of action that is non-personal, not because it is sub-personal, but because it is (shall we say?) *transpersonal*. There

³⁰‘Where there is charity and love, there is God’. Traditional antiphon sung at the washing of feet at the Maundy Thursday Eucharist.

may be agency of a kind ‘greater than’ the intentional agency of an individual person.

If a non-reductivist account of social or group agency is defensible, then social and group action may be a case in point. Arguably, a group or collective (consisting in the right kinds of relationships amongst individuals) can exercise *its own* agency even though it is not itself a (full-fledged) personal (moral) agent.³¹

In any case, there is conceptual room for a type of agency ‘higher’ than personal agency. On a euteleological account that identifies God as ultimate *telos* or final cause of the Universe, God is *thereby* the ultimate explainer of the Universe’s actuality. God is then the agent (in the broad sense) of the Universe’s existence, and its actuality counts as God’s action (in the broad sense). But this divine agency, so far from being sub-personal, quite transcends any personal agency. For Aristotle, final causes are principles of intelligibility, fundamental for understanding (certain sorts of) processes—but they are not active agents that do anything.³² On a euteleological account, by contrast, God, as the Universe’s ultimate end, is *thereby* that which actualises the Universe . . . but *not* by being the Personal Substance that is its producer.³³

A euteleological explanation of the Universe’s existence is different, we think, from John Leslie’s (1979, 1989) extreme axiarchism, according to which it is *the value*, or *ethical requiredness*, of the good that effects its own realisation. That would seem to attribute a productive capacity to an ideal, which is most puzzling. On a euteleological account, the idea is, rather, that the Universe actually has a realised ultimate purpose (on a Christian view, love understood as perfect relationship), and that explains why the Universe exists as it does, *without its having any overall productive cause at all* (including—*per impossible*—its realised *telos* operating as productive cause). Does the Divine Love, then, actualise what contingently produces it? That would be nonsense with ‘actualise’ read familiarly as ‘produce’. Yet the claim may be true (if breathtakingly so!) when read aright, using a broad notion of ‘actualise’.³⁴

A euteleological account, then, does have a way of understanding divine action in creating the Universe. What of divine action *within* the Universe? That there is such action in the familiar sense of productive causation seems absolutely

³¹On the status of groups and group action, see for example May (1987), De George (1979) and Copp (1979). More recent work includes Shockley (2007) and Smiley (2010).

³²See Randall (1960, p. 128).

³³The euteleological account thus contrasts with emanation theories, which have been, we think, the most common alternatives within the theistic traditions to the idea of creation by a supernatural Person. God remains *the producer* of the Universe on emanation theories, albeit not *qua* intentional agent, whereas the euteleological approach assumes that neither God nor anything else has (nor could have?) the status of producer of the Universe.

³⁴This breathtaking idea—of something that is the ultimate creator of what contingently produces it—is in fact familiar within the Christian tradition. Mary gives birth to the Christ, the Living Word who ‘was with God and was God’ ‘in the beginning’ (John 1:1). As one of the Marian Antiphons puts it, ‘*tu quae genuisti, Natura mirante, tuum sanctum Genitorem*’ [‘you who, while Nature marvelled, gave birth to your holy Creator’].

essential to the theistic religious traditions.

Consider the Christian version of the euteleological view on which we have concentrated, according to which the divine *telos* is love. On that view, the obvious candidate for divine activity within the Universe is *the activity of love*. But what is that? The phrase might naturally be taken to refer to the loving actions of personal agents. Do we say, then, that divine action is just that, the good deeds of finite agents? That seems too reductionist. If divine activity *just is* the loving activity of finite agents, then divine activity will be something finite. To avoid that conclusion, we must surely say rather that loving actions *accord with, manifest, and even participate in*, the agency of the divine will. But that appears to imply the standard view that God is a personal super-agent.

Here is another possibility. The activity *of love* may properly refer to the productive action *of love itself*—that is, *of the relation* in which realised love consists. (Our first thought, which led to the reductionist view, failed because it focussed on the active power of *loving* agents, when it is the active power *of love* that we ought really to consider.) But does this suggestion even make sense?

Yes. Recall our earlier remarks about collective action. Arguably, it is the complex concrete relation amongst the individuals in the collective that has the productive power. Realised love, then, may perhaps possess active power *as a relation* amongst individuals. It is at least coherent to hold, we think, that a child, for example, may be acted upon by her parents' loving relationship—where this is not reducible to her being acted upon lovingly by both Mummy and Daddy. Of course, persons-in-relationship do act as individual agents, but *the relationship itself* may possess an emergent active power that transcends individual powers.

But why regard *as divine* the power of loving relationship (even if it is conceded, as we are here suggesting is possible, that it is an emergent power)? How could such a power possess the infinity, the unlimitedness, of divine power? Arguably, it may do so because of its supremacy as the most imaginably powerful force *for good*—a force for good than which none greater can be conceived? Love's power can seem weak, since it does not insist on its own way.³⁵ It cannot achieve its effects through dominating control, and makes vulnerable those who participate in it—for that is how personal loving actions are best understood theologically, not as the self-sufficient exercise of virtuous individuals, but as their participating in the divine activity. And—on the Christian understanding—the weakness of remaining true to the 'demands' of love even to the point of giving up one's life is affirmed in resurrection into a transformed participation in the divine life (of concretely realised love). On a view that rejects supernatural personal agency, these saving effects must (amazingly!) be attributed to the *active power of love itself*.³⁶

³⁵The English Standard Version translation of I Cor 13: 5.

³⁶If these effects are not supernatural, it might be supposed that they would then be explicable in principle within natural science—which seems extremely implausible. But this need not follow: see our earlier comments on two distinct senses of 'supernatural', note 11 above.

But how could the activity of love count as *divine* activity within the Universe if it is the activity of something relational that emerges within that Universe only after a long, chancey, process of evolution? One might complain that such activity, even if its *ethical* greatness fits it for divinity, will be highly dependent for its existence on much else. It cannot therefore be *ontologically* fit to count as the activity of the divine.

This unfavourable conclusion would follow if ontological supremacy were solely a matter of supremacy in the order of productive efficient causes. But ontological supremacy may rest, rather, on being that which ultimately explains all that is. On a euteleological account, in the absence of any ultimate Producer, the Universe as a creation *ex nihilo* may be ultimately explained in terms of its Point, realised love. What we are now able to add is that the divine status of love's activity *in* the Universe depends on love's being the ultimate *telos* of the Universe, and *thereby* its Source. This does entail the high paradox that the existence of what realises the Source of all depends on contingencies that it, the Source, actualises (just in the sense that it explains their actuality). But that is what it is for the Goal to be also Source—and that's the only kind of Source a creation that is literally *ex nihilo* can have.

A euteleological account of the Universe's existence as explained in terms of its having the supreme good (the existence of perfectly loving relationships) as its ultimate *telos* seems, then, to have the resources to provide *an* interpretation of divine action, both in creating the Universe and in saving activity within it. It is, of course, a further question whether it is justifiable to understand the Universe according to such an account. There is, after all, much in the Universe that is *dysteleological*, resisting inclusion in an overall picture that takes the point of what exists to be love and its enjoyment. The *particular* difficulty posed by the standard Argument from Evil will disappear, when there is no personal first productive cause and sustainer of horrors whose moral perfection is at stake. But it does not follow that there is no longer any problem of evil. The problem of evil for a euteleological account will be a different one—namely, explaining how a world that contains much that is unloving and resistant to love can nevertheless have the existence of love as its ultimate end. A response to this 'dysteleological problem of evil' is certainly needed—and it will be important to consider whether the difficulty posed by evil for belief in the existence of God as conceived on a euteleological account is more tractable than the difficulty it poses for belief in the personal omniGod. We are content here, however, to have—we hope—undercut confidence in the assumption that an adequate notion of divine action is obtainable only if God is the personal omniGod, and to have suggested ways in which divine action may be accommodated on one kind of non-personal understanding of what it is for the Universe to be the Universe of God's creation—namely, a euteleological account of the Universe as directed upon a supremely good end for the sake of which it exists though it lacks any ultimate producer.³⁷

³⁷We are grateful to audiences at the American Philosophical Association Pacific Division Meeting (San Diego, April 2011) and at the annual conference of the Australasian Association

REFERENCES

- Acton, (Lord) John. 1907. "Letter of 5 April 1887 to Bishop Creighton." In *Historical Essays and Studies*, edited by J. N. Figgis and R. V. Laurence. London: Macmillan.
- Alston, William. 1989. *Divine Nature and Human Language: Essays in Philosophical Theology*. Ithaca, NY: Cornell University Press.
- Alston, William. 1994. "Divine Action: Shadow or Substance?" pp. . 41–62. University Park, PA: The Pennsylvania State University Press.
- Aquinas, T. 1975. *Summa Contra Gentiles: Book One: God*. Notre Dame: University of Notre Dame Press. 1259–1264.
- Aquinas, Thomas. 1945. "*Summa Theologica, Part I.*" In *Basic Writings of Saint Thomas Aquinas*, edited by Anton C. Pegis. Vol. 1. New York: Random House. 1265–1272.
- Aristotle. 1915. "*Nicomachean Ethics.*" In *The Works of Aristotle*, edited by William David Ross. Oxford: Clarendon Press.
- Bishop, John. 1989. *Natural Agency: An Essay on the Causal Theory of Action*. Cambridge: Cambridge University Press.
- Bishop, John. 1998. "Can There Be Alternative Concepts of God?" *Noûs* 32: 174–188.
- Bishop, John. 2003. "Prospects for a Naturalist Libertarianism: O'Connor's Persons and Causes." *Philosophy and Phenomenological Research* 66: 228–243.
- Bishop, John. 2007. "How a Modest Fideism May Constrain Theistic Commitments: Exploring an Alternative to Classical Theism." *Philosophia* 35: 387–402.
- Bishop, John. 2009. "Towards a Religiously Adequate Alternative to OmniGod Theism." *Sophia* 48: 419–433.
- Bishop, John and Ken Perszyk. 2011. "The Normatively Relativised Logical Argument from Evil." *International Journal for Philosophy of Religion* 70: 109–126.
- Burrell, David. 1993. *Freedom and Creation in Three Traditions*. Notre Dame, IN: University of Notre Dame Press.
- Burrell, David. 1998. "The Attributes of God: (a) Simpleness." pp. . 70–74. London: Cassell.
- Chisholm, Roderick M. 1966. "Freedom and Action." pp. . 11–44. New York: Random House.
- Cooper, John W. 2006. *Pantheism and the Other God of the Philosophers: From Plato to the Present*. Grand Rapids, MI: Baker Academic.

of Philosophy (University of Otago, New Zealand, July 2011), and to Thomas Harvey and Jeremy Reid, for helpful comments on earlier drafts of this paper. We are specifically indebted to Thomas Harvey for the notion of a 'dysteleological' problem of evil.

- Copp, David. 1979. "Collective Actions and Secondary Actions." *American Philosophical Quarterly* 16: 177–186.
- Davidson, Donald. 1963. "Actions, Reasons, and Causes." *Journal of Philosophy* 60: 685–700.
- De George, Richard T. 1979. "Social Reality and Social Relations." *Review of Metaphysics* 37: 3–20.
- Flint, Thomas P. 1998. *Divine Providence: The Molinist Account*. Ithaca: Cornell University Press.
- Forrest, Peter. 2007. *Developmental Theism: From Pure Will to Unbounded Love*. Oxford: Oxford University Press.
- Hartshorne, Charles. 1984. *Omnipotence and Other Theological Mistakes*. Albany, NY: State University of New York Press.
- Helm, Paul. 1980. "God and Spacelessness." *Philosophy* 55: 211–221.
- Johnston, Mark. 2009. *Saving God: Religion After Idolatry*. Princeton, New Jersey: Princeton University Press.
- Leibniz, Gottfried Wilhelm. 1951. "Essays on the Justice of God and the Freedom of Man in the Origin of Evil." In *Theodicy: Essays on the Goodness of God, the Freedom of Man and the Origin of Evil*, edited by Austin M. Farrer. London: Routledge & Kegan Paul.
- Leslie, John. 1979. *Value and Existence*. Oxford: Basil Blackwell.
- Leslie, John. 1989. *Universes*. London: Routledge.
- Levine, Michael. 1994. *Pantheism: A Non-theistic Concept of Deity*. London: Routledge.
- Levine, Michael. 2011. "Pantheism." *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/archives/spr2011/entries/pantheism/>.
- May, Larry. 1987. *The Morality of Groups: Collective Responsibility, Group-based Harm, and Corporate Rights*. Notre Dame, IN: University of Notre Dame Press.
- Nagasawa, Yujin. 2008. "A New Defense of Anselmian Theism." *Philosophical Quarterly* 58: 577–596.
- O'Connor, Timothy. 2000. *Persons and Causes: The Metaphysics of Free Will*. Oxford: Oxford University Press.
- Phillips, Dewi Zephaniah. 1965. *The Concept of Prayer*. London: Routledge & Kegan Paul.
- Plantinga, Alvin. 2000. *Warranted Christian Belief*. Oxford: Oxford University Press.
- Randall, John Herman. 1960. *Aristotle*. New York: Columbia University Press.
- Shockley, Kenneth. 2007. "Programming Collective Control." *Journal of Social Philosophy* 38: 442–455.
- Smiley, Marion. 2010. "From Moral Agency to Collective Wrongs: Re-thinking Collective Moral Responsibility." *Journal of Law and Policy* 19: 171–202.

- Swinburne, Richard. 1977. *The Coherence of Theism*. Oxford: Oxford University Press.
- Wright, Larry. 1973. "Functions." *Philosophical Review* 82: 138–168.

GOD KNOWS THE FUTURE BY ORDERING THE TIMES

Ryan Byerly

Many contemporary philosophers who are proponents of classical theism are also libertarians (call these libertarian classical theists).¹ As classical theists, they maintain that God has infallible foreknowledge of the future free actions of human persons.² And, as libertarians, they hold that actions are free only if they are not causally determined and only if those who perform these actions could have done otherwise.³ Explaining how all of this could be true is a difficult task which requires providing an account of the *mechanics* of infallible divine foreknowledge. The task is to say *how* God could know these future actions infallibly where these actions are not causally determined and where the persons who perform them could have done otherwise.

Perhaps the most discussed answer to this task on behalf of libertarian classical theists is the Molinist account of infallible divine foreknowledge. But many have found Molinism unsatisfying. My aim here is to articulate an alternative, or a family of alternatives, to Molinism. According to the view I will propose, God infallibly foreknows the future by ordering the times. I won't argue that this view is *superior* to Molinism or to other accounts of the mechanics of infallible divine foreknowledge, though it does have some advantages over the former view. I only aim to show that it deserves to receive attention from philosophers of religion alongside Molinism and other accounts of the mechanics of divine foreknowledge, since it accomplishes the same aims that those views attempt to accomplish.

In section one, I'll explain more thoroughly the task of the libertarian classical theist. In section two, I present an initial statement of the view I want to propose concerning the mechanics of divine foreknowledge. In section three, I show that this view is consistent with the actions of persons not being causally determined in any sense which libertarians will find threatening. And in section four I show that this view is consistent with persons being able to do otherwise than they

¹This trend is chronicled in Rudder Baker (2003).

²It is, of course, consistent with classical theism to hold that God doesn't have infallible foreknowledge of future events, having instead only infallible knowledge of these [see, e.g., Stump and Kretzmann (1991)]; but I will be focusing in this paper on views according to which God's knowledge of the future is foreknowledge.

³Some might think that these two conditions are roughly the same, or that one entails the other and so both needn't be mentioned. But it isn't immediately clear that this is so. So, I list each separately. It is important to note that while denying causal determinism is definitive of libertarianism, in some rare cases libertarians have not required the ability to do otherwise. For discussion of such so-called Frankfurt-libertarian views, see Timpe (2006).

in fact do. At least, I show that *if* libertarian classical theism is consistent at all, then the account of the mechanics of foreknowledge I propose is consistent with the absence of causal determinism and the presence of the ability to do otherwise. By so doing, I offer an alternative way of defending the consistency of libertarian classical theism.

The Libertarian Classical Theist's Task

We can see the difficulty facing libertarian classical theists if we think about *how* someone might come to know a contingently true proposition about the future.⁴ The only way of doing this with which we are familiar will involve inference from what is known about the past together with what is known about the laws of nature. For instance, I might know that it will rain in Mississippi tomorrow morning on the basis of my knowledge that it rained earlier this evening in Texas and my knowledge of the law-like behavior of weather patterns in the southeastern United States. I might know that the formerly blue litmus paper will be red in a few minutes on the basis of my knowledge that the paper was just sprinkled with acid and my knowledge of the laws governing the interaction of acid and litmus paper.⁵ And so on. But if this way of knowing is the only way of knowing contingent propositions about the future, and not just the only way of knowing such propositions with which we are familiar, then libertarian classical theism is in jeopardy.

For, suppose that the only way to know future contingent propositions is on the basis of predictions from the past and laws of nature. And suppose that God knows the future actions of persons. It follows either that God's knowledge of the future actions of human persons is not infallible or that the actions of those persons are causally determined. For, if God's knowledge of the future on the basis of prediction from the past and laws is to be infallible, then the past and laws must ensure that the future goes as God believes, since infallibility requires that one's evidence ensures what one believes.⁶ The only conceivable way the past and laws could ensure that things go as God believes is if the past and laws causally determine the future. So, if God's knowledge is to be infallible, then the actions of persons are causally determined. And this is just to say that either God's knowledge is fallible or the actions of human persons are causally determined.

This conclusion directly threatens the libertarian classical theist's view outlined above. For, the libertarian classical theist is committed to both God's infallible foreknowledge of human free actions and to the claims that actions

⁴We don't want to talk here about contingently true propositions which are simply true *in* the future. For, such propositions could include propositions *about* the past, which we could know in ways other than by consulting the past and laws of nature.

⁵I might not know a whole lot about the relevant laws; but it is only insofar as I have some grasp of them, even if I don't think of them *as laws*, that I obtain knowledge.

⁶See, e.g., David Lewis's classic discussion and limited defense of infallibilism in Lewis (1996).

are free only if they are not causally determined and only if those who perform those actions could have done otherwise. But, given the conclusion just stated, these commitments are inconsistent. For, this conclusion implies that if God infallibly foreknows the future actions of persons, then these actions *are* causally determined. But, then, those actions cannot be free, given the libertarian's commitments. So, it is false that God infallibly foreknows the future *free* actions of persons. Furthermore, libertarians find it plausible that if a person's action is causally determined, then that person couldn't have done otherwise. It again follows that God can't know the future *free* actions of persons, since libertarians maintain that free actions are actions where the subject of those actions could have done otherwise. Thus, if the only way to know future contingent propositions is by prediction from the past and laws of nature, then libertarian classical theism is inconsistent.

It is incumbent upon libertarian classical theists, therefore, to defend the view that there is some way to know contingent propositions about the future, even to have *infallible* foreknowledge of these, which does not imply causal determinism and which does not imply that human persons could not act otherwise than they do. One way to do this is by providing a model of God's foreknowledge which makes it possible for God to know the future actions of persons infallibly while at the same time not requiring causal determinism and allowing that the persons who perform these actions could have done otherwise. To provide such a model is to take on the task I described above as the task of providing an account of the mechanics of divine foreknowledge.⁷

As I have mentioned, the most discussed such model is probably Molinism.⁸ Though I cannot go into all of the ins and outs of this view, or into the variety of Molinist views there are,⁹ I will provide a brief sketch of this family of views here. Molinists maintain that God has infallible foreknowledge of the free actions of creatures on the basis of a combination of parts of his middle knowledge and his free knowledge. God's middle knowledge consists (at least partly) in his knowledge of true subjunctive conditionals whose truth is independent of any of his volitions. Included among these conditionals are what are sometimes called subjunctives of freedom—subjunctives which specify, for any agent S, circumstances C, and action A, whether S *would* freely do A or *would* freely do something other than A *were* C to obtain. God's free knowledge is his knowledge of contingent propositions which depend upon his volitions. The relevant portion of this knowledge used together with his middle knowledge to get infallible foreknowledge of human free actions is God's knowledge of which circumstances he wills to bring

⁷I borrow this latter phrase from Viney (1989).

⁸For some representative contemporary defenses of this view, see Flint (1998), Craig (1990), and Kvanvig (1986). Another view with some contemporary defenders [e.g., Hunt (1993)] is the simple foreknowledge view according to which God knows the future by simply *seeing* it in a certain way.

⁹I am thinking chiefly of the contrast between standard and maverick Molinism of the sort explicated in Kvanvig (2002).

about. On the basis of his free knowledge of the circumstances he wills to bring about and of his middle knowledge of subjunctives of freedom, God is able to infallibly foreknow what all creatures will freely do. And, the Molinist maintains, God's knowing these actions in this way does not impinge upon human freedom since these actions are not causally determined and since the agents who perform them could have done otherwise.¹⁰

I am not interested here in rehearsing objections to this picture.¹¹ But the fact that such objections are plentiful should provoke readers to share some interest in what I am interested in doing here. What I *am* interested in doing is providing an alternative to Molinism which accomplishes the same things for libertarian classical theism that Molinism does. I'll begin to do this in the next section.

God Knows the Future by Ordering the Times

On the model I will propose, God has infallible foreknowledge by virtue of his self-knowledge, his knowledge of the times, and his infallibly competent deductive powers. God begins with knowledge of every possible time. God then wills for some of these times to be ordered in a particular way so as to constitute the history of the actual world. By virtue of his self-awareness he knows that he wills this. His willing that the times be ordered in such-and-such a way entails that the times *are* ordered in such-and-such a way, since he is omnipotent. And God knows that he is omnipotent, too, so he deduces that the times are ordered in such-and-such a way. Finally, the times being ordered in such-and-such a way entails everything which occurs in the history of the world, including everything which occurs at every future time. Since God knows this, too, he can competently deduce what will happen at every time, including the future times. Indeed, he can know infallibly what will happen at every time, since his evidence entails everything that happens at every time.

I should say some more about some important elements of this story. Start with the times. The view I propose borrows from a view with an impressive pedigree in the philosophy of time according to which times are some sort of maximal or nearly maximal abstract representational object.¹² This contrasts with views according to which times are concrete entities consisting of individuals which instantiate properties and so forth. I am not entirely persuaded that the view I will present here could not be adopted by an advocate of concrete times. But it is easier to see how the view would go for a friend of abstract times, so I shall start there.

As I will explicate the view, times are nearly maximal, consistent propositions. A maximal proposition is a proposition which, for every proposition *p*, includes either *p* or not-*p* as a conjunct. The times in my theory needn't *quite*

¹⁰Such a defense is surely at work in Flint (1998).

¹¹For some representative objections, see Adams (1991) and Hasker (1995).

¹²See e.g., Chisholm (1979), Davidson (2003, 2004), Fine and Prior (1977), Zalta (1987), Crisp (2007), Bourne (2006).

be maximal. They may include, for every proposition p *except for propositions about relations between times*, either p or not- p as a conjunct.¹³ This is their essential nature, anyway. What makes a time the particular time that it is has to do only with propositions it includes which don't say anything about that time's relation to other times. But, contingently, a time may include propositions about which other times it is related to. Thus, the times in my theory say everything about what is going on at those times except for how other times are related to those times, so long as they are not related to any other times. If they are contingently related to other times, they may say this as well. I'll discuss this further in a moment; but for now let this suffice for a presentation of *near* maximality. Times must also be *consistent*. To be consistent, a time must be such that possibly, every conjunct of it is true. An example of a time would be a proposition like \langle Obama is President and George Bush is not President and the angels are singing and two plus two is four and ... \rangle where in the "..." we fill in either p or not- p for every proposition p (other than propositions about the relations between times, if the time has no such relations) such that possibly every conjunct in this big conjunction is true.¹⁴

The story about God's foreknowledge above also requires that at least some of these times can stand in relations to one another. The relations they stand in to one another are primitive, unanalyzable *earlier than* relations to one another. This is just what we should expect such a theory to say since, intuitively, some times are earlier than others and some times are later than others. Yesterday is earlier than today and tomorrow is later than today. Times which are related to one another by these relations form a series of times. We can call such a series of times a series of *e-related* times. Any time which is earlier than a time which is earlier than the present time we will say is also earlier than the present time; and any time which is later than a time which is later than the present time we will also say is later than the present time.

We can use our series of e-related times to make sense of when some proposition was true, when some proposition *will* be true, and when some proposition *is* true in the present-tensed sense of "is" (i.e., "is currently"). A proposition p *was* true just when that proposition is entailed by a time which is earlier than the present time. A proposition p *will* be true just when that proposition is entailed by a time which is later than the present. And, a proposition is true, in the present-tensed sense of "is," just when that proposition is entailed by the present time. The present time, we will say, is the true time. Times, on this view, take turns being true. Indeed, propositions are true at some times and false at others. On this view, then, propositions can change their truth-values.¹⁵

¹³This sort of nearly maximal account of times is presented as just the sort of account of times an ersatz should prefer in Finch and Rea (2008). They don't say whether times which are in fact e-related might be *fully* maximal.

¹⁴I use the angle brackets " $\langle \dots \rangle$ " to indicate whatever proposition is expressed by "..."

¹⁵And thus the view follows the fundamental insight which led Prior to develop his tense logic [see Prior (1996)].

We might ask whether all of the times are e-related to other times or whether only some are.¹⁶ There are actually three options here.¹⁷ First, we might say that every time is part of the one and only series of e-related times. This would be a mistake, though. For, times are nearly maximal consistent conjunctions. Some nearly maximal consistent conjunctions include conjuncts like <unicorns fly about>. But, then, if every time is part of one series of e-related times, then some times according to which unicorns fly about are earlier or later than the present time, since the present time, too, is part of this one series. And surely we don't want our account of times to entail either that unicorns once flew about or that they will eventually fly about, since it could be that as a matter of contingent fact they never fly about.

A second option would be that all times are part of some series of e-related times or other, but that not all are part of the same series. On one such natural proposal, there would be one series of e-related times which constitutes the series of e-related times which represents the actual world, and other series for other worlds. If we took this view, we would have to find some way to distinguish the actual world's series of e-related times from those of other possible worlds. We might try, for instance, a primitive actuality property.

On the other hand, if we find that adopting another primitive like this makes the theory too complex, we can just say that there is only one series of e-related times. It includes only those times that did, do, or will represent the actual world. Other times could have been e-related to one another; but as a contingent matter of fact they are not. I'll talk here as if there is only one series of e-related times, and as if it is the one which is constituted by only those times which did, do, or will represent the actual world. Only times which are part of this series, then, will include propositions about their relations to other times as conjuncts.

Enough about the times. I should also say something about what God *does* with them. On the view sketched above, God *orders* these times. That is, he brings it about that they are e-related in the way that they are by willing this to be so. He wills, for each particular time in our series, that it be earlier than some other time which be earlier than some other time and so on.

This shouldn't be objectionable in itself. Defenders of the sort of picture of times I have sketched here often take it as a primitive, unexplained matter of fact that certain times are e-related to one another rather than other times. Thomas Crisp, for instance, writes the following:

So the suggestion here is that the earlier than relation connects certain abstract times and not others, though which abstract times it connects is something that could have been different. Well, one might wonder, why wasn't it different? Why does the earlier than relation connect just the times it does? What explains the fact that it connects

¹⁶The discussion in the next few paragraphs follows the presentation in Crisp (2007).

¹⁷Well, there is also a fourth, which I shall ignore in the text. One could maintain that there are many series of e-related times (as on the second proposal discussed in the text) but that nothing distinguishes them and nothing needs to—since all such series are ontologically on par. This is to wed the view presented here with something like modal realism.

these times and not others? . . . [A]s plausible an answer as any to [these] questions is that it's a brute, contingent fact that the abstract times come temporally ordered as they do. Explanation has to come to an end somewhere, and it's not unreasonable to suppose that it bottoms out in the contingent fact that certain times are earlier than certain other times. (Crisp 2007, p. 104)

Such theorists shouldn't find it in principle objectionable, of course, if this contingent fact *does* have an explanation. Crisp, in this passage, is trying to defend the view that it is *acceptable* if the fact doesn't have an explanation—not that it is unacceptable if it does. And, indeed, certain theists—for instance, those who are committed to a strong principle of sufficient reason—may have a special motivation for thinking that there *is* an explanation of this contingent fact. My suggestion is that the explanation has to do with God's volitions. God wills that the times be related in a certain way, and by virtue of his omnipotence he is able to order them as he wills.

It is because of this feature of the present view that I said earlier that times are essentially only *nearly* maximal rather than maximal. For, if the times were essentially maximal and so included conjuncts about how they are related to other times, then God couldn't just relate them as he wished. He couldn't take a time that said it was earlier than a time t and make that very time not be earlier than t . For, in that case, that time would say both that it was earlier than t and not earlier than t , in which case it would not be a time since it would not be consistent. There may be some other way to develop a view analogous to the one I am offering here which does allow that the times essentially imply things about how they are related to one another—perhaps for instance by saying that instead of *ordering* the times, God *actualizes* some particular series of times. Such a view would structurally parallel the one I develop here; so I will not present both. I shall stick with the approach here. The times are essentially only nearly maximal, not fully maximal, and God orders them as he wishes. Those which he orders by earlier than relations to one another are the only ones which contingently say something about their relations to each other.¹⁸

I have talked as if God can just order these nearly maximal times in any way he wishes. This may be misleading. For, there are likely some significant constraints which govern the ways in which God can order the times. For instance, nothing said above rules out the possibility that causal claims may be included in those propositions which constitute a time. But, this will imply some limitations on the ways in which God may order times, given common views about the directionality of causation. For, suppose that God orders a time t_3 third in the order of times. And, suppose that t_3 includes as one of its conjuncts the proposi-

¹⁸The view of times I advocate here is not *entirely* without cost. It implies, for instance, that the present could have been just as it was and the past and future different—a claim sometimes called the principle of temporal recombination. Many may count this implication a boon rather than a bane, since they find the principle intuitive. Some, however, particularly *presentists*, have denied this principle in order to escape an objection to their view known as the grounding objection. See, e.g., Crisp (2007). The view I advocate here implies that if the presentist is to escape the grounding objection, she must find some other way of doing so.

tion that event e is caused by event e' . Suppose further that causation always, or at least in instances like this one, is such that causes temporally precede their effects. Given these assumptions, there will be some limitations to what times can be ordered earlier than t_3 in our series. God cannot choose times which do not include the existence of e' as the first and second times in the series, for instance. The nature of causation, at least in this instance, will bar this possibility.¹⁹ This abstract example illustrates the way in which the metaphysics of causation may place constraints on the ways in which God can order the times. Of course, for times which include many more causal claims as conjuncts, the limitations will be more stringent.

Other sorts of limitations on the ways in which God can order the times derive from the nature of libertarian freedom. Many libertarian classical theists are attracted to ideas like those of Richard Swinburne (2004) according to which a world with sensible libertarian free choices is one which requires certain law-like regularities. Sensible free decisions may require a world where, with law-like regularity, events of type e follow from events of type e' , for example. Without such regularity, the world would be an epistemological mess—agents wouldn't be able to make sense of the world well enough to choose good or ill. It is, on this picture, strictly speaking possible for God to intervene and prevent an event of type e from following from an event of type e' . But, because by nature God places special value on sensible free decisions, he will not engage in such intervention in any widespread manner. This again places some constraints on the ways in which God can—can consistently with his own nature, anyway—order the times.

Despite these constraints, there can still be considerable flexibility in the ways in which God can order the times, given the libertarian classical theist's assumption of the denial of comprehensive causal determinism of all events. For those events which are not causally determined, God is free to choose between times which include them and times that don't consistently with his own natural preference for a world with widespread regularities. This feature of the present view will be important in later sections, where I will return to it. For now, let this suffice for a presentation of the nature of times and the constraints governing the way in which God may order them.

The features highlighted above are the primary metaphysical elements of the theory. The primary epistemological elements of the theory are what it says about God's self-awareness, his knowledge of times, and his deductive knowledge.

Part of the theory is that God knows what he wills by his self-awareness. This self-awareness may involve introspection, or something weaker. Some may worry that introspection is a bit too strong, that it involves too much effort on God's part. Perhaps God knows what he wills, including his willing that certain times be e -related in a particular way, through some sort of self-awareness which

¹⁹One might attempt to avoid these constraints by denying the assumptions about causation—by, for instance, insisting that all causation is simultaneous. But, endorsing this view would add unnecessary complexity to the present view. So, I avoid it in the text. For an overview discussion of causation, including its directionality, see Schaffer (2007).

falls short of introspection.²⁰ Either sort of self-awareness may plausibly lead to knowledge and even infallible knowledge for someone sufficiently good at using the method required. And we may suppose that God is sufficiently good here.

The theory also requires that God knows that he's omnipotent—that whatever he wills goes. Though this sort of self-knowledge on God's part is often taken for granted—it must be, for instance, an important part of the Molinist's story, too—it's not obvious how it goes. Since this is a sort of self-knowledge, we might imagine that God has it by virtue of some kind of self-awareness as in the case of his knowledge of his will. But, it is also, according to the classical theist, a metaphysically necessary fact. As such, God might know it by virtue of *a priori* reasoning. Either way, it is unlikely that the present account of God's infallible foreknowledge will be found wanting because it relies upon the claim that God has infallible knowledge of his omnipotence. It is hard to see how an account of infallible foreknowledge *wouldn't* require this; so there is no special problem for the present view.

God must also know the times. But to know the times is simply to know all of the consistent combinations of propositions, excluding propositions about the relations between times. Again God may know this *a priori*. It may even be that such consistent combinations of propositions simply are ideas in God's mind. Some have defended this sort of view about possible worlds, and times on the present account parallel popular accounts of possible worlds.²¹ Since other accounts of the mechanics of divine foreknowledge, like Molinism, also require God to have substantial modal knowledge, it should not be especially objectionable that this account does the same.

Suppose then, with the present account, that God does have infallible knowledge of his will, his omnipotence, and of the times. And suppose, with the story above, that God wills for these times to be ordered in a particular way. God then has infallible knowledge that he wills the times to be ordered in the particular way that he ordains, and he has infallible knowledge that what he wills goes. If God also is infallibly competent at deduction, which is not unreasonable to suppose, then he can have infallible knowledge that the times *are* e-related in the particular way he has willed. And this will be plenty to ensure his infallible foreknowledge of all that occurs in the history of our world.

For, suppose that God has infallible knowledge that the times are e-related in the particular way he has willed. He will then know, for any time *t*, its relations to every other time. But, then, for any future truth *F*, God will know *F*. For, *F* will be true just in case it is entailed by a time which is later than the present

²⁰I'm particularly attracted here to Uriah Kriegel's idea that conscious states (like God's willings) are self-representational and that we might be able to know them through their self-representation [see Kriegel (2009)].

²¹For accounts of possible worlds which parallel the account of times presented here, see Plantinga (1974), Chisholm (1976), Fine and Prior (1977), Adams (1974), Zalta (1983). For discussion of views according to which possible worlds and other abstracta are divine thoughts or something similar, see Morris (1987) and Menzel (1990).

time. By hypothesis, God knows infallibly the relations between the present time and every other time, including whatever time entails F. God knows infallibly, then, that F is entailed by a time that is later than the present. And this is just to say that God knows that F will be true—indeed, he even knows at which future time F will be true. And he knows this infallibly, since by virtue of his infallible deductive competence he deduces it from other things he infallibly knows.

In this section, my goal has been to present an account of how God could have infallible foreknowledge by ordering the times. In the next two sections, I shall argue that this account is consistent with the absence of causal determinism and with persons having the ability to do other than what they in fact do.

No Causal Determinism

Here I want to argue that the account of divine foreknowledge developed in section two is consistent with the absence of responsibility-undermining causal determinism, i.e., the sort of causal determinism which would undermine a person's moral responsibility for her actions. I'll offer a positive argument for this claim, and then reply to some arguments against this claim.

The positive argument is as follows. The view developed in section two is consistent with persons being the agent-causes of their actions. But agent-causation is inconsistent with responsibility-undermining causal determinism.²² So, the view developed in section two is consistent with the absence of responsibility-undermining causal determinism.

First, let me show that the view developed in section two is consistent with agent-causation. The sort of agent-causation involved here is that which is typically contrasted with event-causation. Agent causation occurs when agents themselves, and not events within agents (like an agent's having a certain desire or belief), cause actions. Agent causation of this sort is endorsed by a number of libertarians. I'll use the conception of agent-causation developed by Bergmann and Cover (2006). Their account is as follows:

(AC) X is the agent-cause of e iff:

- (i) X is a substance that had the power to bring about e
- (ii) X exerted its power to bring about e, and

(iii) Nothing distinct from X (not even X's character) caused X to exert its power to bring about e.

To show that the view of divine foreknowledge discussed in section two is consistent with this kind of agent causation of future actions, I must show that that view is consistent with there being some agent X and event e such that X and e satisfy (i)–(iii) of AC at some future time.

This is not difficult to show. Let t be some future time—some time later than the present. The account of divine foreknowledge from section two will be

²²This is not *quite* to say that agent-causation is inconsistent with causal determinism. Some have argued that compatibilism is consistent with agent-causation [see, e.g., Markosian (2012)]. The account of agent-causation developed below, though, is at least consistent with the absence of causal determinism.

consistent with (i) if t entails that there is some agent who has the power to bring about an event e . There is nothing problematic about this; so, the view of section two is consistent with (i). Likewise, there is no problem with t entailing that this agent exercised its power to bring about e so as to bring about e . This implies that the view of section two is consistent with (i) and (ii). All we need now in order to show that the view of section two is consistent with the conjunction of (i)-(iii), then, is to show that t needn't require that something distinct from our agent caused it to exert its power to bring about e . And this is clearly correct. For, the view of section two is compatible with there being *no* causation at all within a world, let alone with there being no cause of an agent's exercises of its powers. So, time t needn't require that our agent is caused by anything distinct from itself to exert its power to bring about e .

Let me say in a bit more intuitive way what is going on here. On the view presented in section two, everything that happens at every future time is entailed by that future time. Indeed, everything that happens at every future time is entailed by the relations each of those future times stands in to the present time. That these times stand in the relations they do to the present time entails everything that will happen at those times. God knows what relations these times stand in to the present time, since he made these times stand in those relations. And this is how he knows what will happen at every future time. But none of this implies *anything* about *any* causal relations between things in the world. God could order times such that there are *no* causal relations at all in a world, save the causal relation between himself and the particular ordering of times he brings about. He would still thereby have infallible foreknowledge of everything that would ever occur, but this would not in the least require causal determinism. In the same way, God can bring it about that a future time which entails that some agent X exercises her power to e so as to bring about e is later than the present time without it being the case that X 's exercising her power to e is caused by *anything*.

Of course, given some of our earlier observations about the nature of causation and of libertarian freedom, and given what we know of God's own character, it is quite unlikely that God would create a world where no causation at all occurred save his creation of that world. It is rather more likely that he would create a world with a great deal of causal regularities—enough even for sensible free decisions to be made by creatures with libertarian freedom. But this needn't imply that *all* events in that world be causally determined. And, more apropos to our present case, it needn't imply that our agent X 's exercising her power to e be causally determined. X 's exercising her power to e may be caused by nothing at all. God may order the times in such a way that agents like X exercise their powers without being causally determined to do so as long as there are enough regularities in the way things go with the rest of the world that agents like X are able to make sensible decisions about how to exercise their powers. That such a combination of widespread regularities and causally undetermined exercises of agential power is achievable is a pillar of many libertarian classical theistic

perspectives, and one that needn't be defended here. Rather, what is of interest here is the power the view presented in section two has to explain how God could infallibly foreknow such exercises of power.

So, the view of section two is consistent with future actions being agent-causally brought about. But it is widely held, especially by libertarians, that an action's being agent-causally brought about is sufficient for the agent who performed that action to have done so responsibly.²³ After all, if X is the agent cause of her action, then the causal buck for her action stops with X. If X isn't responsible for the actions she agent-causes, nobody is. But surely somebody is—X is. Agents who agent-cause their actions are therefore responsible for those actions. But this implies that the view of section two is consistent with the absence of a responsibility-undermining causal determinism. For, if such a responsibility-undermining causal determinism were *present*, then agents couldn't be responsible for those actions they agent-causally brought about.

Each premise in the argument above has now been supported. The view of section two is consistent with agent-causation. But, agent-causation implies the absence of responsibility-undermining causal determinism. So, the view of section two is consistent with the absence of responsibility-undermining causal determinism.

Someone will object that, surely, the view of section two *does* imply that agents who exercise their powers to bring about effects are caused to do so. They are caused to do so by God. For, God causes the times at which these agents exercise these powers be e-related to the present time in just the way he wants, and the fact that these times stand in such relations to the present *entails* what these agents do. So God surely does cause these agents to do what they do.

But this objection is unpersuasive. It relies upon some transfer principle like the following: if X causes e and e (or the proposition that e occurs) entails p, then X causes p (or, X causally brings it about that p). The most general such principles are clearly false, and it is difficult to see how some more limited such principle which applied to the case of God but not to other objectionable cases could be devised without being suspiciously *ad hoc*. The general principle just stated is clearly false in the case where p is a necessary truth, for instance. For, where p is a necessary truth, the principle would imply that anytime anyone causes anything, she also causally brings it about that this necessary truth is true, since it is entailed by every proposition. And this is surely wrong; we don't causally bring it about that necessary truths are true whenever we do anything. The principle also plausibly fails generally for contingent truths. One of the other reindeer can cause Rudolph's nose to be covered in black soot without thereby causing Rudolph to have a nose, though <Rudolph's nose is covered in black soot> entails <Rudolph has a nose>.

The mistake underlying this argument is that causation and entailment are

²³See, *inter alia*, O'Connor (2000) and Rowe (2006).

being confused. This is a natural mistake since, after all, entailment seems to be a very strong relation and causation a strong relation but not quite as strong as entailment. If entailment is stronger than causation in the sense that entailment relations entail causal relations, then the principle above would be a simple principle of transitivity: if x causes e and e causes g then x causes g .²⁴ Fortunately, entailment relations *don't* entail causal relations. For instance, to borrow an example from Bergmann and Cover, <the moon orbits the earth and X causes e > entails < X causes e >, but the former does not cause the latter. Nor does the moon's orbiting the earth and X 's causing e cause X 's causing e . Entailment and causation are simply not the same thing.²⁵

The objection fails. It is reasonable to conclude that the view of section two does not require that for an agent to exercise her power to e , God or anyone or anything else must cause her to exercise her power to e . She can exercise her power to e without being caused to do so, and when she does, on Bergmann's and Cover's account, she agent-causes e . So long as agent-causation is sufficient for responsibility, the view of section two is consistent with the absence of responsibility-undermining causal determinism.

Our objector may wish to revise her original statement of the objection. "Fine," she says, "neither God nor anyone nor anything else need cause an agent's exercise of her powers on your view. But, still, that she exercises her powers as she does is *entailed* by something which occurred in the past—God's ordering of the times. On typical statements of causal determinism, this will be sufficient for any exercise of power on the part of this agent to be causally determined. So, you're not off the hook so easily."

There is something right about this objection and something wrong about it. What's right about it is that typical statements of causal determinism do imply that if an action is entailed by something which occurred in the past, then it is causally determined. For, typical statements of causal determinism define it as follows: every event is the consequence of the past and laws of nature.²⁶ And here consequence is typically understood as *logical* consequence. Accordingly, a particular event is causally determined if it is impossible for the past and laws of nature to remain as they are and that event to be different. Further, the view discussed in section two *does* imply this concerning every future event. Every future event is such that it is impossible for the past and laws of nature to be just as they are and for this event to be different. This is because one past event was God's ordering of the times, and this ordering logically entails that every future event occurs just as it does. So, given typical statements of causal determinism, the view of section two implies that the actions of persons are causally determined.

²⁴This isn't to say that transitivity is indisputable for causation. For resistance to the claim that the causal relation is transitive, see McDermott (1995), Hall (2004), and Ehring (1987).

²⁵For similar discussion, see "A Tale of Two Cronies" in Kvanvig (2012).

²⁶For a helpful overview of causal determinism, see Hofer (2010).

But where the objection goes wrong is in implicating that this makes any unique trouble for the present view—or any trouble at all for the present view. For, given this sort of definition of causal determinism, any view according to which God has infallibly true beliefs about the future in the past will imply causal determinism. Further, views like Molinism which have God willing things to happen which in turn, together with other past truths (specifically, the past truth of the subjunctives of freedom), entail the entire future, will likewise imply that all actions are causally determined. Indeed, even views which allow that propositions like <it *will be the case* that so-and-so does such-and-such> are true in the past will imply causal determinism. But this has not stopped libertarians, including Molinists, from attempting to account for the consistency of God's infallible foreknowledge and human free action. And the reason is because there is a problem here with the way in which causal determinism is being defined.²⁷ Causal determinism, on the definition above, doesn't say *anything* about causation. An action can be *causally* determined on this view without being *caused* at all. This consequence may be a welcome one where the definition of causal determinism is being used for certain purposes in the philosophy of physics where the concept of causation is given a dismissive attitude. But in discussions of free will, especially where agent-*causation* is in view, such a definition of causal determinism misses the point. The agent-causal theorist won't be worried about a causal determinism which only implies that agents' exercises of power are entailed by some past truth; what worries her is the sort of causal determinism which would imply that the causal buck never stops with the agent. And this is hardly implied by the account of foreknowledge sketched in section two.

Still, the objector may stubbornly insist, there's something worse about the view in section two than those other views when it comes to assessing whether a responsibility-undermining causal determinism is true. For, on those views it is a *soft* past fact which entails what happens in the future and on your view it is a *hard* past fact. We should think of causal determinism as saying that necessarily, the same *hard* past facts and laws of nature lead to the same future facts. Then, your view implies causal determinism and those other views don't.

The problem here is with the hard fact/soft fact distinction. It is difficult to see what would make facts like God's past beliefs about the future or the conjunction of his past willings of circumstances and the subjunctives of freedom any different from God's past willings to order the times in a certain way with respect to whether these count as hard or soft facts. On several ways of characterizing soft past facts—past facts which, in some sense, we can do something about—God's past willings concerning the order of the times are indeed soft past facts. For instance, they are facts over which we have counterfactual power.²⁸ Were an agent who in fact did A not to have done A, God's willings concerning the ordering of the times would have been different. Also, like other putatively

²⁷For a similar argument, compare Byerly (2011).

²⁸For accounts of soft facts which appeal to counterfactual power, see Saunders (1966) and Plantinga (1986).

soft facts, God's willings concerning the ordering of the times explicitly mention times other than past times; and on some accounts (e.g., Warfield (2010)) this is sufficient for those facts to be soft facts.²⁹ Further, it is plausible that what an agent agent-causes at a time t is in some appropriate sense *explanatorily* prior to God's willing to order the times as he does.³⁰ God orders the times as he does (at least in part) *because* of what agents do at those times.³¹ This sort of explanatory posteriority is likewise an initially attractive account of what makes a past fact a soft past fact.³² So, the objector here faces the difficult task of specifying what makes a past fact a hard past fact in such a way that God's past beliefs and willings concerning the circumstances he will bring about are not hard facts, but God's willings concerning the ordering of the times *are* hard facts. This, I submit, is a challenge which cannot be met.

We have seen no way to defend the view that the account of section two implies a responsibility-undermining causal determinism. But there is a powerful argument from its consistency with agent-causation for the claim that it *doesn't* imply a responsibility-undermining causal determinism. So, until there is some further reason brought forward for thinking that this view does imply a responsibility-undermining causal determinism, it is reasonable to conclude that it does not. At least, if the reasoning in this section is sound, then it is reasonable to conclude that *if* libertarian classical theism is consistent at all, then the model of the mechanics of foreknowledge presented in section two is consistent with the absence of causal determinism.

Perhaps, though, this does not yet imply that the view is consistent with the future actions of persons being performed *freely*. For, according to the libertarian, not only must free actions not be causally determined, they must also be such that those who perform them could have done otherwise. Indeed, some libertarians might object to the account of agent-causation used above precisely

²⁹Warfield's account here follows the tradition of Adams (1967) and Hoffman and Rosenkrantz (1984). Fischer (1992) calls this approach the "entailment criterion" approach and lists several other advocates of it.

³⁰There is a growing literature devoted to explanatory priority in the context of discussions of the freedom-foreknowledge problem, with no consensus about what explanatory priority involves [see, e.g., Adams (1991), Craig (1998), and Hasker (2000)]. I don't offer an account of explanatory priority here; I only suggest that in at least some intuitive sense, what persons do at times explains why God orders the times as he does.

³¹One might contend that while a proposition like $\langle S \text{ does } A \text{ at } t \rangle$ might explain why God orders t as he does, the defender of the view sketched here will have to be an eternalist if she wants S to be the (partial) truth-maker for $\langle S \text{ does } A \text{ at } t \rangle$. Finch and Rea (2008) argue for a conclusion like this. If they are correct, so be it: someone who wants to advocate the solution to the freedom-foreknowledge problem sketched here must be an eternalist. This is no special cost of the view advocated here, though, since other libertarian views like Molinism must say the exactly the same thing, if Rea's and Finch's argument succeeds.

³²This strategy is discussed explicitly in Fischer et al. (2009) and is defended in Freddoso (1983).

because it doesn't require that the agent could have done otherwise.³³ So, to complete a defense of the consistency of the account of the mechanics of foreknowledge sketched in section two with libertarian commitments concerning free action, it would be desirable to show that that view is consistent with agents having the ability to do otherwise. This I attempt in the next section.

The Ability to Do Otherwise

The question of whether the view that God knows the future by ordering the times is consistent with the ability of persons to do otherwise than they in fact do is quite similar to the question of whether the former view implies causal determinism. For, plausibly, the question of whether persons are able to do something other than they in fact do is a question about what persons can do holding certain facts fixed.³⁴ The hard question is just which facts to hold fixed. And discussion of this question parallels discussion of causal determinism in the previous section, since which facts we might hold fixed parallel which facts we might include as those which could lead to responsibility-undermining causal determinism.

Suppose, first, that we propose that a person is able to do otherwise than she does if, and only if, holding fixed *all facts whatsoever about the past and laws of nature*, it is possible for her to do otherwise. If this is what is required for a person to be able to do otherwise, then the view proposed in section two is not consistent with persons being able to do otherwise. But neither is Molinism nor any view which allows God to have infallible beliefs in the past nor any view which countenances past truths like <it *will be* the case that so-and-so does such-and-such>. For, all such views imply that there are some past truths which, whenever they are true, the future actions of persons cannot be different than they are. This, I take it, is a reason for the libertarian to *reject* the present account of what is required for someone to be able to do otherwise in the sense in which such an ability is required for free action, just like in the last section the fact that some definitions of causal determinism implied that the actions of persons must be causally determined was a reason for the libertarian to reject the claim that the absence of *this* kind of causal determinism was required for free action. In order for a person's action to be free, it must be that she was able to do otherwise—but she needn't be able to do otherwise *holding fixed all truths whatsoever* about the past and laws.

So, we might try to restrict which past facts must be held fixed. Perhaps we should suggest that only the hard facts, only those facts which, intuitively, we have no control over, be held fixed. Thus, a person is able to do otherwise if, and only if, holding fixed all facts over which she has no control, it is possible that

³³Contrast, e.g., the account of agent-causation developed in Rowe (2006) with that of Bergmann and Cover (2006). Rowe requires explicitly that the agent have the power to do otherwise.

³⁴This is the "restricted possibility" view of ability. For defense, see Lewis (1983) and Lehrer (1976). Fischer (1992) explicitly uses the location of "holding fixed" certain facts when determining whether someone has the requisite sense of ability. For discussion, see Maier (2010).

she does otherwise than she in fact does. But, given what we saw in the previous section, it is quite difficult to assess how this suggestion would cast special doubt on the consistency of the account of section two and the claim that persons can do otherwise than they do. For, it would seem like whether or not facts about what God wills about the relations between times are hard or soft facts stands or falls with whether or not facts about God's past beliefs or past willings of circumstances are hard or soft facts. Thus, if the view in section two is inconsistent with persons having the sort of ability to do otherwise which is required for free action, then it is impossible for God to have infallible foreknowledge of free actions and surely the Molinist proposal doesn't successfully show how he could have such. Conversely, if it *is* possible for God to have infallible foreknowledge of free actions and if the Molinist proposal *is* a consistent proposal about how God could have such infallible foreknowledge, then so is the proposal of section two. This should be enough motivation for the libertarian to conclude that the proposal of section two is consistent with persons having the ability to do otherwise which she requires for free action.

And perhaps even more can be said on behalf of the conclusion than the proposal of section two is consistent with persons having the sort of ability to do otherwise required for free action. Specifically, we may attempt a direct argument for this conclusion much like the direct argument for the compatibility of the view of section two with the absence of responsibility-undermining causal determinism from the previous section. For, plausibly, this ability to do otherwise can be captured by something like the following account:

S did A and *was able to do other than A* just when (i) S exercised a basic power of hers to do A so as to bring about A, (ii) nothing distinct from S caused S to exercise this power, (iii) S had a basic power to do something other than A, and (iv) nothing with which S didn't identify would have prevented S from exercising this power, were she to have attempted to exercise it.

Such an account is plausible because it yields the correct results about whether a person could or could not have done otherwise in a wide range of cases. Persons who have no basic power at all to do otherwise than they do are not able to do otherwise, on this view. Persons like the unwilling addict—the addict who doesn't identify with his addiction, but rather dis-identifies with it—whose psychological problems would prevent them from exercising powers they do have to perform actions are not able to perform those actions, on this view.³⁵ Persons who are externally constrained in such a way that they cannot exercise powers they have to perform actions are not able to perform those actions, on this view.³⁶ And all of these results seem entirely accurate, since we tend to think that the persons just described don't perform the actions (or omissions) they

³⁵These cases play a significant role in discussions of the ability to do otherwise in the literature. See, e.g., Howard-Snyder (2008). For more on the notion of what a person "identifies with," see Frankfurt (1971) and the more expanded discussion of "Real Self" views in Wolf (1990).

³⁶I am thinking here of Frankfurt-style cases. See, e.g., Frankfurt (1969).

perform freely because they weren't able to do otherwise in the sense required for freedom.

But suppose that this account of the ability to do otherwise is correct. If so, then it is clear that the account of the mechanics of infallible divine foreknowledge put forward in section two is consistent with persons having the ability to do other than what they do. We saw in the previous section that this account is consistent with conditions (i) and (ii); these are just a condensed version of conditions (i)-(iii) in AC, Bergmann and Cover's account of agent-causation. Further, there is no inconsistency in letting times entail of a person who has a power to perform an action A that she also has a power to do something other than A. So, condition (iii) here is satisfied. And, there is nothing about the account from section two which requires that there is something with which an agent doesn't identify which would prevent her from exercising a power she does have, were she to attempt to exercise it instead of the one she in fact exercises. So, all of (i)-(iv) may be satisfied by the account of section two. Thus, if the account of *able to do otherwise* presented above is correct, then the account of the mechanics of divine foreknowledge presented in section two is consistent with the ability of persons to do otherwise.

Now, I don't put too much stock in the above account of the ability to do otherwise. Analyzing this concept has proved extremely difficult, and may even be impossible.³⁷ Offering an analysis of it is, according to some, tantamount to offering an account of free action itself.³⁸ And I am very skeptical that the above account offers a full account of free action which is entirely counterexample free. I do think it plausible, however, that such an account may provide a useful guide to when a person has the ability to do otherwise required for freedom. In any event, I would suggest that, in the absence of some compelling reason to think that the account of section two is inconsistent with persons having the ability to do otherwise, the reasons here adduced for the consistency of these two positions are sufficient to motivate further interest in that account. We have found no persuasive argument that the account of section two faces a conflict with the ability to do otherwise not also faced by Molinism and libertarianism more generally, and we have seen an argument that, if we follow a useful guide concerning what is required for the ability to do otherwise, then the account of section two is consistent with persons having the ability to do otherwise.

Conclusion

In this paper, I have sketched an account of the mechanics of infallible divine foreknowledge and argued that this account is consistent with the absence of causal determinism and the presence of the ability of persons to do otherwise than what they in fact do. Or, perhaps more accurately, I have argued that *if* any

³⁷Howard-Snyder (2008), for example, takes the concept as primitive; the best we can do is ostend to instances where it is clearly absent and say that what we mean by "able to do otherwise" is whatever is *lacking* there.

³⁸van Inwagen (2008) counsels that we think of the problem of free will in this way.

account which accepts that God has infallible foreknowledge of free actions, that causal determinism is false, and that persons have the ability to do otherwise is consistent, then the account proposed here is one such account.

I do not argue that the account sketched here is superior to other accounts which would also attempt to show how these three views are consistent. Though it does have some advantages over other such accounts,³⁹ it also has its own difficulties. Chief among these will be difficulties having to do with the metaphysics of time presupposed by the account. My contention here is that a view like that I have sketched is worthy of further investigation to determine whether such difficulties are ultimately telling against it or not. And this is as much as someone ought to hope for, I would think, in an initial statement of an account of the mechanics of divine foreknowledge.

I want to close by emphasizing that there is actually a variety of views like that I have sketched here which are worthy of attention. As I said in section two, it may be worthwhile for someone to investigate whether a view like that discussed here which appealed to concrete rather than abstract times could also maintain the consistency of libertarian classical theism. Further, as I also mentioned there, it may be that an account along the lines of that discussed here which allowed times to be essentially fully maximal, and not just essentially only *nearly* maximal, could work just as well as the account discussed here. This too deserves attention. So I invite friends and foes of the sort of view presented here to consider whether there is a workable view in the neighborhood which could give so many what they want—consistent libertarian classical theism.

REFERENCES

- Adams, Marilyn. 1967. "Is the Existence of God a "Hard" Fact?" *The Philosophical Review* 76.4: 492–503.
- Adams, Robert M. 1974. "Theories of Actuality." *Noûs* 8: 211–221.
- Adams, Robert M. 1991. "An Anti-Molinist Argument." *Philosophical Perspectives* V: 343–353.
- Bergmann, Michael and Jan Cover. 2006. "Divine Responsibility without Divine Freedom." *Faith and Philosophy* 23: 381–408.
- Bourne, Craig. 2006. *A Future for Presentism*. Oxford: Clarendon Press.
- Byerly, T. Ryan. 2011. "Ockhamism vs Molinism, Round Two: A Response to Warfield." *Religious Studies* 47 (4): 503–511.
- Chisholm, Roderick. 1976. *Person and Object*. LaSalle: Open Court.
- Chisholm, Roderick M. 1979. "Objects and Persons: Revisions and Replies." *Grazer Philosophische Studien* 7 (8): 317–388.
- Craig, William Lane. 1990. *Divine Foreknowledge and Human Freedom*. Leiden: E. J. Brill.

³⁹With reference to Molinism, for instance, the present account doesn't require conditional excluded middle nor does it face a grounding objection like Molinism does.

- Craig, William Lane. 1998. "On Hasker's Defense of Anti-Molinism." *Faith and Philosophy* 15.2: 236–240.
- Crisp, Thomas M. 2007. "Presentism and the Grounding Objection." *Noûs* 41 (1): 90–109.
- Davidson, Matthew. 2003. "Presentism and the Non-Present." *Philosophical Studies* 113: 77–92.
- Davidson, Matthew. 2004. "Critical Notice of Theodore Sider, Four Dimensionalism." *Philosophical Books* 45: 17–33.
- Ehring, Douglas. 1987. "Causal Relata." *Synthese* 73: 319–328.
- Finch, Alicia and Michael Rea. 2008. "Presentism and Ockham's Way Out." In *Oxford Studies in Philosophy of Religion*, edited by Jonathan L. Kvanvig. Vol. I. Oxford: Oxford University Press.
- Fine, Kit and Arthur Prior. 1977. *Worlds, Times, and Selves*. Amherst: University of Massachusetts Press.
- Fischer, John Martin. 1992. "Recent Work on God and Freedom." *American Philosophical Quarterly* 29 (2): 91–109.
- Fischer, John Martin, Patrick Todd, and Neal Tognazzini. 2009. "Re-reading Nelson Pike's 'Divine Omniscience and Voluntary Action'." *Philosophical Papers* 38 (2): 247–270.
- Flint, Thomas P. 1998. *Divine Providence: The Molinist Account*. Ithaca: Cornell University Press.
- Frankfurt, Harry G. 1969. "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy* 66: 829–839.
- Frankfurt, Harry G. 1971. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68: 5–20.
- Freddoso, Alfred J. 1983. "Accidental Necessity and Logical Determinism." *Journal of Philosophy* 80: 257–278.
- Hall, Ned. 2004. "Two Concepts of Causation." In *Causation and Counterfactuals*, edited by John Collins, Ned Hall, and L. A. Paul. Cambridge: MIT Press.
- Hasker, William. 1995. "Middle Knowledge: A Refutation Revisited." *Faith and Philosophy* 12.2: 223–236.
- Hasker, William. 2000. "Anti-Molinism is Undefeated!" *Faith and Philosophy* 17.1: 126–131.
- Hofer, Carl. 2010. "Causal Determinism." *The Stanford Encyclopedia of Philosophy*. <http://plato.stanford.edu/entries/determinism-causal/>.
- Hoffman, Joshua and Gary Rosenkrantz. 1984. "Hard and Soft Facts." *Philosophical Review* 93 (3): 419–434.
- Howard-Snyder, Daniel. 2008. "The Puzzle of Prayers of Thanksgiving and Praise." In *New Waves in Philosophy of Religion*, edited by Yujin Nagasawa and Erik Wielenberg. New York: Palgrave-McMillan.
- Hunt, David. 1993. *The Providence of God*. Leicester: Intervarsity Press.

- Kriegel, Uriah. 2009. *Subjective Consciousness: A Self-representational Theory*. Oxford: Oxford University Press.
- Kvanvig, Jonathan L. 1986. *The Possibility of an All-Knowing God*. New York: Saint Martin's Press.
- Kvanvig, Jonathan L. 2002. "On Behalf of Maverick Molinism." *Faith and Philosophy* 19.3: 348–357.
- Kvanvig, Jonathan L. 2012. "A Tale of Two Cronies." In *Destiny and Deliberation: Essays in Philosophical Theology*. Oxford: Oxford University Press.
- Lehrer, Keith. 1976. "'Could' in Theory and Praxis: A Possible Worlds Analysis." In *Action Theory*, edited by Myles Brand and Douglas Walton. Dordrecht, Holland: D. Reidel.
- Lewis, David. 1983. *Philosophical Papers*. Vol. I. Oxford: Oxford University Press.
- Lewis, David. 1996. "Elusive Knowledge." *Australasian Journal of Philosophy* 74.4: 549–567.
- Maier, John. 2010. "Abilities." *The Stanford Encyclopedia of Philosophy*. [Http://plato.stanford.edu/entries/abilities/](http://plato.stanford.edu/entries/abilities/).
- Markosian, Ned. 2012. "Agent Causation as the Solution to All the Compatibilist's Problems." *Philosophical Studies* 157: 383–398.
- McDermott, Michael. 1995. "Redundant Causation." *British Journal for Philosophy of Science* 40: 523–544.
- Menzel, Christopher. 1990. "Theism, Platonism, and the Metaphysics of Mathematics." In *Christian Theism and the Problems of Philosophy*. Library of Religious Philosophy 5. Notre Dame: University of Notre Dame Press.
- Morris, Thomas. 1987. *Anselmian Explorations*. Notre Dame, IN: University of Notre Dame Press.
- O'Connor, Timothy. 2000. *Persons and Causes: The Metaphysics of Free Will*. Oxford: Oxford University Press.
- Plantinga, Alvin. 1974. *The Nature of Necessity*. Oxford: Clarendon Press.
- Plantinga, Alvin. 1986. "On Ockham's Way Out." *Faith and Philosophy* 3: 235–269.
- Prior, Arthur. 1996. "A Statement of Temporal Realism." In *Logic and Reality: Essays on the Legacy of Arthur Prior*, edited by B. Jack Copeland. Oxford: Clarendon Press.
- Rowe, William L. 2006. *Can God Be Free?* New York: Oxford University Press.
- Rudder Baker, Lynne. 2003. "Why Christians Should Not Be Libertarians: An Augustinian Challenge." *Faith and Philosophy* 20: 460–478.
- Saunders, John T. 1966. "Of God and Freedom." *Philosophical Review* 75: 219–225.
- Schaffer, Jonathan. 2007. "The Metaphysics of Causation." *The Stanford Encyclopedia of Philosophy*. [Http://plato.stanford.edu/entries/causation-metaphysics/](http://plato.stanford.edu/entries/causation-metaphysics/).

- Stump, Eleonore and Norman Kretzmann. 1991. "Prophecy, Past Truth, and Eternity." *Philosophical Perspectives* 5: 395–424.
- Swinburne, Richard. 2004. *The Existence of God*. New York: Oxford University Press. 2nd edn.
- Timpe, Kevin. 2006. "A Critique of Frankfurt-Libertarianism." *Philosophia* 34: 189–202.
- van Inwagen, Peter. 2008. "How to Think about the Problem of Free Will." *Journal of Ethics* 12: 327–341.
- Viney, Donald. 1989. "God Only Knows? Hartshorne and the Mechanics of Omniscience." In *Hartshorne: Process Philosophy and Theology*, edited by Robert Kane and Stephen Phillips. Albany, NY: SUNY Press.
- Warfield, Ted. 2010. "Ockhamism and Molinism—Foreknowledge and Prophecy." In *Oxford Studies in Philosophy of Religion, Volume II*, edited by Jonathan L. Kvanvig. Oxford: Oxford University Press.
- Wolf, Susan. 1990. *Freedom Within Reason*. Oxford: Oxford University Press.
- Zalta, Edward N. 1983. *Abstract Objects: An Introduction to Axiomatic Metaphysics*. Dordrecht, Holland: D. Reidel.
- Zalta, Edward N. 1987. "On the Structural Similarities between Worlds and Times." *Philosophical Studies* 51: 213–239.

EVIL AND THE PROBLEM OF ANOMALY

Trent Dougherty and Alexander R. Pruss

1. The Problem of Anomaly

As Kuhn's *Structure of Scientific Revolutions* (Kuhn 1970) has made particularly evident, scientists often maintain their theories in the face of anomalies or apparent counterexamples. There is good pragmatic justification for such "stubbornness". Belief in a theory is motivating. If theories were abandoned at the first sight of anomaly, motivation for further development of the theory—development that at least sometimes overcomes the anomalies—would often be insufficient, and the progress of science would be severely hampered.

But it seems to be epistemically irrational to hold on to a scientific theory in the face of an apparent counterexample: The theory implies that every F is a G . We are faced with an apparent case of an F that is not a G . Hence, apparently, the theory is false. If the Popperian line of reasoning (Popper 1959, 86–87) were followed, that would spell the end of that theory. One might conclude from this that by holding on to theories in the face of anomalies scientists sacrifice their private epistemic good¹ for the sake of the good of the discipline.

Such a view would, however, lead to an unacceptable scepticism about too many major scientific theories. For major scientific theories do tend to have anomalies, as reflective and honest scientists will admit. (Indeed, Kuhn claims that every theory always faces puzzles.²) This, then, is the problem of anomaly: *How can we epistemically rationally avoid scepticism about scientific theories that are subject to apparent counterexample?*

We shall suggest an answer, discussing the difference between mere anomaly mongering and seriously disconfirmatory use of anomalies. And our suggestion will have corollaries for the inductive problem of evil. Influential formulations of the problem, such as Rowe's fawn-type case (Rowe 1979) and Tooley's focus on the Lisbon earthquake (Plantinga and Tooley 2008), are relevantly like the offering of anomalies for scientific theories. The theist can respond to such anomalies

¹Some (perhaps Kuhn himself, but we do not assert a view about that) do not conceive of scientific progress in terms of *epistemic* goods at all. Here, we identify with a strongly realistic tradition, from which our investigation in this paper take place.

²"To be accepted as a paradigm, a theory must seem better than its competitors, but it need not, and in fact never does, explain all the facts with which it can be confronted" (Kuhn 1970, p. 12) "no paradigm ever solves all the problems it defines" (Kuhn 1970, p. 109). We have focused on the strongest kind of failure to explain: the apparent counter-instance. But there are degrees of failure to "fit in" more generally.

in ways that relevantly parallel how an evolutionary theorist should respond to many of the challenges raised by creationists and Intelligent Design advocates. Part of what this involves is the thesis that it is not necessary for rational retention of religious belief in light of the facts about evil that one be able to offer a theodicy in the sense of a story which claims to report God's actual reasons for allowing the evil, but that, rather, it is sufficient to be able to tell a story that is sufficiently plausible (how plausible will be the subject of some discussion below), and sometimes no story need be told at all. In this regard, our thesis is like standard sceptical theism. However, we will avoid the difficulties in containing the scepticism that have plagued sceptical theism. Our account is designed, after all, to *block* a sceptical move. It is our opponents who will have trouble avoiding scepticism about theories such as Darwinian evolution. Finally, we shall discuss how the discussion can move forward from Rowe-style inductive arguments from evil, by offering suggestions on what the arguer from evil would need to do to transcend anomaly mongering.

To illustrate our case, we shall have to use the familiar method of appealing to an artificially simple case. This cannot be avoided without greatly complicating the presentation. We are confident, however, that what we say can be extended to serious examples from the history of science. One example which Kuhn (1996, 39) uses to illustrate rational theory retention in the face of anomaly is the retention of Newton's inverse square law in light of the progression of the apogee of the moon's orbit. Newton's calculations only predicted about half of the progression which was observed. For over half a century, physicists tried to figure out what was wrong, but no serious consideration was given to revising the inverse square law. Part of why this is so is surely because it was to be expected that a theory of such scope would have many, many places where there would be apparent discrepancies.

2. Scientific theories and anomalies

2.1 How not to argue from anomalies

Whenever a theory T uncontroversially entails the universal generalization

$$(1) \forall x(Fx \supset Gx),$$

an *anomaly* for T is an a such that it seems that Fa but it does not seem that Ga .³ The locution "it does not seem that Ga ", however, is ambiguous between the weak reading that it is not the case that it seems that Ga and the strong reading that it seems that $\sim Ga$. These two readings, then, define weak and strong anomaly respectively. If our theory says that all ravens are black, then an apparent raven that looks non-black is a strong anomaly and an apparent raven that fails to look black is a weak anomaly. How can a raven fail to look black without looking non-black? A trivial case is where the raven is in a dark place where it cannot be seen at all. Or there might be some light, but not enough

³If necessary, we can assume that the conjunction $\text{Seems}(x, F) \& \text{Seems}(x, \sim G)$ entails $\text{Seems}(x, F \& \sim G)$.

for the raven to look any particular dark colour—the raven will then look dark, rather than black, non-black, midnight blue or non-midnight-blue.

Merely weak anomalies present no challenge to a theory unless there is reason to think that if a satisfied G , then this would probably be seen.⁴ In any case, if we can show that it can be epistemically rational to retain a theory in light of a *strong* anomaly, it will *a fortiori* follow that it is so for weak analogies. We shall therefore, restrict consideration to strong anomalies, and use the term “anomaly” without further qualification to mean “strong anomaly”.

Strong anomalies, however, appear to not only incrementally but absolutely disconfirm a theory. After all, the following appears to be a cogent non-deductive argument:

- (2) If T is true, then $\forall x(Fx \supset Gx)$.
- (3) It seems that Fa .
- (4) It seems that $\sim Ga$.
- (5) Therefore, $\sim \forall x(Fx \supset Gx)$. (Ampliatively from (3) and (4))
- (6) Therefore, $\sim T$. (By *modus tollens* from (2) and (5))

This is the Basic Anomaly Argument (BAA).

However, BAA is dubious. Consider as our toy example the theory T that every fox is a canid. Now, the world contains many foxes. Some foxes more obviously look like canids than others. Suppose that we want to shake someone’s belief that all foxes are canids, and so we search through all the world’s foxes and pick out the most feline-looking fox there is, call it Felix. It could then be true both that Felix seems to be a fox (Felix might be the offspring of two obvious foxes and have the usual high degree of DNA match to its parents) *and* that Felix seems to be a felid rather than a canid.

It would, however, be fallacious to conclude that Felix disconfirms T . The reason is that there was a selection bias in the choice of Felix.⁵ We went out looking for the most feline-looking fox we could find, and that we found one should have been no surprise given the background information that foxes are arranged in a broad spectrum of felinity of appearance.

In fact, if finding Felix disconfirms T , we could have obtained significant disconfirmation for T simply by armchair investigation based on uncontroversial

⁴In fact, in some cases a merely weak anomaly can provide a small degree of confirmation *for* the theory. For instance, if it is part of the theory of electrons that electrons are too small to see with optical microscopes, the failure of optical microscopic examination to reveal any particles pushing the pieces of dust in a Brownian motion experiment provides a small amount of confirmation of the theory of electrons by disconfirming the alternate hypothesis that the larger bits of dust are moved around by microscopically observable dust particles. That anomalies can count as evidence for a theory is perhaps the mirror image of Swinburne’s 1971 point that an instance of a generalization can count as evidence against it. A grasshopper just barely inside the Andrew County line, disconfirms the generalization that all grasshoppers are in Andrew County.

⁵We bracket here discussion about the way *prima facie* seemings are incorporated into *ultima facie* seemings. A sheep or a sheep dog could seem to be a sheep in some respects and seem to be a dog in others. We leave open the question how or whether these can be combined into an overall seeming state.

common knowledge. Uncontroversial common knowledge tells us that there is a large degree of variation of appearance among foxes, and that foxes are not all that different from cats in appearance. This gives us good reason to think that there is some fox that looks more like a felid than like a canid. Thus, there probably is a fox that to a competent observer looks like it is not a canid. By parallel reasoning to the Felix case, that would give us significant evidence against *T*. But that is absurd.

Similarly, we can expect more serious scientific theories that entail non-trivial universal generalizations about large numbers of instances also to have strong anomalies. We know both that our perceptual faculties are not perfectly reliable and that most natural kinds, except those defined at the molecular and lower levels, exhibit significant random variation. Suppose, then, that our skill at visual identification of canids gives us a false negative in 0.01% of the cases. This means that for every 10,000 foxes that are canids, we should expect one that looks to us like a non-canid. Hence, if there are, say, a 100,000 foxes under observation it should, if anything, be surprising if we don't find one that looks like a non-canid, whether or not all foxes are canids.

For a less trivial example, consider Naturalistic Evolution (NE), the claim that all currently extant non-domesticated species on earth gradually developed from a unicellular ancestor by random mutation under non-intentional selective pressures. (Thus, views on which some non-domesticated species developed under selective pressures intentionally arranged by aliens or God do not qualify.) NE implies the universal claim that for every currently extant non-domesticated species on earth there is a gradual chain of random mutations and non-intentional selective pressures that produced that species from a unicellular ancestor. But over a million non-domesticated multicellular extant species have been identified.⁶ Given innate limitations of fossil data, the great amount of variation between species, as well as the limitations of our cognitive faculties, it is entirely unsurprising given NE that there are non-domesticated species which, even after significant investigation, will seem not to be the product of a gradual chain of random mutations and non-intentional selective pressures tracing back to a unicellular ancestor. Therefore, when creationists and Intelligent Design (ID) advocates point out the fact that there exist such species, they are pointing out something that should be no surprise to NE theorists. They are merely engaging in anomaly mongering. This is especially clear when, as is not uncommon, the examples were chosen specifically after "sifting" for the purpose of providing an apparent counterexample to NE, an obvious case of selection bias. While the particular identification of such species may require empirical work, the expectation that there *be* such species can be generated by mere armchair biology.

Why, then, does BAA seem so plausible? How can it be that it seems like we have a counterinstance, and yet this is not significantly disconfirmatory of

⁶Bisby et al. (2011) list 1,059,854 species in Kingdom Animalia and 238,386 species in Plantae. Only a very small proportion of species is domesticated.

the theory? On the face of it, this is very hard to understand. One explanation is that there are two senses of “seems”. If we understand the “seems” as an all things considered epistemic seeming, where among the things considered is our evidence for T as well as the entailment from T to the non-existence of a counterinstance, then of course if it seems to us that we have a counterinstance, that is a serious problem for T . But this is an uninteresting observation: when the total evidence on balance supports something that uncontroversially entails $\sim T$, then T is evidentially ruled out for belief.

On the other hand, if we understand the “seems” in BAA as what appears to us when we bracket T and the evidence for T , then the ampliative inference in BAA is unwarranted. The question of whether an anomaly makes it epistemically irrational to believe in a theory depends on what an anomaly is. Is it a case of something that *all things considered* seems to be a counterinstance to the theory or a case of something that seems to be a counterinstance to the theory when we bracket the theory and its evidence?⁷

That the precession of the perihelion of Mercury is an anomaly for Newtonian physics is a claim independent of what the evidence for Newtonian theory is, for it is simply an apparent counterinstance to a generalization (regardless of what evidence we have for or against that generalization). Even if the evidence for Newtonian theory were overwhelming, the precession would still be an anomaly. This is so even when warrant transfers from the theory to the negation of the anomaly claim, providing a rebutting defeater for the evidence for the anomaly claim.⁸ An undercutting defeater would put us at ease by removing the appearance⁹ of counterexample, but a rebutting defeater leaves the troubling appearance of counterexample. Thus, we can take anomalies to be things that appear to be counterinstances when we bracket the evidence for the theory.

The next section will regiment the above line of thought by offering a formal account of when anomalies have force and when they do not. This will allow for a more careful account of anomaly mongering as well.

2.2 Anomalies and Anomaly Mongering

If we understand anomalies as *apparent* counterinstances or cases that look like counterinstances independently of the evidence for the theory, then it becomes clear that an anomaly does not render belief in the theory irrational. It is quite normal to have an instance a of F such that our only reason for believing that a is G is that we accept a theory T that entails that every instance of F is an instance of G , and it can even be the case that had we not accepted such a theory, we would have reasonably thought that a is not G . For instance, we can be looking under a

⁷This approach may seem similar to one advocated by Wykstra (1984, p. 91). In fact, it might seem to be an application of it. However, we have no intention here of asserting some general principle of confirmation in this case.

⁸And thus is our approach in the application in Section 3 different from the “Moorean shift” strategy mentioned by Rowe (1979) and applied by Geivett (1995).

⁹We postpone a discussion of an ambiguity here until its application in Section 3.

microscope at a properly functioning cell. The cell might be look to us like it has no nucleus, but we could, and often should, argue: “Properly-functioning cells of this type have nuclei; this is a properly-functioning cell of this type; it looks like it doesn’t have a nucleus; hence, it must have a nucleus that we somehow overlooked.” After all the probability of the theory that properly-functioning cells of this type have nuclei is presumably *much* greater than the probability of the universal generalization that we would never overlook a nucleus.

If we reject this line of thought, insisting that belief in the theory is irrational in cases where absent the theory we would have reasonably thought that we have a counterexample, then scepticism about many major scientific theories ensues. Many major scientific theories have consequences that we would have rationally disbelieved had the theory not predicted them. The table in front of me appears as if it had no vacuum in it; it is only the atomic theory of matter that would give me reason to think that it is mostly vacuum. The mammalian eye appears¹⁰ directly designed by an engineer; it is only evolutionary theory that gives us reason to think that it evolved without any engineer being directly involved.

The mere fact that theory *T* is subject to anomaly is typically unsurprising, and does not constitute significant evidence against *T*. In fact, the non-existence of anomalies for a wide-ranging theory might make us suspicious that the theory is empirically trivial.

At the same time, this line of thought needs to be tempered by the fact that anomalies can pile up in such a way that the theory is no longer tenable, and in principle a single but particularly telling anomaly might by itself do that. Criticizing a theory by pointing out the mere fact that there are anomalies for the theory is a sign of mere anomaly mongering. But there is more that one can do with anomalies than monger them.

Go back to the fox case. We supposed that the false negative rate for canid identification was 0.01%, and pointed out that then we would expect about one in ten thousand foxes to look like a non-canid. But what if we observed one in a thousand foxes looking like a non-canid? Or what if we observed one in ten thousand foxes not only *looking* like a non-canid but *smelling* like one, whereas the false negative rate for combined sight plus smell identification of canids was 0.001%? In both cases, we would be observing significantly more apparent counterinstances than we would expect if the theory that all foxes are canids were true.

In cases like this, sampling methods are important. If, for instance, our sampling method favoured those foxes that have a feline look¹¹, finding that a mod-

¹⁰Evolutionary naturalists often feel no threat in admitting that some biological features have an “appearance of design.” “Biology is the study of complicated things that give the appearance of having been designed for a purpose” Dawkins (1996, p. 1). “Natural selection is the blind watchmaker, blind because it does not see ahead, does not plan consequences, has no purpose in view. Yet the living results of natural selection overwhelmingly impress us with the illusion of design and planning.” Dawkins (1996, p. 21).

¹¹One of the authors generated such a sample by appropriately tailored Internet searches.

erately large proportion of the sampled foxes looked like non-canids could be statistically insignificant. And this shows why it was that we should be entirely unmoved¹² by the non-canid appearance of Felix, the most feline-looking fox in the world. For, in the initial story, we found Felix precisely by looking for feline-looking foxes.

This observation has an important consequence for intellectual integrity in scientific research. Integrity in research requires that if one is trying to disconfirm T , one cannot simply look for the most glaring anomalies and present those as one's evidence. Ideally one should define a selection procedure for one's cases ahead of time and make sure that if the procedure is biased in favour of the production of apparent counterinstances, the statistical significance of these apparent counterinstances is discounted for the bias. And if one did not define a selection procedure ahead of time, but tried several selection procedures until one got one that produced the result, then one should adjust the significance of the results based on how many degrees of freedom one had (Simmons et al. (2011) and Abdi (2007)). Moreover, it is crucial that all the non-anomalous cases that one found along the way be factored in as evidence for T . If after looking at ten thousand foxes, one finally found one that looked non-canine, one needs to take into account the evidence in favour of T provided by the 9999 foxes that *did* look canine.

Anomalies can pile up to make a theory untenable. But the piling up needs to be carefully considered statistically. Finding that there are *many* apparent counterexamples to a theory need not be significant, unless one takes into account rates of potential experimental error. For instance, if we found a thousand galaxies that appear incompatible with a theory about galactic development, that might not be significant if we had to examine a million galaxies to come up with these and the rate of relevant observational or analytical error was one in a thousand.

But quantity is not the only way that an anomaly can be genuinely problematic for a theory. We could have a single case anomaly that makes a theory untenable. Imagine that we find Felix, the most uncanine-looking of foxes, and we continue to study Felix using multiple experimental methods, and to *all* of the methods, Felix looks both like a fox and a non-canid. At this point, things start looking bad for the theory. But again we cannot avoid using statistics. There is, presumably, a non-zero false-negative rate r for the combination of the experimental methods for identifying canid character. And we have to take into account just how many foxes we had to examine to find Felix. If the number of foxes we had to examine to find Felix was of the order of magnitude of $1/r$, then it is not particularly surprising that there was such a fox on the theory that all foxes are canids. But if we examined significantly fewer than $1/r$ foxes, then indeed we have strong evidence against the theory.

¹²Note that this is a normative claim. In a perfectly ordinary sense of "look" it would still be true to say that it doesn't look like a fox.

Moreover, a very large quantity of token anomalies need not significantly disconfirm when there are many *types* of cases and all the token anomalies fall under only one of the types. It could well be that there is only one type of F which appears to be $\sim G$, but that that type has many tokens. If being F is anything like a natural kind, then the explanatorily and statistically salient level for counting may be the genus. So suppose we return to the site where we found Felix and found a large enough community of Felix-type foxes to designate a new subspecies of fox. The discovery of this new subspecies would not provide significantly more disconfirmation of T than the discovery of Felix alone.

These examples raise a serious question. How are we to tell whether the existence of an anomaly is in fact a problem for a theory? The Bayesian has a simple answer. Let T be the theory and let E_n be the proposition that we have found at least one anomaly of some type (e.g., a fox that looks and smells like a canid) among n cases that we have examined. Then, we can measure the degree of confirmation that E_n provides to $\sim T$ by, e.g., the Bayes factor¹³ $P(E_n | \sim T) / P(E_n | T)$: when this ratio is much smaller than one, we have significant disconfirmation of $\sim T$ (i.e., confirmation of T); when it is close to one, we have no significant confirmatory or disconfirmatory effect; and when it is much greater than one, we have significant confirmation of $\sim T$ (i.e., disconfirmation of T). Because $P(E_n | \sim T) \leq 1$, when $P(E_n | T)$ is very close to 1, the degree of confirmation to $\sim T$ will not be much greater than 1, if it will be greater at all. When we expect an anomaly given T , the fact that we find one is not much evidence, if at all, against T .

Suppose that the n cases we have examined are conditionally independent given T and conditionally independent given $\sim T$ —this is the situation we get in independent testing of a hypothesis.¹⁴ If T is true, no case will be a counterexample to T . But, nonetheless, due to observational error, there is some small probability $\varepsilon > 0$ that the case will *seem* to be a counterexample, i.e., that there will be an anomaly. The probability that at least one case seems to be a counterexample is then: $P(E_n | T) = 1 - (1 - \varepsilon)^n$. In other words, the probability of an anomaly increases exponentially with the number of cases observed. When n is significantly greater than $1/\varepsilon$, the amount of disconfirmation provided by the mere existence of an anomaly will be insignificant. And if the prior probability of T is high, then in those cases the reasonable thing to say is not that the anomaly significantly disconfirms T , but that the prior evidence for T makes it likely that the anomaly is only *apparently* a counterexample for T . On the other

¹³For our purposes, other plausible measures of confirmation will give similar results.

¹⁴One might initially think that independent tests for hypotheses require that the outcomes of the different tests be independent *simpliciter* rather than independent conditionally on the hypothesis and conditionally on the negation of the hypothesis. But that is incorrect. Suppose I have a coin, and a hypothesis H that it is loaded so that it has probability of $2/3$ of coming up heads. Suppose the one alternative is that it is a fair coin. Then each toss of the coin counts as an independent test of the hypothesis, and indeed the outcomes of the tosses are independent given H as well as independent given $\sim H$. However, the outcomes of the tosses are not independent given one's full subjective probability measure.

hand, if ε is of the order of magnitude of $1/n$ or smaller—i.e., when our observational judgment that something is actually a counterexample to T is reliable in comparison to the number of trials—then $P(E_n|T)$ will be significantly less than one, and the Bayes factor $P(E_n|\sim T)/P(E_n|T)$ may be significantly greater than one. For instance, if $\varepsilon=1/n$, and n is large, then $P(E_n|T)$ will be asymptotically equal to $1-1/e \approx 0.632$, and if $P(E_n|\sim T)$ is close to 1, then the Bayes factor will be close to 1.6, resulting in a moderate probability boost to $\sim T$. Jeffreys (1961, p. 432) has a Bayes factor of 1.6 as the boundary between “barely worth mentioning” and “substantial”.

If we observe a greater proportion of anomalies in our data set, then the evidence against T will be stronger. But it still matters just how large the error rate ε is. After all, among n data points, if our false positive rate in identifying counterexamples to T is ε , we would expect to see about $n\varepsilon$ anomalies if T is true.

Now that we have sketched the evidential dynamics of anomalies, we will consider a number of ways of responding to them, over and above leveling the charge of anomaly mongering, yet short of successfully incorporating the (now former) anomaly into a theory.

2.3 Responding to anomalies

2.3.1. The bare response Pseudoscientific objections to scientific theories, such as some of the objections to evolution, sometimes resort to mere anomaly mongering. The pseudoscientist examines a large number of cases until she finds an especially anomalous one, and typically does so without any statistical care. She then presses the especially anomalous case that she has found. The scientist, in turn, is fully within her epistemic rights to be unimpressed when the selection procedure that led to the anomalous case has not been specified in ways that make for statistical significance.

In the preceding section we saw that if an anomaly is found by testing randomly chosen cases, that is only significantly bad news for the theory if the probability ε that a non-counterexample to the theory would look like a counterexample (we might add: of this degree of anomalousness, if we have a way of quantifying the degree of anomaly) is of the order of magnitude of $1/n$, where n is the number of cases tested, or smaller.

However, not only do typical pseudoscientists not give us estimates of ε and a value for n , but their procedure is typically methodologically even worse. Rather than choosing cases uniformly at random, they are most likely consciously or unconsciously employing a non-random (or at least not uniformly random¹⁵) search algorithm designed for finding anomalies. Perhaps they start by looking at cases which, while apparently fitting the theory, do so with a strain. Then

¹⁵Some search algorithms, simulated annealing being a standard example, make use of randomness, but marshal that randomness carefully and do not merely take cases uniformly at random. We will speak loosely of a “random” search algorithm as one where cases are uniformly chosen at random.

they may consider how modifying the cases would make it harder for them to fit the theory, and focus their quest for anomalies on those areas of the search space most likely to yield cases like that. Employment of such a search algorithm will increase the number of anomalies to be expected after the examination of n cases beyond $n\varepsilon$, because ε was the probability that a random case would look like an anomaly without being one.

Real scientists also engage in such optimized searches for cases, whether in favour of or against a theory. However, here one needs to distinguish between *exploratory* and *confirmatory* (or *disconfirmatory*, as the case might be) research. In exploratory research, one is trying to refine a plausible and interesting hypothesis, or to design a future experiment, and one may do so by looking at particularly suggestive cases in light of which one then tailors the hypothesis or experiment. For instance, an ESP researcher might examine correlations between answers on an exam where ordinary cheating is impossible to form hypotheses about what sorts of conscious or unconscious telepathic communication could be going on. Moreover, this exploratory research might involve all sorts of non-random search algorithms for finding interesting cases, such as following up on anecdotally provided leads or concentrating one's search in the vicinity of cases on the borderline of being interesting, and moving it in accordance with an increased gradient of being interesting (perhaps with some random variation for good measure). But it is crucial to distinguish this initial exploratory research aimed at designing a hypothesis or a future experiment from research aimed at confirmation or disconfirmation of a hypothesis. Once it is time to confirm or disconfirm the hypothesis, one will ideally work as independently as one can of the non-random cases that inspired the hypothesis or experiment. Or, with sophisticated statistical work, one can sometimes turn a piece of exploratory research into a piece of confirmatory research by carefully analyzing the details of one's exploratory search algorithm and taking those details into account.

Keeping a clear distinction between exploratory and confirmatory research is one of the marks of good science as opposed to bad science or pseudoscience. Witztum et al. (1994) published a paper in a respected peer-reviewed statistics journal arguing that the dates of death of rabbis could be found encoded near the names of the rabbis in the Book of Genesis to a degree that significantly exceeds what we would expect in a chance text. The most plausible criticism of WRR's work has been that of McKay et al. (1999), whose speculative but not implausible naturalistic explanation of WRR's results can be put in terms of the distinction between exploratory and confirmatory research: the statistical results may be due to undisclosed exploratory research that allowed WRR to refine the hypothesis they were testing (e.g., by tweaking the form of names used for the rabbis), without the further tests being made appropriately independent of the results of the exploratory research. For another example, Wagenmakers et al. (2011) accuse D. J. Bem (2011), a respected academic psychologist, of presenting ESP research that is in fact exploratory but treating it as if it were confirmatory of the hypothesis that ESP is real (in fact, it is the work of Wagenmakers et al. that has

alerted us to the distinction between confirmatory and exploratory research). Of course, it is clear that Bem intends his work to be confirmatory—he intends to ensure that he isn’t simply on a fishing expedition for ESP evidence (see also Bem et al. (2011)). The point here isn’t whether WRR and Bem were in fact confusing exploratory with confirmatory research, but rather that, as both sides in each debate will admit, the question of keeping the two kinds of research carefully delineated is crucial. And while both of these cases involve confirmation, similar points come up for disconfirmation.

In the absence of serious statistical work of confirmatory or disconfirmatory research, the scientist can dismiss the pseudoscientific critic’s objections as mere anomaly mongering. But a bald “This is just an anomaly” response, even when it is epistemically justified, is intellectually unsatisfying and rhetorically unhelpful. Moreover, even individually statistically insignificant anomalies can pile up, and so such a bare response leaves the theorist open to further attack along similar lines by an opponent who returns with improved statistical methodology.

2.3.2. The Best Response Section 2.3.1 covered the weakest (yet in a way sufficient) response to anomalies: assert that they are “mere” anomalies, i.e. that they are just the result of anomaly mongering. By contrast, the best kind of response to an anomaly would be to show, independently of the evidence for the theory, that the apparent counterinstance of the theory is not in fact a counterinstance. Technically, this would make the anomaly cease to be an anomaly—given the new data, the case would no longer appear to be a counterinstance for the theory.¹⁶ Maybe Felix’s identity as a canid could be shown by blood tests, and that would be the end of the anomaly.

But often this kind of response is not available. For instance, the apparent counterinstance may be unavailable for detailed examination. Perhaps Felix’s owner refuses to let us take blood samples (if Felix’s owner is the one who is pressing the anomaly, then we should be *very* suspicious at this point). Maybe we have gathered all the data about a distant galaxy that our research budget allows us to gather. Or possibly the case is historical and we have already mined the documentary and archaeological data to our utmost.

2.3.3. Intermediate Responses In between the extremes of the bare statement that the case at hand is a mere anomaly that had inadequate statistical methodology and a full independent defense of the theory’s fit with the case lies a practice of offering a variety of stories showing how the case might fit with the theory. Each of the stories has the property that if it is true, then the case fits with the theory. At the same time, the stories go beyond the mere assertion of

¹⁶This is an epistemic sense of “appears.” Sometimes one still experiences psychological doubt after one has concluded that there are insufficient rational grounds for them. Most anyone who has walked in a cemetery at night has experienced this. Alternatively, the situation could be taken to be analogous to the way a pencil in the water still appears to be bent, even though one has dismissed the evidential value of this appearance. On this model, the anomaly would remain an anomaly, but it would be drained of any epistemic significance.

experimental or perceptual error. That is, they do not simply suggest that the appearance of exception is due to an error in measurement (broadly construed) or a misapprehension of the case—they either do not talk about errors of measurement or misapprehensions at all, or else they offer a putative explanation of the *specific* error or misapprehension. The more probable (either independently of the theory or conditionally on the theory) and compelling the story, the more intellectually and rhetorically effective the response is. We will call these responses “conciliatory stories”.

The least powerful conciliatory story would be a logically possible story that shows how what we observe in the anomalous case is *logically* compatible with the theory. We can call such a story a “weak defense”. Weak defenses are rarely given in serious scientific work, because when we are dealing with empirical data they are too easy to give. We could always suppose logically possible elves who meddled with our experiments, say.

Slightly better than logical possibility is epistemic possibility. There are two ways to think about epistemic possibility. On the first, something is epistemically possible when it has a non-zero epistemic probability. Logical possibilities can have zero mathematical probability and perhaps zero epistemic possibility. But knowing that the story has non-zero possibility is consistent with near statistical certainty that it is false. A stronger form of epistemic possibility consists in consistency with all that we know to be true. But this is also quite weak: there are often too many options consistent with all we know.

A better kind of conciliatory story than the previous two will be not only logically possible and logically compatible with the theory, and will not only be true “for all we know”, but will *fit* with what we know about the world, whether or not we have any reason to believe the story. Evolutionary theorists when faced with a case where it is difficult come up with an evolutionary pedigree for some biological feature sometimes resort to these kinds of conciliatory stories. There may be little independent evidence for a given story, and it may not be especially probable conditionally on general evolutionary theory. However, neither will it be surprising given evolutionary theory. It’s just the *sort* of thing one would expect, and there’s no evidence against it.

Critics derisively call stories like these “just-so stories”. But there is nothing epistemically wrong with just-so stories. Just-so stories fit in a not implausible way with what we know about the world and show how a theory can apply to an apparently anomalous case. Granted, insofar as these are mere just-so stories, we have no reason to believe that they are *true*, and so it would be mistaken to think that a just-so story gives us any significant positive evidence for the theory. Just-so stories are primarily defensive moves by the theoretician. But the availability of a just-so story weakens the appearance that there is a counter-instance. Furthermore, the availability of just-so stories can lend coherence-based positive epistemic status to a theory over a rival that lacks even such stories.

A bare statement that we are dealing with a mere anomaly *can* be a rationally sufficient response. Such a statement when combined with a just-so story,

however, can both give us more understanding of the anomalous case and be rhetorically more powerful.

One step above a mere just-so story is what one might call a “maybe-so story”. A maybe-so story S not only is such that if true, it reconciles the anomaly with the theory, and not only does it fit with our knowledge of the world, but it has the additional property that it is not very unlikely to be true given the theory: $P(S|T\&K)$ is not very low, where K is our background and S is the story.

A collection of several maybe-so stories may well undercut any evidentiary force against the theory that remains in a collection of anomalies after one has taken into account the fact that we expect some¹⁷ anomalies if the theory is true. Suppose my maybe-so stories are S_1, \dots, S_m . Let S^* be the disjunction of all the stories. Then even if the conditional probabilities of the individual maybe-so stories is fairly low on $T\&K$, the probability of the disjunction S^* on $T\&K$ may be moderately high. But:

$$(7) P(E|T\&K) \geq P(E|S^*\&T\&K)P(S^*|T\&K).$$

If the particular anomalous observation E is very likely on S^* and S^* has moderately high probability on $T\&K$, then $P(E|T\&K)$ will be fairly likely. In typical cases E will not be certain on $\sim T\&K$, and often will be not be all that likely, since while it is very likely on $\sim T\&K$ that there would be a significant number of anomalies for T , it may not be all that likely that these anomalies would take the precise form of E . So $P(E|\sim T\&K)$ will not be much larger than $P(E|T\&K)$, and may even be smaller, and so in such a case E will not provide significant evidence against T , and if the maybe-so stories are collectively likely enough given T and E is likely on the disjunction of these stories, E may even provide evidence for T .

A strong form of “maybe-so” story we can call the “may-well-be story.” This kind of story is about as probable as not. Giving this kind of story a name draws attention to the potentially important distinction between a story that is, say, about 10% likely and a story that is nearly 50% likely.

Finally, there is the “probably-so story”. This story is like the maybe-so story, but $P(S|T\&K)$ is not only not very low, but is in fact greater than a half. Probably-so stories are one step below completely removing the anomaly by showing that it fits with the theory. Their individual effect is like the effect of a disjunction of maybe-so stories. Probably-so stories need not be strong enough to command assent or belief or acceptance, but strong probably-so stories will be at or near the level of believability. Plausibly, when one believes such a story, the normative force of the anomaly has been neutralized.

A useful strategy, then, will be to disjoin a number of just-so stories to form a disjunctive maybe-so story, or a number of maybe-so stories to form a disjunctive probably-so story. We do, in fact, do this in practice. We may not know why it is that, say, some galaxy has the shape it does given our present theories of galaxy

¹⁷In this way, an event can be improbable, yet unsurprising (van Inwagen invokes this distinction, most recently in his 1996, 172 n7), for, though somewhat improbable, it is a token of a type such that it is at least moderately probable that some token of that type be tokened.

formation, but we may have three or four just-so stories, whose disjunction is not unlikely given our theory.

Critics of evolutionary theory deride biologists when they respond to criticisms with just-so stories. In so doing, they neglect the fact that their criticism of evolutionary theory may amount to mere anomaly mongering, and hence may not need any rational response at all. The giving of a just-so story is rationally supererogatory when the number of anomalies fails to be of the order of magnitude of $n\epsilon$ or when the search algorithm that generated the anomalies was not a uniformly random one and complex statistical analysis was not performed, but nonetheless giving the just-so story is *helpful*, as long as one is clear that the just-so story is a defense of the theory rather than evidence for the theory, and can be particularly important for helping to make vivid a scientific response to the public. Moreover, sometimes the just-so stories rise to the level of maybe-so stories. Plus, it may well be possible to find several maybe-so stories and disjoin them, and at that point this might not only refute the criticisms of the theory, but provide additional evidence for the theory.

In fact, the ascent from just-so to maybe-so or probably-so can happen with a single story:

We think it is time to stop running from Just-So Story as an epithet and to embrace its merits: not that science ends up being a Just-So Story, but that it generally begins as one, emerging from curiosity, questioning, and uncertainty. In the best cases, it then progresses to reasoned conjecture, to asking “What if?” and “Could it be?” and then, if the imagined story seems worth pursuing and is in fact pursuable, to validation . . . and, if productive, to further refinement. (Barash and Lipton 2009)

3. The Problem of Evil

3.1 *The problem of evil as a problem of anomaly*

As common experience makes sufficiently clear, religious believers often maintain their beliefs in the face of arguments to the contrary. This is so even in the case of the most powerful argument against religious belief: the argument from evil. There can be good pragmatic justification for such stubbornness. Religious belief is motivating. If religious beliefs were abandoned at the first sight of counter-evidence, motivation for certain forms of life—forms of life the religious believer often takes to require religious belief—would often be insufficient. And it is plausible that at least some of these forms of life are valuable, at least in the case of some believers.

But it seems to be epistemically irrational to hold on to a religious belief in the face of apparent refutation, and it might appear that the existence of certain kinds of evils in the world constitutes *prima facie* refutation of some forms of religious theism, especially Judaism and Christianity. Call an instance of evil “unjustified” provided that God would be morally required to prevent it.

Theism implies that our world contains no unjustified evils.¹⁸ But we are faced with an apparent case of unjustified evil. Hence, apparently, theism is false. One might conclude from this that by holding on to theism in the face of evil theists sacrifice their epistemic good for the sake of the practical good of certain forms of life.¹⁹

Such a view would, however, lead to an unnecessarily sceptical view of religious belief. For theism is a grand story: a story of everything. Grand stories tend to have anomalies, as reflective and honest theorists will admit. This, then, is the problem of evil as we shall consider it: *Can one epistemically rationally avoid abandoning theism in light of apparently unjustified evil? If so, how?*

We can sharpen this problem and at the same time show that the problem of evil is an instance of the problem of anomaly, namely that in connection with the problem of evil, theism is an instance of a theory facing an anomaly. Compare the formalization of the problem of evil with the formalization of the problem of anomaly above.

- (8) If T is true, then $\forall x(Fx \supset Gx)$.
- (9) It seems that Fa .
- (10) It seems that $\sim Ga$.
- (11) Therefore, $\sim \forall x(Fx \supset Gx)$. (Ampliatively from (3) and (4))
- (12) Therefore, $\sim T$. (By modus tollens from (2) and (5))
- (13) If theism is true, then every evil has a justification.²⁰
- (14) It seems that E is an evil.
- (15) It seems that E lacks justification.
- (16) Therefore, E lacks justification (Ampliatively from (14) and (15))

¹⁸The term “unjustified” is thus to be distinguished from “gratuitous.” Say that an evil is “gratuitous” provided that it, or an evil of at least as great magnitude, is not necessary for a greater good or for the prevention of a greater evil. The notion of an unjustified evil differs significantly from that of a gratuitous evil, since as Peter van Inwagen (2006) has argued, God may be justified in permitting gratuitous evils. For instance, some great good may require that there be *more* than an amount s_0 of suffering. But whatever amount of suffering greater than s_0 takes place, that is gratuitous, since a lesser amount of suffering (though still greater than s_0) would have been sufficient. Such suffering is gratuitous but may be justified. Moreover, there may be evils such that the evil itself is not necessary for a greater good or for the prevention of a greater evil, but refraining from prevention of the evil is necessary for a greater good or for a worthwhile chance at a greater good. For instance, suppose theological compatibilism and Molinism are both true, and imagine that God puts Sally in a position where she freely chooses between a great good and a minor evil and is more likely to choose the good than the evil, and that in this case her freely choosing the great good would also exceed, by an order of magnitude, the value of her non-freely going for the great good. God is then justified in not preventing Sally from choosing the evil, since given theological incompatibilism and anti-Molinism, the only way he could prevent her from choosing the evil would be to take away her freedom, which would thus lose God the probable chance of her freely choosing the good. If Sally nonetheless ends up choosing the evil, the evil may be gratuitous, but it will be justified in the sense that God is justified in not preventing it.

¹⁹Some forms of fideism would reject rejection of religious belief based on evidence, others may not (Bishop 2007). We are only interested in investigating non-fideistic responses to evil.

²⁰If E is a justified evil, then a *justification* of E is the reason for E 's justification. Since there is always a *prima facie* case against permitting an evil, whenever an evil is justified, there has to be a reason for E 's justification.

(17) Therefore, theism is not true.

It should be clear from a side-by-side comparison that the two arguments are of the same logical form. The problem of evil is a problem of anomaly. High-level scientific theories are not always theories of everything (though some do dream of a “final theory” or “theory of everything”²¹), but they are still, at times—Newtonian theory, quantum theory, evolution, etc.—quite grand theories.²² On the basis of the similarities of form and scope, we assert the following two theses: *General Analogy Thesis*: Just as apparent counter-examples need not undermine the epistemic rationality of maintaining scientific theories, so apparently unjustified evils need not undermine the epistemic rationality of maintaining theistic belief. *Specific Analogy Thesis*: The several strategies for responding to anomalies to scientific theories briefly surveyed above are all promisingly open to the theist.

After some brief discussion of the argumentative form of the problem of evil displayed above, we will draw out the details of the Specific Analogy Thesis.

Note that we will not discuss the inference from (14) to *E* being in fact evil—in the kinds of cases that the argument from evil tends to be based on, this inference is uncontroversial except to error theorists about value.

3.2 Excursus on “seems to lack justification”

There are two ways to read (15) and thus the inference to (16). There is a weak reading of (15) which is negative and a strong reading of (15) which is positive: (15a) It is not the case that we know of a justification, even after thinking hard about it.

(15b) *E*, some token instance of horrendous evil, seems to lack justification in virtue of seeming to be unjustified.

The inference from (15a) to (16) is essentially the core of the original version of Rowe’s famous version of the argument from evil (Rowe 1979). It has come to be called the “noseeum” inference by critics because it seems to have the same bad-making features as an inference from “I can’t see any noseeums (very tiny biting bugs) in the tent” to “There are no noseeums in the tent.” (see McBrayer (2010) and Dougherty (2011b) for more background and overview of this problem).

One form of reasoning behind (15b) is based on what Dougherty (MS, 2011b) calls the “Common Sense Problem of Evil”, namely that some evils *E* appear to be unjustifiable, and hence unjustified. This is also present in Rowe’s many papers on evil.

The earliest and most wide-spread attempt to block these kinds of inferences focuses on a principle devised by Wykstra called CORNEA, short for Condition Of ReasoNable Epistemic Access (Wykstra 1984, 2009). Versions of the principle attempt to state conditions on when appearances count as evidence, conditions

²¹Famously, Weinberg (1993).

²²On certain not unpopular forms of reductionism, physics will count as a theory of everything. A theory of everything does not, of course, actually offer an explanation of each token event. However, they do seek to provide unified frameworks from within which every token event could be explained if we knew the relevant details surrounding the event.

which the appearance of gratuitous evil are thought not to meet. However, the principle and general strategy has been subjected to persistent objections (e.g., see Howard-Snyder (1992), the essays by Russell, Draper, and Gale in Howard-Snyder (1996); Graham and Maitzen (2007); and McBrayer (2009)). Van Inwagen has based a similar response on an admittedly skeptical basis (see van Inwagen (1995, chapter 3). For a similar but more general line of thought see also Alston (1996) who defends agnosticism about the main premise of the argument from evil), and Bergmann has based a critique of the inference from (15a) to (16) on a set of “skeptical theses” (Bergmann 2001).

Views like those of Wykstra, van Inwagen and Bergmann go by the name “sceptical theism” because they all express scepticism about our ability to grasp God’s reasons for allowing evil, even if such reasons there be. Many have pointed out that the sceptical bases of skeptical theism seem to have sceptical consequences well outside the narrow confines they intend, especially regarding moral knowledge (Almeida and Oppy (2003), Jordan (2006), Schnall (2007), Maitzen (2009) and Sehon (2010), though, in fairness, see Bergmann and Rea (2005)).

So the literature contains responses to both the strong and weak interpretations of (15). Our line of thought is consonant with sceptical theism but does not rely on premises which threaten generalized scepticism. We are saving scientific theories from too easy a refutation while accounting for how anomalies might pile up. Our account also offers a much richer and detailed analogy, for the analogy with anomalies is more regimented than those of parent and child (Wykstra (1984, 1996), but see Dougherty (2012)), of master chess player and novice (Alston 1996), or of investigating deep space (van Inwagen 1995). In what follows, we will unpack our analogy by noting points of similarity between what we have said concerning the problem of anomaly in science and plausible claims about the problem of evil.

3.3 Unpacking the analogy

3.3.1. Selection bias We had noted that one way in which Basic Anomaly Argument could go wrong was via selection bias. One knows in advance that canid features will be widely distributed across a vast spectrum and one can simply look for the most feline-looking fox to cause problems. This is classic anomaly mongering.

Similarly, one knows in advance that there will be a wide variety of kinds of instances of evil, some of which will have obvious justifications and some of which will not. This is so even within the class of evils for which one really thinks there are justifying reasons. One expects that, even among evils which have justifying reasons, one will find a broad spectrum of ease with which the reasons cannot be discerned. It should be no surprise that one was able to find a case where it was not obvious what the function of suffering was, or even where persistent scrutiny turned up no reason. If one was merely scanning for the most problematic kind of case, this is an obvious case of selection bias.

And this seems to be just what happens in the literature on the problem of

evil. Rowe (1979) takes the example of a fawn burning in a forest fire, Tooley (2008) takes the example of the Lisbon earthquake (from Voltaire, of course), and others use more horrific examples either from fiction (perhaps historical fiction or *The Brothers Karamazov*) or from the headlines. These kinds of instances are selected precisely for the fact that they stand out as the hardest kinds of cases to justify.

But if sufficiently many other instances of suffering are such that their occurring fits in squarely with theism, then the instances just alluded to are mere anomalies. We have here another instance of the failure to distinguish between exploratory and confirmatory/disconfirmatory research discussed in Section 2.3.1.

Moreover, even if most instances of evil didn't fit well with theism, instances of suffering are special cases of *happenings*. If, as the theist claims, most happenings fit together well in the theist's narrative (theistic narratives will, of course, differ by kind of theist), then the presence of this background of coherence underscores the anomalous nature of the instances of evil used in the argument from evil. The idea here is that the theist may find that most things fit well with her theism, and that instances of horrendous evil are exceptions that it is rational not to take as undercutting theism. For the opponent of theism to seize upon these cases without serious statistical analysis is anomaly mongering.

Furthermore, it bears mentioning that very, very few actual instances of evil have been investigated in any detail for whether they contributed to a greater good. Consider what that investigation would look like. It would be a kind of reverse of forensic examination: instead of looking for causes, we would look for effects and then consider whether the effects could have been as favorably produced at significantly²³ lesser cost. The investigation would require a similar degree of empirical seriousness as forensic or historical studies, but would need to combine that with serious ethical insight.

The particularity of token happenings makes it difficult to give detailed explanations of them generally. I see an apple drop from a tree as I ride my bike down the road. Does the theory of gravity explain *that* falling of *that* apple? Presumably, but I have no particular data about its mass, rate of fall, etc. Even if I got off my bike and walked across the orchard to examine this apple there is now no way to test whether its falling was in conformity with the theory of gravity. Such are the limitations of particularity, and many instances of evil bear the same problem with regard to assessing their place in a grand structure of reasons.

The point we want to underscore here is simply that whether some token event has some kind of explanation is not always straightforward, and may require extended investigation, and that few instances of suffering have been examined in the relevant kind of way in much detail. And to search for a justification is to look for an explanation of why God would be permitted to allow the evil.

²³The reason for the "significantly" qualifier is because of van Inwagen type considerations. Cf. note 19, above.

This objection alone could be devastating to the argument from evil because it undermines inferences to lack of justification from premises like (15) above. Since the data presented as counter-instances to the generalization that there are no unjustified evils were collected according to a clear selection bias, and no corrective statistical adjustment was done, those data are of little evidential value.

3.3.2. Caveat: Deceiving appearances It was noted above that it is not uncommon for theories to have consequences which are contrary to appearance, where we only believe that appearances are deceiving because we have a satisfying theory according to which they deceive. Consider the examples of the appearance of deep density in “solid” objects and the appearance of direct design in the mammalian eye. These appearances can be rightly rejected as deceiving, and not because of direct empirical observation to the contrary. For in many cases there is no such direct empirical observation. The table, for example, never loses its appearance of solidity, the pencil in the water never loses its appearance of crookedness, etc.). Likewise, it is not irrational to believe that some instances of suffering serve greater goods or are otherwise justified *only* because we have an otherwise satisfying theory according to which that is so.

Furthermore, there is intuitively an important moral difference between suffering completely unconnected to one’s own wrongdoing and suffering connected in some relevant way to one’s own wrongdoing. But there is room for our being deceived in regard to innocence. Apart from cases of those incapable of wrongdoing, such as small children, we should expect some error rate in our ability to identify whether someone’s suffering is connected in some relevant way to their own wrongdoing. The person may appear totally innocent or undeserving of the suffering in question, but it is uncontroversial that the appearance of innocence is sometimes deceiving. As disappointments when the apparently innocent scandalously turn out not to be so teach us, humans are imperfect at identifying innocence. We do not lay much emphasis on this, especially since this consideration needs to be counterbalanced at least to some degree by the fact that we can also be wrong in thinking that someone’s suffering is deserved, whether because the deed the person is deemed guilty of never occurred or because the person was for some reason inculpable or insufficiently culpable.

3.3.3. How many anomalous evils should we expect? Consider this analogy with naturalistic evolution (NE). NE entails the following universal generalization (or something relevantly similar):

(NUG) For every extant non-domesticated species on earth there is a gradual chain of random mutations and non-intentional selective pressures that produced that species from a unicellular ancestor

But as we noted, there are over a million extant non-domesticated multicellular species. We further noted that there are well-known limitations on our ability to gather data on this subject. Some limitations are external—such as the incompleteness of the fossil record or, at times, inability to probe genetic

material—and some of the limitations are internal—it is a grand theory with vast implications and our cognitive resources are finite (similar considerations apply to quantum physics and cosmology).

So it should be no surprise, then, if, even after considerable investigation, there appear to be counter-instances—whether strong or weak—to NUG.

Theism entails the following universal generalization.

(TUG) Every instance of evil is justified.

But as is obvious, instances of evil are nodes in a vast, complex network of events. Further, there are obvious limitations on our ability to gather data. Some of these limitations are external—such as the difficulty of obtaining the relevant information—and some are internal—the connections are of enormous proportions and our cognitive resources are finite. The relevance of the combination of complexity of subject matter and cognitive limitations is to greatly increase the expected number of anomalies. Alas, for both NE and theism it is impossible to generate anything like a precise figure. Nevertheless, it is plausible that, while theism in itself a simple hypothesis (Swinburne 2004, Chapter 5), working out the theistic narrative for history is of a considerably greater order of magnitude of complexity than NE. Indeed, the history of the origin of species and descent of man is only a subset of the total story of the universe and human history. So we should expect considerably many more anomalies for theism than for NE.²⁴

So, just as it can be rational for advocates of evolution who are satisfied with its narrative to continue believing the theory despite observations that do not appear to fit with that theory, or even appear not to fit with it, so theists satisfied with their narrative can be rational in continuing to hold their religious beliefs. A few caveats are in order, however.

First, the conclusion—both in the naturalistic evolution example and the theism example—is that it is *possible* to be epistemically rational in continuing believing in the face of anomalies. However, it is also possible for people to be irrational in continuing to believe a theory given a collection of anomalies. This is because, as noted, anomalies can pile up and set the ground for the rational necessity of a paradigm shift. And even apart from quantity, the quality of an apparent counter-example can be good enough to require disbelieving a theory. That is, when one's total evidence gives one an all-things-considered reason to take the anomaly to be genuine, then even an otherwise well-confirmed theory should be abandoned.

Also, we have appealed here to the notion of one's being satisfied with a theory. It is beyond the scope of this paper to rigorously define when one should be satisfied with a theory, but, as is fairly standard, we take confirmation to be a notion that combines subjective and objective elements. It is plausible that in some form, a scientific theory is a complex answer to a series of why

²⁴Interestingly, Alvin Plantinga suggested in correspondence that perhaps all living creatures constitute an anomaly to NE, for “in nearly every species there will be features or organs that seem (I think in your sense) not to be the product of a gradual chain of random mutations and non-intentional selective pressures tracing back to a unicellular ancestor”.

questions. Whether one takes a why question to be answered will depend on what theoretical virtues they prize. Reasonable individuals will weigh various virtues differently. So there will be reasonable differences in whether one takes a theory to have satisfactorily answered their questions. However, there are limits to reasonableness. Not just any weighting of theoretical virtues will be acceptable. How to sort this out fully is a difficult question that deserves much attention, but we think the rough sketch we have just made is plausible and widely-accepted.

It is beyond the scope of this paper to defend either the thesis that evolutionary theory is reasonably taken to be a satisfactory theory or that theism is. Swinburne has offered the case that theism is rationally acceptable on precisely the grounds that other large-scale scientific theories are, and we think something like what he says is correct. There may be useful externalist notions of rationality, but we are here dealing with a notion of epistemic rationality which is perspectival (Foley 1993, Kvanvig 2003). All we assume in either the case of naturalistic evolution or theism is that one can be reasonable to continue belief in the presence of anomaly so long as one is reasonably satisfied that one's view was sufficiently well confirmed independently of the anomaly.

We now go on to illustrate how the different degrees of responses to anomaly can be applied to the case of the problem of evil.

3.4 Responding to evil: Levels of conciliation

We have already covered a potentially devastating objection to the argument from evil: the argument depends on biased sampling. To search for anomalies to a grand theory and, unsurprisingly, find them is to engage in anomaly mongering. The advocate of naturalistic evolution is right to dismiss young-earth creationist anomaly mongering resulting from simply searching for hard cases. Likewise, the theist is justified in dismissing atheistic anomaly mongering resulting from a search directed at finding particularly problematic evils. Still, in neither the theistic or the evolutionary case is this the most satisfying response, justified though it is. What would be most satisfying would be a complete explanation of how the anomaly's appearance of counterinstancehood was misleading.

The least strong level of conciliation above mere rejection of anomaly mongering is the mere defense. A mere defense is a *logically* possible story that shows how the appearance of unjustified evil is logically compatible with theism. Though this gambit is appropriate for dealing with the logical problem of evil, it is, with respect to the evidential problem, scarcely better than noting that the anomalies are mere anomalies, for logical possibilities do not as such carry much weight by way of evidential force. The most one could say is that if there is a logical possibility that p , then the epistemic probability of p is greater than zero (and it needn't even be this high where mathematical probability²⁵ applies instead of epistemic probability or the two converge). Suppose I'm defending

²⁵A dart with a perfectly defined point that lands at position x in a continuous target zone had zero mathematical probability of landing at x , but it was logically possible for it land at x .

theory T from anomaly A via a mere defense, D . I note that D is logically possible and entails $T \& A$. The problem was that $P(T|A)$ was low or zero (or at least $P(T|A) \ll P(T)$). If all we can say for D is that $P(D) > 0$ then the most we can get from the fact that D entails $T \& A$ is that $P(T \& A) > 0$. But that is of practically no use in defending the epistemic rationality of believing T .

The same goes for the kind of defense which is proposed as *epistemically* possible. The weak view of epistemic possibility is evidentially on par with what we get from the strong view of logical possibility: non-zero epistemic probability. The stronger interpretation of epistemic possibility equates it with what is true “for all we know,” i.e., what is logically consistent with what we know to be the case. But this is again very weak, for most of us, too many options are logically consistent with what we know.

Maybe there is a less weak notion of “real possibility.” This is the kind of literal defense given by defense attorneys in a trial trying to establish reasonable doubt to avoid conviction. It is not clear that these alternative theories of the crime have to be more than merely true for all we know, but it seems that the typical legal defense is stronger than something which we merely can’t know not to be true, since, plausibly, reasonable doubt requires more than this. We can reasonably disbelieve what we do not know not to be the case. Van Inwagen’s (2006, xiii, 66–67, 78, 84–85, 88, 104, 113, 132, 162) language of what is true “for all we know” and “real possibilities” fits with both of the last two interpretations of epistemic possibility, if they are different, though he himself disavows belief in logical or epistemic possibilities (personal correspondence).

The next level of conciliation concerns stories which, though they are not solidly backed independently of T , nonetheless have some degree of plausibility conditionally on T . The stories fit organically with the theory being defended and not improbable on independent grounds. Many of the accounts given by evolutionary theorists in the absence of more direct evidence are like that.

For instance, the best present hypothesis about the evolution of the cecal appendix is that it provides a “safe-house” for gut bacteria and developed from a small pouch for gut flora in response to selective pressures such as those induced by diarrheal illness. Laurin, Everett and Parker [LEP] (2011) support this by citing such evidence as the recurrence and occurrence of a trait structurally like this one in various kinds of organisms, a consistent morphology that allows the trait to function as such a “safe-house”, as well as epidemiological studies that show that diarrheal illness provides selective pressure in developing countries. If we bracket the extensive evidence for evolutionary theory in general, the cited evidence primarily supports the claim that gut bacteria stay safe in the appendix and that this is beneficial to the organism. The hypothesis that the appendix *evolved as a safe-house* is then incrementally confirmed given the claims of evolutionary theory that such benefits tend to provide evolutionary explanations of heritable structure. And even so, the degree of confirmation appears not to be so high. Despite once using the phrase “proof by deduction” (p. 576), though notably in scare quotes, for one of their arguments for the claim that the appendix

in fact keeps beneficial bacteria safe, LEP are very cautious when summing up arguments, emphasizing that no direct experimental data on the function of the appendix is available (pp. 576–577), and using such phrases as “some degree of confidence”, “seems likely” and “points toward” (p. 577). And when they briefly sketch their evolutionary history they are even more cautious:

it may be *hypothesized* that a small, appendix-like pouch which functioned primarily as an immune structure maintaining the gut flora evolved prior to the evolution of a cecum that functioned effectively as a digestive structure. Although the morphology that preceded the cecal appendices in Euarchontoglires and Diprotodontia *remains unknown*, data on monotremes indicate that a narrow, appendix-like structure *may* have preceded the evolution of a larger cecum with an effective digestive structure. (LEP 2011, 577, emphases added)

Given this story, any discomfort one might feel with evolutionary theory in light of the existence of the appendix should evaporate. However, it must be emphasized that even before the development of this theory, it was rational to treat the case of the appendix as a mere anomaly, and in their brief historical survey of speculation and research on the appendix (p. 569) it does not even occur to the authors to consider the question that the appendix might be a challenge to evolutionary theory, and it is no part of their purpose to provide an apologia for evolutionary theory.²⁶

For a global example (one can also give local stories for particular evils) on the side of evil, consider the story embodied in the proposition L , that God wants us to acknowledge our cognitive limitations by weaving the purpose of the kind of suffering in question into a story too large for us to grasp. This coheres well with theism, and though we have no evidence for it independently of arguments for theism, perhaps, neither do we have any independent evidence against it. So L meets two criteria. First, $P(L|K_0)$, where K_0 is the background independent of the arguments for theism, is not too low (though less than half by far) and, second, $P(L|T\&K)$, where K is the full background, is also not too low, though higher than the preceding probability.

It is worth calling to mind—to elucidate the notion of plausibility in play here—van Inwagen’s distinction between improbability and surprisingness (van Inwagen 1996, n7), which we have suggested a model for in Section 2.3.3. The proposed story F to cover the evolutionary purpose of the appendix is somewhat improbable, in that we have no direct evidence for it. And not only that, but it is even true that $P(F|NE)$ is not high. However, given NE , F is not at all surprising. It is just the sort of thing we’d expect to be true were NE true. The same goes for L above. It is not independently probable, nor is it very probable on T , but it is just the sort of thing one would expect if T were true, or at least it is not that surprising. And of course there are many evil-contextualizing stories

²⁶The authors do mention the popularly accepted theory that the appendix has no function. However, even there they do not think about it as a challenge to evolutionary theory (presumably because evolutionary theory while giving pride of place to selective explanations can also give non-selective explanations, say based on exaptation).

S_1, \dots, S_n such that each of them satisfies the two probabilistic conditions mentioned above. It could be that God has no direct intent to hide his ways from us, but rather this is just a consequence of the great good of the overall pattern of the universe combined with the good of there existing beings of approximately our capabilities that we cannot grasp the purpose of natural evil. Or it could be that in fact the purpose was made known to humanity, but many of us have ignored the revelation. All of these are realistic possibilities in the sense that we have no direct evidence against them, and though perhaps we also have no direct evidence for them, they are unsurprising given theism. When there are several such stories this is reason to believe that there is some natural (or quasi-natural) kind which is sufficiently highly probable on the theory in question and which generalizes the stories. And of course the disjunction of S_1, \dots, S_n could be quite probable even though no particular disjunct is. This will leverage this second level of conciliation to a strength equivalent to the fourth level. But first let us turn to the third level.

This third level of conciliation concerns stories such that though they are not more probable than not, they have significant independent probability (say between roughly 20-30% just as a representative range). These are stories of which we sometimes say things like “I’m tempted to believe this” or “I’m not sure I believe this, but I find it really compelling” or similar phrases which situate our credence in the story between the merely non-trivial and the believed. Here is a story which has that status for one of the authors.

Finite creatures, unlike God, cannot appreciate goodness for goodness’ sake. Rather, they must not only be drawn toward the good, but pushed toward it “from behind” out of horror of the evil. And it really must be horror to keep us from continually turning away from the path to beatitude. Furthermore, unless a finite person both empathized with the pain of each sufferer and saw herself as a potential perpetrator of the great suffering of others, she would think that somehow she was different and could self-establish a utopia without messing it up like everyone else before her. One of us finds this plausible but doesn’t quite believe that it is the case, or at least wavers between belief and suspension of judgment. This status is much more than what has traditionally been called a “defense” yet falls short of what is often required to be called a “theodicy,” for it is advanced not as an account of what is taken to be the case, but what is taken to likely be the case. The confidence might cross the 50% probability barrier without crossing over into belief, since probabilities slightly over 50% are insufficient for belief—for instance, if you see a completely indistinct human figure in the distance, you would not believe that the figure is female simply on the grounds that 51% of the people in the world are female, but only that it is more likely than not. Naturally, a disjunction of such likely stories will have considerable force. But even one such story can have significance.

Finally, we move from just-so stories and likely stories to probably-so stories: stories which are not just “likely stories” in the sense just treated, i.e. ones which have considerable plausibility and doxastic pull, but ones in which our

credence is sufficient for belief. A likely-so story removes, in a way, the appearance of counterinstancehood, and so removes the anomaly. Why only “in a way”? For it explains why the event looked like a counterinstance in a manner that allows us to understand why that appearance is misleading, but that doesn’t always mean that the appearance itself is removed. We could address this issue by distinguishing between usages of “appears,” as is sometimes done (Wykstra 1984 cites Chisholm and others in his footnote 19). For example, the pencil half-submerged in the glass of water still appears to be bent (“visually”, we might say), even when we know enough physics to be quite sure (“rationally”) it is straight. Similarly, understanding why suffering must occur (in general) cannot be expected to make each instance of suffering appear justified, even while we keep in mind that the appearance is misleading (especially since we rarely know whether there is some particular good brought about by this suffering and if so what it is).

One story which we both believe is that something relevantly like the “soul-making” theodicy which goes back to St. Irenaeus and was defended in modern times by John Hick (2007) and Richard Swinburne (2004, Chapter 11) Swinburne (1998)) is true. We are not saying we endorse all that Hick and Swinburne say about this matter. Rather, we are saying that we believe that something like it is true. The core idea is that certain great goods of virtuous activity are such that the proposition that they obtain logically entails that evil occurs and that a world with exercises of virtue that depend on evils and the observed evils, and where the virtues develop in part by the effort of their possessors, is better than a world where there are no evils but neither is there exercise or development of virtue. Some of the key virtues we have in mind are: empathy, compassion, mercy, generosity, forgivingness, and humility (being humble, that is; this is distinct from shame, though shame has its place). The picture is not that God causes or plans directly that particular evils occur so that these virtues might be exercised. Rather, it is that God creates a world where “the sun shines and the rain falls on the just and unjust alike,” a world where we are to a large extent on our own, but one with a nature we must learn to respect and which we can never fully control. In such a world we learn what damage we can do, how and why we must depend on God, how valuable a human being is, and we can exercise the above key virtues.

In a way it is like letting one’s kids participate in a mountain bike club or a chess club, despite the risks, because of the virtues which might be developed. It seems odd to describe the motivation in these terms—“I accept an increase in my child’s risk of harm in order for her to gain virtue”—yet that is not far off. This is a normal and virtuous part of parenting despite the knowledge that the child might break a bone biking or experience deep disappointment and self-doubt at losing a key game. These are not trivial matters to a child, even though we know children often exaggerate their suffering in a way (though in a way not, for they truly can feel crushed at a loss in a race or match). Yet we are confident such risks can be for the best. We believe that something like this explains why

there is a profusion of natural evil in the world. And so even though some events might remain mysterious in the particulars (as most particular events do) and seem not to fit it, we also have a story—a story we believe—which explains how that appearance is misleading.

This is not the place to fully defend the Irenaean theodicy. Rather, our purpose has been to illustrate how each of the levels of conciliation can function. And, of course, they can all function together for a combined effect. We believe the theist has a satisfying explanation of a wide variety of phenomena, and therefore that facts about evil are anomalous with respect to this theory. We believe that the theist has a number of just-so stories, a number of likely stories, and even promising, believable theodicy, and therefore that the theist is reasonable to continue to believe in the face of evil. Treating the problem of evil as a problem of anomaly reveals that this is the case, and that it might even have been the case without any of the stories.

Notice that a story at the fourth level need not be very precise. As we put it, we believe that something relevantly like the Irenaean story is correct. But there need not be a particular precise form of the story that we believe. This is very much what scientists do. While Laurin et al. (2011) do not express much confidence about their particular evolutionary scenario for the appendix and it probably does not rise to the level of belief for them, they might well believe that something like it is true. There are thus two ways of leveraging conciliatory stories from lower levels to higher ones. One way is by disjoining particular stories, and the other way is by embedding the stories within the scope of a “something relevantly like the following is true” operator.

3.5 Objection: The practice of philosophy

“But what you call ‘anomaly mongering’ is what we philosophers do! We hear a generalization, and then we go looking for a counterexample, without the slightest concern for anything statistical.”

However, one difference between typical philosophers’ counterexamples and the cases of evils and scientific anomalies is that a lot of philosophers’ examples live in possible worlds where one can stipulate the exact empirical conditions, and so some of the cognitive limitation worries do not apply.

Moreover, the best kinds of counterexamples to philosophical theories tend not to be mere isolated cases where the theorist can simply bite the bullet and say that they are merely apparently counterexamples to the theory. They may hint at the direction in which the theory should be emended, highlight a crucial feature of a situation that previously was not seen or force the theorist to choose between giving up the theory under discussion and giving up some other theory to which she is attached. Such counterexamples are not mere anomalies, and their rational effect goes beyond the mere fact that they appear to be counterinstances. For instance, Gettier cases (Gettier 1963) show us that justification and truth should not be merely coincidentally connected in cases of knowledge. Frankfurt cases (Frankfurt 1969), on the other hand, show that characterizing freedom

modally is unsatisfactory because modal facts can be externally changed without affecting what is intrinsic to a situation (in this respect, they are like C. B. Martin (1994) objections to counterfactual analyses), and hence the theorist should focus on what is intrinsic to the agent. Cases where utilitarianism requires one, say, to execute those known to be innocent show an apparent tension between the utilitarian thesis and the plausible principle that one must never act unjustly.

3.6 Objection: Too many anomalies?

Another objection has already been anticipated. We have acknowledged the obvious: anomalies can pile up and make a theory untenable despite all we've said. Some comments by Rowe (1998, p. 4), Rowe (1998, p. 533) indicate that he takes it that this is the case with respect to evil and theism. Tooley says that letting n be “the number of states of affairs each of which, judged by known rightmaking and wrongmaking properties, it would be morally wrong to allow” then “Taking n to be in excess of a billion would seem, therefore, to be a very conservative estimate” (Plantinga and Tooley 2008, p. 141, p. 142). This is a disputed first-order issue and so orthogonal to our point here, but it is worth framing a brief reply, lest the main thesis advanced in this paper, though compatible with Rowe's and Tooley's claims, be undermined in importance.

If the claim is that there are so many anomalies, i.e., apparently unjustified evils, then we need serious statistical estimates as to how many anomalies we would expect if theism were true and serious statistical estimates of how many anomalies there are. The average person probably suffers seriously many times a year—many people seriously suffer daily, though some affluent Westerners escape serious overt suffering for years at a time. Over a period of fifty years, with a population at its current level, we can expect of the order of a trillion instances of serious suffering. It seems quite plausible that given theism the expected error rate in evaluations whether an evil is justified or not will be no less than one percent, and perhaps quite a bit higher if we find plausible some of the stories from Section 3.3. With a trillion cases of serious suffering, that would give us ten billion cases of apparently unjustified serious suffering over fifty years, and Tooley's “billions” of cases would not be surprising. Without serious statistics—and this back-of-the-envelope sketch is no substitute for them—it is difficult to argue that we are seeing more cases than we would expect.

Of course, it might be thought that the above enormous estimate of the number of cases of serious suffering is itself an anomaly of a different sort—perhaps we would not expect trillions of cases of serious evils if theism were true. But notice that once we are no longer specifically dealing with apparently unjustified evils, this is only one anomaly—the anomaly of there being very many evils. Plus, while talk of “billions” and “trillions” is rhetorically effective, when asking whether there are quantitatively more serious evils than we would expect given theism, we really should talk about evils on a per capita basis, and maybe a per capita per annum basis.

It is not our point here to give anything like a full response to arguments

from the numbers of cases of serious suffering or of apparently unjustified evil. Such atheological arguments would need to be more carefully developed before fuller responses could be given.

3.7. *Objection: Scientific theories are better confirmed*

“But mature scientific theories are much better confirmed than theism, and hence while anomalies are not enough to destroy the credibility of a scientific theories, evils are sufficient to destroy the credibility of theism.”

Now, first of all, the reason for thinking that scientific theories are much better confirmed than theism had better not rest on the evidential argument from evil. For we are precisely evaluating the impact of the evidential argument from evil on the credibility of theism, and hence the relevant probability is that which theism has prior to the evidential argument from evil. There is, however, a strong cumulative case for theism based on cosmological, design, religious experience and ontological arguments (e.g., Swinburne (2004)). The main difficulty facing theism is precisely the problem of evil. If that is bracketed, the case for theism may look very strong. Or at least a theist would be reasonable in thinking so.

But in any case our argument does not rest on the strength of theism to resist anomaly. Rather, it rests on an analysis of the weakness of those arguments that are based on the mere existence of anomalies without serious work on whether the numbers of anomalies are significantly greater than we would expect on the theory. This analysis works whether the theory to which anomalies are found has the confirmatory status of mature science, or is a theory that has yet to reach that status, as was the case for evolutionary theory in 1859 after Darwin published his *On the Origin of Species*. And it is a merit of this account of anomaly that it does not require the theories that are defended from anomaly to be particularly strong. Just as we should not say that mature scientific theories stand refuted by anomalies, so too we should not say that major scientific theories almost always stand refuted in their infancy just because anomalies are almost always easily found in the infancy of a scientific theory.

4. Conclusion

It is quite normal to have an instance a of F such that our only reason for believing that a is G is that we accept a theory T that entails that every instance of F is an instance of G , and it can even be the case that had we not accepted such a theory, we would have reasonably thought that a is not G . We can expect scientific theories that entail non-trivial universal generalizations about large numbers of instances to have strong anomalies. If it is rational to believe such a theory absent the anomalies, it can still be rational to believe the theory given the anomalies, unless serious statistical work shows otherwise. The anomalies present research problems for scientist who can then seek to find conciliations at multiple levels.

Likewise, if it is rational to accept theism when one brackets Rowe-type arguments from evil, it can still be rational even given those arguments, since

apparently unjustified evils can be reasonably seen as mere anomalies, unless serious statistical work shows otherwise. But, again, the anomalies—or at least types of anomalies—provide research problems for the theistic philosopher who can offer conciliations at a variety of levels.

REFERENCES

- Abdi, Hervé. 2007. “Bonferroni and Šidák corrections for multiple comparisons.” In *Encyclopedia of Measurement and Statistics*, edited by N. J. Salkind. Thousand Oaks, CA: Sage.
- Almeida, Michael and Graham Oppy. 2003. “Skeptical Theism and Evidential Arguments from Evil.” *Australasian Journal of Philosophy* 81: 496–516.
- Alston, William. 1996. “The Inductive Argument from Evil and the Human Cognitive Condition.” In *The Evidential Argument from Evil*, edited by Daniel Howard-Synder. Bloomington, IN: Indiana University Press.
- Barash, David P. and Judith Eve Lipton. 2009. *How Women Got Their Curves and Other Just-So Stories*. New York: Columbia University Press.
- Bem, Daryl J. 2011. “Feeling the Future: Experimental Evidence for Anomalous Retroactive Influences on Cognition and Affect.” *Journal of Personality and Social Psychology* 100: 407–425.
- Bem, Daryl J., Jessica Utts, and Wesley O. Johnson. 2011. “Must Psychologists Change the Way They Analyze Their Data?” *Journal of Personality and Social Psychology* 101 (4): 716–719.
- Bergmann, Michael. 2001. “Skeptical Theism and Rowe’s New Evidential Argument from Evil.” *Noûs* 35: 278–296.
- Bergmann, Michael and Michael Rea. 2005. “In Defense of Skeptical Theism.” *Australasian Journal of Philosophy* 83: 241–251.
- Bisby, F.A., Y.R. Roskov, T.M. Orrell, D. Nicolson, L.E. Paglinawan, N. Bailly, P.M. Kirk, T. Bourgoin, G. Baillargeon, and D. Ouvrard. 2011. “Catalogue of Life.” www.catalogueoflife.org/col, queried August 30, 2011.
- Bishop, John. 2007. *Believing by Faith: An Essay in the Epistemology and Ethics of Religious Belief*. Oxford: Oxford University Press.
- Darwin, Charles. 1859. *On the Origin of Species by Means of Natural Selection, or the Preservation of the Favoured Races in the Struggle for Life*. London: John Murray.
- Dawkins, Richard. 1996. *The Blind Watchmaker*. New York: W.W. Norton.
- Dougherty, Trent. 2011a. “The Common Sense Problem of Evil.” Keynote Address at Annual Meeting of the Australasian Philosophy of Religion Association.
- Dougherty, Trent. 2011b. “Recent Work in the Problem of Evil.” *Analysis Reviews* 71: 560–573.
- Dougherty, Trent. 2012. “Reconsidering the Parent Analogy: Unfinished Business for Skeptical Theists.” *International Journal for Philosophy of Religion* 72 (1): 17–25.

- Foley, Richard. 1993. *Working Without a Net: Essays in Egocentric Epistemology*. New York: Oxford University Press.
- Frankfurt, Harry G. 1969. "Alternate Possibilities and Moral Responsibility." *The Journal of Philosophy* 66: 829–839.
- Geivett, R Douglas. 1995. *Evil and the Evidence For God: The Challenge of John Hick's Theodicy*. Philadelphia: Temple University Press.
- Gettier, Edmund. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23: 121–123.
- Graham, Andrew and Stephen Maitzen. 2007. "Cornea and Closure." *Faith and Philosophy* 24: 83–86.
- Hick, John. 2007. *Evil and the God of Love*. London: Palgrave MacMillan.
- Howard-Snyder, Daniel. (ed). 1996. *The Evidential Argument from Evil*. Bloomington, IN: Indiana University Press.
- Howard-Snyder, David. 1992. "Seeing Through CORNEA." *International Journal for Philosophy of Religion* 32: 25–49.
- Jeffreys, Harold. 1961. *The Theory of Probability*. Oxford: Oxford University Press.
- Jordan, Jeff. 2006. "Does Skeptical Theism Imply Moral Skepticism?" *Philosophy and Phenomenological Research* 72: 403–417.
- Kuhn, Thomas. 1970. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press. 2nd edn.
- Kuhn, Thomas. 1996. *The Structure of Scientific Revolutions*. Chicago: University of Chicago Press. 3rd edn.
- Kvanvig, Jonathan L. 2003. "Propositionalism and the Perspectival Character of Justification." *American Philosophical Quarterly* 40.1: 3–18.
- Laurin, Michael, Mary Lou Everett, and William Parker. 2011. "The Cecal Appendix: One More Immune Component With a Function Disturbed by Post-Industrial Culture." *Anatomical Record* 294: 567–579.
- Maitzen, Stephen. 2009. "Sceptical Theism and Moral Obligation." *International Journal for Philosophy of Religion* 65: 93–103.
- Martin, Charles Burton. 1994. "Dispositions and Conditionals." *The Philosophical Quarterly* 44: 1–8.
- McBrayer, Justin. 2009. "CORNEA and Inductive Evidence." *Faith and Philosophy* 26.1: 77–86.
- McBrayer, Justin. 2010. "Skeptical Theism." *Philosophy Compass* 5 (7): 611–623.
- Mckay, Brendan, Dror Bar-Natan, Maya Bar-Hillel, and Gil Kalai. 1999. "Solving the Bible Code Puzzle." *Statistical Science* 14: 150–173.
- Plantinga, Alvin and Micahel Tooley. 2008. *Knowledge of God*. Oxford: Blackwell Publishers.
- Popper, Karl. 1959. *The Logic of Scientific Discovery*. London: Hutchinson.

- Rowe, William. 1979. "The Problem of Evil and Some Varieties of Atheism." *American Philosophical Quarterly* 16: 335–341.
- Rowe, William. 1998. "Atheism." In *Routledge Encyclopedia of Philosophy*. pp. , 533–534.
- Schnall, Ira M. 2007. "Sceptical Theism and Moral Scepticism." *Religious Studies* 43: 49–69.
- Sehon, Scott. 2010. "The Problem of Evil: Sceptical Theism Leads to Moral Paralysis." *International Journal for Philosophy of Religion* 67: 67–80.
- Simmons, Joseph P., Leif D. Nelson, and Uri Simonsohn. 2011. "False-positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant." *Psychological Science* 22: 1359–1366.
- Swinburne, Richard. 1971. "The Paradoxes of Confirmation — A Survey." *American Philosophical Quarterly* 8: 318–330.
- Swinburne, Richard. 1998. *Providence and the Problem of Evil*. Oxford: Oxford University Press.
- Swinburne, Richard. 2004. *The Existence of God*. New York: Oxford University Press. 2nd edn.
- van Inwagen, Peter. 1995. *God, Knowledge, and Mystery*. Ithaca: Cornell University Press.
- van Inwagen, Peter. 1996. "The Problem of Evil, the Problem of Air, and the Problem of Silence." In *God, Knowledge, and Mystery: Essays in Philosophical Theology*, edited by Daniel Howard-Snyder. Ithaca: Cornell University Press.
- van Inwagen, Peter. 2006. *The Problem of Evil*. Oxford: Clarendon Press.
- Wagenmakers, Eric-Jan, Ruud Wetzels, Denny Borsboom, and Han L. J. van der Maas. 2011. "Why Psychologists Must Change the Way They Analyze Their Data: The Case of Psi." *Journal of Personality and Social Psychology* 100: 426–432.
- Weinberg, Steven. 1993. *Dreams of a Final Theory: The Search for the Fundamental Laws of Nature*. London: Hutchinson Radius.
- Witztum, Doron, Eliyahu Rips, and Yoav Rosenberg. 1994. "Equidistant Letter Sequences in the Book of Genesis." *Statistical Science* 9: 429–438.
- Wykstra, Stephen. 1984. "The Humean Obstacle to Evidential Arguments from Suffering: On Avoiding the Evils of Appearance." *International Journal for Philosophy of Religion* 16: 73–93.
- Wykstra, Stephen. 1996. "Rowe's Noseeum Arguments from Evil." pp. . 126–151. Bloomington, IN: Indiana University Press.
- Wykstra, Stephen. 2009. "CORNEA (1984-2009): In Memoriam?" Annual Philosophy of Religion Conference, MS.

ANSELM ON FREEDOM AND GRACE

James A. Gibson

Anselm of Canterbury devoted much attention to creaturely freedom and its relation to God's grace.¹ In answering the question, "why is there evil if God made creation as good?" Anselm developed a sophisticated incompatibilist account of freedom. According to incompatibilist accounts, an agent is free with respect to a choice only if causes outside of the agent do not determine that she will make that choice. By viewing rational creatures, human and angels, as free agents, Anselm recognized that such agents could not be guaranteed to always choose rightly. As a result, he argued that the good of freedom of choice accounts for how evil could come into the world without making God blameworthy.² But he worried that his account of freedom conflicted with what the Bible says about God's grace in the restoration of fallen human creatures: "the Bible speaks at times as if that grace alone seems to avail for salvation and free choice not at all, but at other times as though our salvation entirely depends on free choice."³ In addressing the question of how God's grace harmonizes with human freedom, Anselm argued that both human freedom and the grace of God have an essential role to play in the change from an agent being unjust to being just.

The central question of this paper concerns whether Anselm provided compatible answers to both questions: (1) why is there evil if God made creation as good; and (2) how, if at all, does God's grace harmonize with human freedom? I will argue that Anselm's answer to the first question makes it difficult to see

Acknowledgments: For valuable discussions and substantive feedback, I am grateful to Joseph Keim Campbell, Vern Glaser, Felipe Leon, Nicholas Owen, Katherin Rogers, Josef Simpson, Kevin Timpe, Stan Tyvoll, and especially Joseph Dowd and Bonnie Kent.

¹On notation in this paper: *DV* = *On Truth (De Veritate)*; *DL* = *On the Freedom of Choice (De Libertate Arbitrii)*; *DCD* = *On the Fall of the Devil (De Casu Diaboli)*; *DC* = *De Concordia*; *CDH* = *Cur Deus Homo*. All translations from *DV*, *DL*, and *DCD* are taken from Anselm (2002). All other translations are from Anselm (1998). All citations of Anselm's texts will begin with the work, followed by the chapter number. If the work contains separate books with subchapters, the book number will precede the chapter separated by a period.

²Two clarifications. First, I will not address whether the value of freedom outweighs the disvalue of evils that result from the use of freedom, since to my knowledge Anselm does not register this as a problem. Second, I am not concerned with natural evils (e.g., earthquakes that cause the loss of life) because Anselm regards the first instance of evil to be the fall of the devil. Since natural evils may be explained, however implausibly, by demonic activity, my focus will be on the adequacy of explaining the possibility of moral evils.

³DC 3.3.

how he can answer the second question plausibly. In particular, I will argue that his incompatibilist account of freedom gives rise to the problem of harmonizing human freedom and God's grace, and that there is no satisfactory way for him to resolve the problem of harmonization unless he is willing to make concessions to an unfavorable theological position known as "Pelagianism."

What follows in §1 is a discussion of Anselm's account of creaturely freedom insofar as it relates to the question of why there is evil.⁴ I will first characterize Anselm's account of freedom and then show how it explains the possibility of evil entering the world. Building on the results of the first section, §2 develops the problem of harmonization. After characterizing how exactly his account of freedom appears incompatible with God's grace, I will examine the attempts by Anselm and contemporary philosophers to harmonize incompatibilist freedom with God's grace. I will argue that Anselm has no satisfactory way to resolve the problem of harmonization consistent with his theological tradition. In §3, I will show that the problem of harmonization dovetails a scholastic debate over whether God could have made a creaturely agent with the ability to sin in the first instant of its existence. Sections 2 and 3 together provide two arguments for the claim that given Anselm's theological context, he must either reject incompatibilism or accept Pelagianism in some respect. Since many contemporary philosophers of religion fall broadly within Anselm's theological tradition and explain the presence of evil through a form of freedom akin to Anselm's, these results will be relevant for the contemporary scene.

1. ON THE POSSIBILITY OF EVIL

1.1. Preliminaries on Freedom of Choice

Anselm reports that the most commonly accepted definition of 'freedom of choice' during his time is "the ability to sin or not to sin."⁵ He rejects this definition at least for the reason that neither God nor a subset of angels can sin, but it would be impious to deny freedom to them. Anselm requires that the definition capture what is essential to all situations in which agents are said to be free. To this end, he defines 'freedom of choice' as "the power to preserve rectitude of will for the sake of rectitude itself."⁶ This section, §1.1, clarifies the meaning of each of the parts of this definition, so that in §1.2 we can see how Anselm's appeal to freedom explains the presence of evil.

The power (*potestas*) to preserve rectitude of will, Anselm tells us, is always present in human nature.⁷ He explains *potestas* by an analogy and by distinguishing three senses of 'will'.⁸ Beginning with the latter, one sense of 'will'

⁴For more comprehensive treatments on Anselm's account of freedom, see Hopkins (1972), Kane (1981), Williams and Visser (2001), Davies and Leftow (2004) and Rogers (2008).

⁵*DL* 1.

⁶*DL* 3.

⁷*DL* 4.

⁸The analogy appears in *DL* 3–4 and the discussion of the different senses of 'will' appears in *DC* 3.11.

refers to the “tool of the will’s action, another [sense] as the affectivity of the tool, and yet another as the using of the tool.” These correspond, respectively, to the faculty that wills, the dispositions that incline the will-instrument (faculty) by which one chooses, and the volition or choosing on the part of the agent. Concerning the faculty, Anselm provides the following analogy. One retains the power to perceive a mountain even if no mountain is within range of sight, for if a mountain were within range then one would see it. If one could not see it due to a dearth of light, one would only lack the opportunity to see it. The power of sight is not destroyed by the lack of the object or the appropriate medium for seeing. So, to complete the analogy, the power to preserve rectitude of will remains part of the agent even if one is not able to use one’s power. The power to will in particular ways is present not just in virtue of having a faculty; it is also necessary to have dispositions that incline the will to make particular choices (volitions).⁹ Without these dispositions, the will cannot move itself. These dispositions will be examined in §1.2.

The reason Anselm speaks about the power of preserving (*servandi*; also translated as “keeping”) rectitude of will is that one must have rectitude of will in order to effectively will rectitude (*rectitudo*) for its own sake.¹⁰ When one lacks rectitude of will, rectitude of will cannot be preserved. But when this condition is preserved for its own sake, Anselm calls this ‘justice’.¹¹ Accordingly, an agent is just when she preserves rectitude of will for its own sake. Being just, then, requires having rectitude of will. So if an agent *has* rectitude of will, which is being preserved for its own sake (that is, if an agent has justice), one should expect Anselm to say that an agent is just. This is precisely what he says: I have said that justice is in every case uprightness (*rectitudinem*) of will maintained for its own sake. Whence it follows that everyone who has this uprightness has justice and is just, since everyone who has justice is just. (*DC* 3.4)

not-having-justice is equivalent to being unjust, and both are blameworthy. . . (*DCD* 16)

More succinctly: having justice entails being just and *vice versa*. Thus, a person can preserve rectitude of will only to the extent that the person is just. If one does not preserve rectitude of will for its own sake, even though one has the faculty to do so, one fails to have rectitude of will. Consequently, one will not effectively will in a way that is just and therefore fails to be just. Having clarified the sense of ‘power to preserve’, it is now appropriate to examine what Anselm means by ‘rectitude of will’.

‘Rectitude’, which Anselm takes to be convertible with ‘truth’ and ‘justice’, has a general and a restricted sense.¹² In the general sense, rectitude is what

⁹*DCD* 12 makes the case that the dispositions are necessary for willing.

¹⁰*DC* 3.12: “no one wills uprightness without possessing uprightness, and no one can will uprightness except by uprightness.”

¹¹*DV* 12: “justice is rectitude of will preserved for its own sake.”

¹²In taking Anselm to distinguish between two senses of ‘rectitude’, I follow Sadler (2008, p. 94). The admission of different senses of justice comes in *DV* 12 when Anselm says that we

ought to be or what is right for something to be. In this sense, one can speak about how things are supposed to be even if agents are not the objects of reference, e.g., fire is supposed to be hot. The restricted sense of ‘rectitude’ is the one pertinent to Anselm’s definition. This sort of rectitude involves a *moral* sense of rightness, where the one who has rectitude (i.e., the just) is praiseworthy and the one without rectitude where there should be rectitude (i.e., the unjust) is condemnable.¹³ When framed within an explicitly theological context, rectitude is “an uprightness which is present in people only when they, for their part, will what God wants them to will.”¹⁴ In this sense, rectitude or justice is not one virtue among other virtues, but the satisfaction of God’s will through willing what is morally required. Preserving rectitude of will, then, is willing in a way that upholds what is morally required.

When one preserves rectitude of will, the will wills “what it ought to and because it ought to.”¹⁵ This brings us to the prepositional expression, ‘for the sake of rectitude itself’. Since we are concerned with the restricted sense of rectitude of will that makes a person just and thereby praiseworthy, Anselm is clear that a person is praiseworthy only if she wills something for the *right* reason. Accordingly, only beings with a rational nature—God, angels, and humans—are capable of being just. By contrast, when a horse grazes, it does what it ought but it is “not aware of rectitude.” Although the horse has rectitude in the general sense, it cannot have rectitude in the restricted sense at least because it cannot act for the sake of rectitude itself. But even rational agents may will in a way consistent with rectitude in the general sense without having rectitude in the restricted sense. A person may will to give money to the poor, but she is not praiseworthy if the donation is willed “for the sake of an empty reputation.” So freedom of choice is concerned with the power to preserve rectitude of will on the basis of a certain kind of reason, *viz.* for the sake of rectitude itself, and willing what is fitting for this reason makes the person praiseworthy.

Summarizing, freedom of choice is the power to preserve rectitude of will for its own sake. When one has rectitude of will, one is able to preserve it for its own sake. If it is preserved for its own sake, one is just in the sense of satisfying what is morally demanded by God and is thus praiseworthy. But if one lacks rectitude of will, one cannot effectively will rectitude for its own sake. Since it is not being preserved, one does not have justice and is thereby unjust.

do not call something ‘just’ “on the basis of that *sort* of justice” (my emphasis). It does not make sense for Anselm to write about different sorts of justice if there were only one. For a contrasting view of the meaning of ‘rectitude’, see Rogers (2008, p. 63).

¹³It is possible to lack rectitude of will without being unjust when there is no obligation to be a particular way: e.g., wearing a watch on the right wrist rather than on the left. In the remainder of the paper, I restrict failing to be just and failing to have rectitude of will to situations where there is the obligation to be just in some respect.

¹⁴*DC* 1.6.

¹⁵All discussion in this paragraph is taken from *DV* 12.

1.2. How Freedom Explains the Possibility of Evil

We have been examining Anselm's definition of freedom in order to understand his answer to the first of our two opening questions: how can there be evil or injustice if God created the world as good? From the given definition, it is not clear how the appeal to freedom is illuminating. After all, Anselm denies that the ability to sin is part of the definition.

In order to see how the appeal to freedom explains the existence of injustice, we must distinguish between a *definition* of 'freedom' and an account of the conditions under which a subject is free.¹⁶ As mentioned earlier, the former specifies what is true of all free agents—God, angels, and humans—insofar as they are free. But the conditions under which God is free differ from the conditions under which creatures—angels and humans—are free. For instance, God has freedom of himself whereas creatures receive freedom as a gift from God. So an account of creaturely freedom of choice will specify what must be the case for creatures to have the power to preserve rectitude of will for its own sake. By giving an account of the conditions under which creatures are free, the relevance of freedom to explaining the presence of injustice will become apparent.

Perhaps the most perspicuous clue to how an account of creaturely freedom explains the presence of injustice is found in Anselm's claim, "All human merit, whether good or evil, come from the two dispositions termed 'wills'."¹⁷ He identifies these two wills in the following passage:

So [God] created them happy with no deprivation. For this reason his rational human creatures received all at once the *will to be happy*, happiness itself, and the *will to be just* (the uprightness which is the very state of justice) and freedom of will as well, without which they could not preserve that state. (*DC* 3.13, italics added)

Anselm distinguishes four things: freedom of will (discussed in §1.1), the two wills, and happiness itself. The two wills are dispositions (*affectuum*) that incline the will-instrument to will according to the ends of being happy and of being just. These dispositions are characterized generally enough to allow specific willings, e.g. desires for something to be the case or volitions to bring about some end, to be, respectively, species of them or manifestations of their influence.¹⁸ The dispositions make possible an agent's effectively willing some action for the reasons of justice and of happiness. For example, donating money to the poor may be the content of what is willed, and the agent may will this for the reasons

¹⁶This distinction follows Kane (1973).

¹⁷*DC* 3.12.

¹⁸There is a scholarly dispute over whether a "Kantian" or "Frankfurtian" interpretation of the two dispositions best explains their role within Anselm's moral psychology. On the Kantian interpretation, the dispositions are first-order desires aimed at alternative objects, i.e. at happiness and justice, but happiness and justice need not always be aimed at exclusively of each other. On the Frankfurtian interpretation, the disposition for justice is a second-order desire and the disposition for happiness is a first-order desire both capable of being aimed at a single object. As far as I can tell, nothing I say in this paper commits me to either interpretation nor affects my arguments in following sections. For more on this dispute, see Tyvoll (2006), Rogers (2008, pp. 66–67), and Williams (2009).

that caring for the poor is the right thing to do and it is advantageous to will this. Finally, happiness itself is distinguished from the will to be happy because one can will something in order to be happy but the thing willed does not in fact bring about happiness.¹⁹

According to Anselm, every rational and non-rational creature with a will has the disposition to be happy.²⁰ Anytime a creature wills something, such as willing to eat a piece of cake, the disposition for happiness is effective in willing toward an end believed to bring about happiness. But rational creatures were created with an additional disposition, the will for justice. When an upright creaturely agent wills something justly, both dispositions work together to bring about that willing. Whereas one's willing is restricted by what one believes will bring about happiness, one's willing is not so restricted by the disposition for justice; the disposition for justice is not always effective in moving rational creatures to will. When one wills and the disposition for justice is ineffective, one wills some action in order to be happy but does so in a morally criticizable way. For instance, if I believe that eating your piece of cake will make me happy, the disposition for happiness inclines me to will that action. But the disposition for justice together with my knowledge that I should first obtain permission inclines me to refrain from willing that action. Given this conflict about what to do, I could will unjustly.

So even though 'freedom' is *defined* as the "power to preserve rectitude of will for its own sake," an *account* of creaturely freedom allows for creaturely agents to will unjustly when circumstances present a possible conflict between their two dispositions. This raises two questions: why did not God create rational agents with only one disposition, such as the disposition for justice; and if both dispositions are necessary, why not make creatures such that the two dispositions could not conflict?

In *De Casu Diaboli*, Anselm argues that both dispositions must be in a created agent if the agent is to will justly.²¹ He presents two thought-experiments of an angel created with only one disposition, one thought-experiment corresponding to each disposition. If God created an angel only with the disposition for happiness, the angel would not be able to will anything other than what it believes contributes to its happiness. Consequently, the angel could not will for the right reason and so could not be just. But more importantly, it could not be just because it would will its own advantage "out of necessity" and the willing would be the "work and gift of God." The same reasoning applies in the case of the angel created with only the disposition for justice, except that this angel could will for the right reason but still out of necessity. Therefore two dispositions are

¹⁹Concerning the point that the disposition for happiness and happiness itself can come apart, see Kane (1981, pp. 92–93).

²⁰*DCD* 12.

²¹*DCD* 13–14.

necessary if the angel wills justly.²²

The preceding reasoning also provides an answer to why God could not make a just agent with two dispositions that could not conflict.²³ A creaturely agent would not be just, Anselm believes, if “he received that willing in such a way that he would not be able to will otherwise.”²⁴ This statement falls within the context of the thought-experiment where an angel is given only the disposition for justice. By “being able to will otherwise,” Anselm has in mind a morally significant option, an unjust willing. If the angel has only the disposition for justice, to will otherwise is not to will some other just thing, but an unjust thing. So if a creaturely agent received two dispositions such that they could not conflict, he could will only what is fitting. But he would will something fitting out of necessity if he wills at all, and thus his willing would fail to be just. This yields a version of the principle of alternate possibilities: *if a creaturely agent is just, the agent must be able to be unjust.*²⁵

One can see why Anselm accepts this principle through the notion of attribution. Anselm takes there to be a relevant difference between created agents—bracketing those creatures in heaven who cannot sin now—and God, given that only the former requires the ability to will unjustly. Anselm writes, God possesses to a perfect degree what he possesses independently, he most of all is worthy to be praised for the good things which he possesses and keeps in his possession, doing this not out of any inevitable necessity, but as I have said earlier, out of an unchangeability which is his peculiar property and lasts for ever. (*CDH* 2.10)

Since God has his properties and abilities from no other source, it is not necessary that God be able to choose among morally significant options. A rational creature, by contrast, possesses neither its will nor the movement of its will independently.²⁶ The creature depends upon God as an external source because it receives its will from God and God has a causal role in moving the creature’s will towards justice. But for the creaturely agent’s willing to be “from itself” and for justice to be attributable to her, her willing must not be necessitated by God’s

²²Anselm’s thought-experiment corresponding to the disposition for justice is a counterpossible: it is not coherent to suppose that an angel really could have only the disposition for justice. Later, we will see that having this disposition is equivalent to having justice itself. So if an angel had only this disposition, the angel would be just. But Anselm argues that the will of an angel with only this disposition would be necessitated and therefore would not be just. Since being simultaneously just and not just is impossible, we should not take Anselm to be describing what he believes is a real possibility.

²³The claim that “God could not make creaturely agents with dispositions that could not conflict” should not be confused with “there could not be creaturely agents with dispositions that could not conflict *now*.” Anselm accepts that a subset of angels (i.e., those who did not sin in the first instant of their existence) and glorified human saints cannot sin *now*. But he accepts that they could have sinned at a previous time. Interestingly, he denies that their ability to sin in the past is relevant to their being praiseworthy when they cannot now sin, cf. *CDH* 2.10.

²⁴*DCD* 14.

²⁵This principle applies only to a subset of creaturely agents; cf. note 23.

²⁶*DCD* 20.

causal activity.²⁷ That is, if a creaturely agent's willing is just, there cannot be causal source outside of the agent sufficient to bring about her willing. Thus G. Stanley Kane writes of Anselm's view,

Self-determination in a creature (as opposed to self-determination in God) is possible only if there are alternatives from which he can choose. . . . If the creature performs a right act because he is compelled by some necessitating force to do so, then *he* cannot take the credit for it and *he* is not just for having done it.²⁸

Being able to choose between at least two options, then, is required for the *willing* to be attributable to the agent. But since the *justice* of the willing is also attributable to the agent, the agent must have been able to will a morally significant alternative.

Summarizing, Anselm's account of creaturely freedom explains how injustice is possible. Angels and humans were created with two dispositions that could come into conflict, thus allowing for them to will unjustly. Although the definition of 'freedom of choice' does not include the ability to sin, the account of creaturely freedom developed here requires that if a creaturely agent is just, the agent must be able to be unjust. Whichever way a creaturely agent wills must be non-necessitated. Only if these conditions hold can just willings and justice itself be attributable to the creaturely agent.

2. ON HARMONIZING FREEDOM WITH GRACE

2.1. *The Problem of Harmonization*

§1 examined Anselm's answer to the first question, how could evil come about if God created the world as good? §2 brings the preceding discussion to bear on the second question, how does human freedom harmonize with divine grace, if it does at all? The immediate concern in §2.1 is first to identify what Anselm takes to be the role of divine grace in restoring unjust agents to being just. This follows with determining the role of human freedom in restoring unjust agents. As a result, the problem of harmonization is stated explicitly.

Rectitude of will was described earlier as a condition of the will that enables one to preserve rectitude. That condition can now be identified more carefully. Whether one wills justly depends upon whether one has the disposition for justice. This is because the disposition for justice enables one to preserve rectitude of will for its own sake. According to Anselm, it is in virtue of having this disposition that a subject is just: "the will to be just is actually justice itself."²⁹ The surrounding context clearly shows that Anselm is not thinking of the will to be just as volition or as the will-instrument.³⁰ But without getting into textual

²⁷DLA 5 provides further discussion.

²⁸Kane (1973, p. 302).

²⁹DC 3.13. This need not conflict with the idea that an agent is just in virtue of its freedom, since whether one keeps this disposition depends upon the use of freedom.

³⁰"From these two affections, which we also call wills, derives all human merit, good or bad. [...] These two wills [...] differ in that the one that is for willing advantage is not itself the thing that it wills, whereas the one that is for willing rectitude is rectitude," DC 3.12; "God

analysis, one can see that it must be the disposition by the following reasoning. Given the supposition that justice applies to the will, justice applies to the will as either the will-instrument, the dispositional state aimed at justice, or to volitions.³¹ The will to be just cannot be the will-instrument since one retains the will-instrument even when one is unjust; and it cannot be the will as volition because we recognize that a person does not cease to be just when asleep, when no willing occurs unless dreaming. Therefore, the will to be just is the disposition for justice.

It follows that when one fails to preserve rectitude of will for its own sake because of sinning and thereby becomes unjust, one loses the disposition for justice. To explain, one has rectitude (i.e., has justice) when one has the disposition for justice. And we saw that everyone who has justice *is* just. Since Anselm connects statements like this with having eternal life (cf. *DC* 3.4) and says that not-having-justice implies being blameworthy (cf. §1.1), when one ceases to be praiseworthy on account of sinning, one loses the disposition for justice—i.e., one loses justice itself.

In order for rectitude of will to be preserved and for the agent to become praiseworthy again, the agent must acquire the disposition for justice. But because this disposition makes possible an agent's effectively willing some action for the right reason, if an agent lacks this disposition, the agent is unjust and not able to effectively use her power to will rectitude for its own sake. So, Anselm writes, the unjust agent is "a slave to sin because of the impossibility of recovering rectitude through its own power."³² It is for this reason that the grace of God is required in order to restore an agent who was once just but is now unjust. Anselm's account of the psychology of rational creatures provides us with a model that clearly identifies what it means for God to extend grace to someone: it is to restore the disposition for justice and thereby make the agent just and praiseworthy.

Anselm, however, denies that the restoration of (adult) humans is by grace alone:

When Sacred Scripture says something in favour of grace it does not at all exclude free choice, and in turn when it speaks in favour of free choice it does not dismiss grace, as though either grace alone or free choice alone is sufficient for salvation. . . . Assuredly (with the exception of what I said about the salvation of infants) the divine sayings are to be recognized as saying that neither grace alone nor free choice alone effects a person's salvation. (*DC* 3.5)

Why should salvation depend on free choice? One's salvation involves acquiring the disposition for justice, since having the disposition for justice entails being just and *vice versa*. But we saw in §1.2 that one cannot be just if being in

has ordered these two wills or affections so that the will which is an instrument might use the one which is justice," *DC* 3.13.

³¹Anselm affirms this supposition: "justice is not rectitude of knowledge or rectitude of action, but rectitude of will," *DV* 12.

³²*DLA* 10.

that state is necessitated by an outside source. Assume for *reductio* that God's grace were sufficient to cause the unjust agent to have the disposition for justice and that in fact God causes the unjust agent to have this disposition. It follows that the agent's having the disposition for justice is necessitated; so the agent is not just. But having the disposition for justice entails being just. Contradiction! Therefore, one cannot become just unless one's free choice is involved in acquiring justice.³³ Hence, salvation depends on human free choice.³⁴

The problem of harmonization can now be stated. In order for an unjust agent to become just and praiseworthy, both the grace of God and human freedom are necessary in bringing about that change. However, if an unjust agent is not able to preserve rectitude of will for its own sake, then supposing that an unjust agent is restored at some moment, it appears that the restoration at that moment is due to the grace of God alone. In that case the agent's becoming just is necessitated by something outside of her. But on Anselm's view such an agent would not be just or praiseworthy. The problem of harmonization, then, is the difficulty of giving an account of how an agent can change from being unjust and condemnable to being just and praiseworthy in a way that preserves a role for the grace of God and human freedom. A satisfactory account must recognize that the unjust agent's acquisition of the disposition for justice cannot be necessitated and she must have the ability to be unjust; otherwise, justice and praise will not be aptly attributable to her. The remainder of §2 will examine whether Anselm can resolve the problem of harmonization.

2.2. Pelagian Solutions

There is one way to answer the problem of harmonization, but it has consequences that Anselm would eschew. Suppose one rejected Anselm's claim that

³³In *DC* 3.3, Anselm writes, "a creature possesses the uprightness which I have called uprightness of the will only by the grace of God." It might be thought that this passage casts doubt upon my claim that free choice is involved in acquiring justice, since God is the only causal source for receiving the disposition. Thanks to Stan Tyvoll for raising this issue.

In response, note that Anselm continues in the same paragraph: "grace alone can save someone when free choice can do nothing, *as happens in the case of infants, whereas in the case of those who have the use of reason*, grace always aids one's innate free choice by giving it uprightness which it may preserve by free choice, because without grace it achieves nothing toward salvation," (italics added). So Anselm need not be interpreted as denying that an adult's free choice is involved in acquiring the disposition for justice. But even if God is the only causal source for receiving the disposition, Anselm still does not need to deny the role of freedom. Consider this analogy. Regardless of how many publications and outstanding teacher evaluations one receives, one may still be denied tenure. The only source of one's receiving tenure is the gracious will of the tenure committee. But one's use of freedom can still be involved in receiving tenure. For instance, one can refrain from the sin of criticizing the work of colleagues in print. So it does not follow from the fact that God is the only causal source of receiving uprightness that one's free choice is not involved in acquiring justice. One may still need to freely refrain from sin.

³⁴It is a consequence of this that if one does not freely do something while awake to enable God to give the disposition for justice, one cannot acquire the disposition for justice while asleep unless one is dreaming. For in the sleep state, one is not conscious and able to will unless one is dreaming. Cf. *DC* 3.11.

having the disposition for justice entails being just. An unjust person could then have the disposition for justice and the ability to do something that would contribute to her restoration. On such an account, one becomes just (restored) not by acquiring the disposition for justice, but by willing (choosing) in a way that keeps God's moral demands, such as by responding in the act of faith.

I will note two problematic consequences of this solution from standpoint of Anselm's Christian theology. To begin, it becomes unclear why one would ever lose the disposition for justice on account of sinning. Anselm's view explains why one would lose it, namely, because having that disposition is having justice itself. But the solution under consideration denies that claim in its aim to preserve the disposition for justice when an agent is unjust. One way to preserve it is to regard it as a natural (i.e., essential) part of human nature, implying that it could not be lost. That would be a form of Pelagianism in the garb of Anselm's moral psychology.³⁵ According to Pelagius, "human nature ensures a permanent capability for sinlessness, and from this both will and act can follow."³⁶

The first problematic consequence of Pelagianism is rejecting, *inter alia*, the doctrine of original sin. According to one part of the doctrine, at least under Augustine's formulation, human beings received an "inherited concupiscence" after the fall of Adam; that is, they received a desire for sin which placed them in a state of moral poverty.³⁷ Not all medieval philosophers formulated the doctrine through Augustine's notion of concupiscence, Anselm and John Duns Scotus being two examples.³⁸ But differences of detail aside, the important point is that the doctrine in its various formulations implies that one lacks the ability to will justly; and of course, lacking the ability to will justly does not imply lacking the ability to will what is fitting, as Anselm's example of giving money to the poor illustrates.

The doctrine of original sin implies that God's grace is necessary for restoration. Pelagius did not deny a role for grace, but he understood grace to be God's act of creating human nature with its capabilities, including the capability to will rightly, and providing moral examples like the Mosaic Law and Christ's sacrifice. But the sense in which God's grace is given is not one that restores the agent from being unjust to being just; it only makes restoration possible. This yields a second unpalatable consequence of Pelagianism: the agent must do something good on her own and thereby merit salvation. This is possible according to Pelagianism (described through Anselm's moral psychology) because the

³⁵Although I cannot delve into a historical discussion of Pelagianism, Pelagianism has had its share of notable critics, particularly Augustine. Later Pelagians such as Caelestius and Julian were sharply criticized by numerous church councils, e.g., at the third ecumenical Council of Ephesus in 431 and the non-ecumenical Council of Orange in 529, which was ratified in 531. For more on the Pelagian controversy, see Schaff (2002, pp. 783–815), TeSelle (1999), and Leyser (1999).

³⁶TeSelle (1999, p. 635).

³⁷Burnell (1999), Rigby (1999).

³⁸Cross (1999, pp. 96–100).

disposition for justice cannot be lost. For anyone trying to solve the problem of harmonization within the confines of orthodox Christian theology, as Anselm is, these consequences are unacceptable.

It might be thought that Pelagianism is *not* a consequence of rejecting Anselm's claim that having the disposition for justice entails being just. In a recent article, C. P. Ragland proposes a solution to the problem that attempts to avoid Pelagianism while retaining an incompatibilist form of freedom like that outlined in §1.³⁹ Although Ragland is not concerned with Anselm in his article, his proposal can be adapted into our discussion quite easily. He follows John Wesley by distinguishing between prevenient grace and convincing grace. Out of God's complete goodness, God gives prevenient grace to everyone, which is necessary but not sufficient for faith; and this grace is in no way merited by the deeds of fallen agents. According to Ragland, "This grace gives [fallen human agents] the ability—absent from fallen human nature by itself—to trust Christ. Whether people choose to *exercise* this ability to having saving faith is a matter of libertarian freedom in the fullest sense—a matter of deliberate choice." Prevenient grace gives them the ability to trust Christ since God's prevenient grace involves "implanting in them the nonnecessitating inclination to accept God's offer of convincing grace." So in our context, the role of prevenient grace is to restore the disposition for justice. The role of convincing grace is to reconcile the human agent with God in response to the human agent's free choice. The moment of receiving convincing grace is the moment when the human agent becomes just.

Several things separate Ragland's proposal from traditional Pelagianism. Ragland does not accept Pelagius's view that God gave human nature the essential ability to will justly. Unlike Pelagianism, one can lose the disposition for justice, and it is "prevenient grace" that restores the broken will of the human agent. In addition, Ragland introduces a kind of grace, "convincing grace," that saves human agents, whereas Pelagius's form of grace only makes salvation possible, contingent upon the free choice of the human agent.

Despite these differences, Ragland's proposal still suffers from at least one of the problems attributed to Pelagianism.⁴⁰ Consider Pelagius's denial that human kind inherited a fallen nature; that is, the denial that human nature lost the disposition for justice. It is nevertheless the case on Ragland's view that God provides prevenient grace to *everyone*. Since this comes in the form of restoring the disposition for justice, it follows that God restores human agents back to the same prelapsarian state of Adam. If prevenient grace is given immediately after Adam's sin, and to all mortal humans, it may be more apt to say that God's prevenient grace preserves the disposition for justice because postlapsarian human

³⁹Ragland (2006). All citations of Ragland follow pp. 360–362.

⁴⁰Corresponding to the problem of whether one can perform some good to cause one's salvation, see Timpe (2007, 285–86, 296 n.14). Timpe rejects Ragland's proposal since he believes that it implies one can be a cause of one's own salvation through performing a good apart from grace. Whether this is so on Ragland's view depends upon which grace is in view. I ignore this complication by raising a different objection.

nature is *eo ipso* insufficient to keep it. In any case, although the human nature could have been damaged, God's prevenient grace ensures that it is not actually damaged.⁴¹ Whereas Pelagius's view of grace makes a morally impoverished human nature impossible, Ragland's view of grace only makes it an unactualized possibility, a modally weaker version of Pelagianism in this respect.

Suppose that Ragland were to restrict the role of prevenient grace to God giving *only* the ability place faith in Christ. Prevenient grace would not provide the ability to will justly in various other ways, which might be restored only after convincing grace is received. But if this were the reply Ragland would make to preserve speaking about broken wills and original sin, it would be *ad hoc*. If the goodness of God requires that prevenient grace be given in at least this restricted sense, as Ragland suggests, then why does not the goodness of God also require the giving of a non-necessitating inclination to will rightly in general? Giving this further inclination would in no way violate freedom of choice; it would enhance it. People could still will unjustly just as Adam could prior to the Eden-incident and just as they can do after receiving convincing grace; but they could all the more preserve rectitude of will for its own sake and avoid causing further evils. Thus it seems to me that if the goodness of God is invoked as evidence for prevenient grace in the restricted sense, it is equally evidence for the restoration of the disposition for justice. At least one aspect of Pelagianism, then, is an unintended consequence of Ragland's solution.

We have been exploring the consequences of an answer to the problem of harmonization, which involves rejecting Anselm's claim that having the disposition for justice entails being just. The upshot is that rejecting this claim requires jettisoning other claims that Anselm believes to be part of an important theological inheritance. So given that we are concerned with whether Anselm can provide compatible answers to this paper's two opening questions, we will assume that a solution to the problem must fit within the confines of Anselm's theological context. Thus, we will assume that having the disposition for justice entails being just. The remainder of §2 will examine attempts to resolve the problem of harmonization without embracing Pelagianism.

2.3. *The Ability To Do Otherwise Solution*

Katherin Rogers argues that Anselm has a successful answer to the problem of harmonization.⁴² She recognizes that "all the causal power to produce a new good, in this case the *affectio* for justice in a fallen soul, belongs to God." She also admits, "God's restoring justice to the fallen creature entails that it now desires to keep rightness of will for its own sake." But given that having the disposition for justice entails being just and it is God who restores the disposition, it appears

⁴¹At least, it is not any less damaged than Pelagius would admit. Pelagius recognized that sin can have bad effects upon the will through habit, whereas later Pelagians went further by denying any effects of sin upon the will. See Schaff (2002, 804 n.1).

⁴²Rogers (2008, 140-41; cf. 78). Unless otherwise noted, all quotes in §2.3 come from Rogers's 2008 work in the pages noted here.

that God's act of restoring the fallen agent necessitates her being just, which is supposed to be impossible. So what is the role of human freedom in making justice attributable to the agent?

Rogers appeals to Anselm's *De Casu Diaboli*, where he says that an angel was "able to give himself justice, since he was able to take it away from himself and also able not to take it away."⁴³ More generally, an agent can be said to cause or do something when she could fail to cause or do it, but does not. Thus Rogers writes,

In morally significant choice, as in the choice to keep or abandon the justice restored by grace, it is indeed up to the creature that one desire wins out over the other. . . . It is up to human free will to keep the justice which grace has restored to it. . . . Anselm's claim is that God gives fallen humanity the grace that is necessary for salvation, and we can choose, on our own, to keep it or throw it away. . . . The best the fallen human being imbued with grace can say for himself is that perhaps he is managing to refrain from being so stupid and so wicked as to throw away the entirely unmerited divine gift of grace.

Since the agent has the ability to be unjust but does not will unjustly, justice can aptly be attributed to her.

It is tempting to interpret Rogers's claim that the agent has the ability to keep justice or throw it away as an ability the agent has *after* God causes her to have the disposition for justice. But at least three reasons can be given to reject this interpretation. (1) In an earlier discussion of "Frankfurt cases," Rogers argues that an angel is free to will otherwise up to and *including* the very moment of choice.⁴⁴ So in our discussion, this suggests that the restored agent has the ability to will unjustly at *the very same moment* God wills that she have the disposition for justice. (2) If an agent does not have the ability to be unjust at the very same moment God gives the disposition, then the agent's having the disposition would be necessitated. This would generate the same *reductio* argument presented in §2.1 for why the agent's salvation depends on free choice. (3) If an agent were not able to be unjust at the moment God gives the disposition, it follows by the principle of alternate possibilities above that the agent would not be just until the following moment. But then Rogers would have to deny that having the disposition for justice entails being just, and we saw in §2.2 that this denial leads to Pelagianism. So Rogers and Anselm are committed to saying that the agent is able to will unjustly at the moment of restoration.⁴⁵

⁴³*DCD* 18.

⁴⁴Rogers (2008, pp. 78–81).

⁴⁵Three readers of earlier versions of this paper each suggested that we must speak about two instants concerning the process of an unjust agent becoming just. God provides the disposition for justice at the first instant, and the agent can preserve or reject that disposition only at the second instant. According to this view, one can obtain a sort of justice when the disposition is restored, but it is not a justice connected with being praiseworthy since only God is causally responsible for acquiring the disposition. However, a choice can be made in the second instant, so that the kind of justice connected with being praiseworthy arises if rectitude is preserved then. Thanks to Joseph Dowd, Katherin Rogers, and Stan Tyvoll for convincing me to discuss this view. My response is two-fold.

Having the ability to be unjust at the moment of restoration introduces a problem: in what sense does an agent have the ability to will unjustly given that God's causing her to have the disposition for justice entails that she is just at that moment? Although neither Rogers nor Anselm specify what sense of 'ability' is relevant at that moment, we can distinguish between at least two senses of 'having the ability to will otherwise', 'could have willed otherwise', and similar locutions.⁴⁶ I will argue that neither sense is adequate for Anselm's purposes.

The first sense involves having a *general ability*. A general ability is a power one has even if the power is not exercised or used at some particular time. Anselm's analogy of being able to see a mountain even when there is no nearby mountain to be seen is an illustration of having such a power. Depending on the general ability in question, general abilities may be lost or acquired. The power to play the guitar might be lost if one loses two fingers in an accident.⁴⁷ The power to preserve rectitude of will for its own sake, by contrast, cannot be lost because it is an essential part of human nature. Turning to the general ability to will unjustly: one retains this ability at the moment of restoration. If God does not will to restore the agent at some moment, then the agent can will unjustly on her own. When God wills to restore the agent, this general ability is not destroyed on account of the ability not being exercised at that moment.

However this cannot be the relevant sense in which an agent is able to will unjustly. Suppose that God creates a deterministic world in which no event contravenes his moral will. Creaturely agents could still have the general ability to will unjustly, though the circumstances are not such that this ability would

Textual Objection: The introduction of a kind of justice connected with the disposition but not with praiseworthiness is not explicitly found in Anselm's work. The present account, it seems to me, is indirectly motivated by an unnecessary interpretation of certain passages (c.f. note 33) together with trying to resolve the tension raised by the problem of harmonization. Moreover, the present view is explicitly denied by Anselm: see note 57 and the surrounding context in *DV* 12 where he connects having rectitude, i.e., justice, with being praiseworthy. Since the receiving, having, and willing of justice take place simultaneously and having justice is connected with being praiseworthy, it appears that the two-instant view is not Anselm's.

Philosophical Objection: If one could have the disposition for justice without being praiseworthy, it is puzzling why one would ever lose the disposition on account of being condemnable through sinning. On my sketch of Anselm's view, ones having or lacking the disposition for justice is tied to the sort of moral appraisal one receives. Sinning and becoming condemnable removes the disposition just as much as having the disposition makes one praiseworthy. The two-instant view denies this connection between having the disposition and one's moral appraisal. So as far as I can see, it lacks an explanation for why one loses the disposition when being condemnable. Thus, claiming that one would lose the disposition through sinning seems *ad hoc* on the two-instant view in a way that it does not on my interpretation.

⁴⁶Of course, there are more senses of 'ability' than the two I will distinguish. I have chosen the two most plausible senses of 'ability' relevant to the moment of restoration. My discussion has benefited from and follows, to a large extent, Campbell (2005), which goes into further detail about the senses of 'ability' in the contemporary free will literature.

⁴⁷Campbell (2005, p. 399): "At an early age, the jazz guitarist Django Reinhardt was a virtuoso. When he was 18 he lost the general ability to play the guitar due to injuries suffered in a fire. Later he relearned to play the guitar using only eight fingers, for two of his fingers were paralyzed in the accident."

ever be exercised. The cost of taking this interpretation of ‘ability’ at the moment of restoration, then, undercuts Anselm’s argument for how freedom explains the possibility of evil. For according to this interpretation of ‘ability’, agents are able to be unjust in worlds that would never include a single evil (setting aside natural evils). But then why did not God create a world like that? If the appeal to freedom explains the possibility of evil, another sense of ‘ability’ must be relevant.

The second sense of ‘ability’, the *all-in* sense, is more amenable to incompatibilism about freedom and determinism. Let Ψ represent the complete set of truths about the past together with the laws of nature that hold relative to a time when a choice is made. If an agent’s willing A at t , the time of choice, is inconsistent with everything entailed by Ψ , then S is not able in the all-in sense to will A at t . So if the laws of nature and the past together entail the falsity of S *does not will justly at the moment of restoration*, then S is not able to will unjustly at that moment.

The all-in sense of ‘ability’ does not help Anselm or Rogers resolve the problem of harmonization. We are considering whether an agent is able—in the all-in sense—to be unjust at the very same moment she wills justly. Even if the past with the laws were not inconsistent with either S *wills justly at t* or S *does not will justly at t* , there is something true at t that seems inconsistent with one of these propositions. The true proposition that S *wills justly at t* is inconsistent with S *does not will justly at t* . But of course the right characterization of all-in abilities should exclude the fact of which choice is made when determining whether or not an agent is able to will otherwise. Nevertheless, there is another fact true at t which is inconsistent with S *does not will justly at t* : it is the fact that *God causes S to have the disposition for justice at t* . Since that fact entails that S *is just at t* , it follows that S *is not able in the all-in sense to be unjust at t* . Were the fact that God causes the fallen agent to have the disposition for justice excluded from the set that determines whether an agent is able to be unjust at t , the agent would indeed be able to be unjust at t . However, she would then lack the ability to be just at t , and so would not be restored.

It appears that there is no relevant sense in which an agent is able to be unjust at the moment of restoration compatible with our discussion in §1. So if God restores an agent to justice, the principle of alternate possibilities, i.e., *an agent is just only if the agent is able to be unjust*, is violated and thus false. But since Anselm is committed to this principle, it follows that one cannot change from being unjust to being just. This would be a devastating result because his theology claims that God the Son became incarnate and died in order to make unjust agents just. The only alternative to adopting Pelagianism in some respect, on the one hand, or adopting compatibilism about God’s causal activity and human freedom and responsibility, on the other hand, is to find something that an agent can freely do as a precondition for God causing the disposition for justice in her. If there is something an agent can freely do prior to or during the moment of restoration which would enable God to bring about this disposition

in her, then it is true that the agent has the all-in ability to do otherwise. So the ability to do otherwise solution would only be incomplete as it stands. The next section examines a solution that, if successful, would complete the present solution.

2.4. *The Quiescence Solution*

In the contemporary literature, G. Stanley Kane first suggested the quiescence solution in addressing problem of harmonization for Anselm; unfortunately, it received no further attention until its articulation by Eleonore Stump when writing on Augustine and Aquinas twenty years later.⁴⁸ Although there is an important difference between Kane and Stump's versions, both take the central idea behind the quiescence solution to be that there is an alternative to accepting grace and rejecting grace: it is being quiescent. For one to be quiescent means that one refrains from rejecting grace (i.e. refrains from sinning) by "not doing anything at all."⁴⁹ That is, one refrains from making a sinful first-order volition without also making a just first-order volition. On Anselm's view, quiescence is a state in which one's will-instrument is not being used for justice or for injustice.⁵⁰

Stump explores different ways in which one's will can become quiescent, but only one way is relevant to acquiring a justifying faith (i.e. becoming restored). On Stump's version of the quiescence solution, following her interpretation of Aquinas, the intellect becomes divided against itself so that a subject's will moves from a state of rejection to a state of inactivity. To see how this works, consider someone with a phobia of needles who is going to a doctor for a shot.⁵¹ The phobic's intellect represents the injection as harmful to her and so her will is opposed to receiving it. But the doctor may exhort her to accept the injection for reasons beneficial to her. Though the reasons given do not necessitate that she accept the injection, the reasons may be weighty enough for her to be unsure whether she should accept it or reject it. If the reasons are weighty enough (and it is contingent whether they are weighty enough), Stump writes, "the intellect becomes locked in indecision, unable to resolve the conflict within itself into one single, integrated judgment. In the face of this blockage in the intellect, the phobic's will becomes quiescent."⁵² When the intellect is divided, the phobic forms a higher-order desire for a will that wills to assent to receiving the shot; this is not forming a first-order volition to receive the shot. So since this higher-order desire conflicts with the desire to reject the shot, the will neither accepts it

⁴⁸Kane (1981, pp. 159–179); ?; Stump (2003, pp. 389–404); Tyvoll (2004); Ragland (2006); and Timpe (2007).

⁴⁹Kane (1981, p. 166); cf. Stump (2003, p. 394).

⁵⁰It is not clear whether Anselm can accept the quiescence solution. Rogers (2008, pp. 137–139) argues that Anselm cannot accept it since it conflicts with his other commitments. Although I find Rogers's arguments plausible, I believe that Anselm needs this solution in order to make room for having the ability to do otherwise in the all-in sense.

⁵¹This example is given in Stump (2001, p. 140) and Stump (2003, pp. 398–399).

⁵²Stump (2003, p. 399).

nor rejects it. Thus Stump claims, “it is appropriate to describe the change of the will to quiescence as the expelling or driving out of the preceding rejection.”⁵³

A crucial feature of Stump’s version is that the higher-order desire is not itself an act of will. Furthermore, the change from a state of rejection to a state of quiescence does not involve a decision on the part of the agent. As a result, Stump argues that the change from a state of rejecting God’s grace to becoming quiescent with regard to grace does not involve the agent willing some good: Consequently, without risk of falling into Pelagianism, we can suppose that it is up to the human willer, and to her alone, whether her will refuses grace or is quiescent with regard to grace. As I have been at pains to show, this is not to say that the human willer at issue looks at the options of refusing grace or becoming quiescent with regard to grace and forms a decision about which of the options should characterize her will. Rather, it is to claim just that control over whether her will acts or fails to act is vested ultimately in her.⁵⁴

When the agent enters a state of quiescence on her own by ceasing to reject God’s grace, God can simultaneously cooperate with her by bringing about a will of faith in her at that moment.

How can God simultaneously cooperate by giving the will of faith, which implies the acceptance of grace at that time, while it also being true that the agent becomes quiescent (i.e. not yet accepting of grace)? The problem is removed by making a distinction between *natural priority* and *temporal priority*. Scotus argued that it is possible for causes and effects to coincide at an instant, but the causes are in a sense prior to their effects:

No cause produces its effect if it is not prior (to it) by nature. It (need not) be prior in time. Even if (it is the case that) the cause were not prior in time before it is causing (its effect) it would still (have to be) prior by nature (to its effect) (which it is causing).⁵⁵

Applying Scotus’s distinction to the quiescence solution, the agent’s becoming quiescent is naturally, but not temporally, prior to God giving faith.⁵⁶ So if change at an instant is possible, there is no contradiction in the quiescence solution.⁵⁷

⁵³Ibid.

⁵⁴Stump (2003, 402).

⁵⁵*Opus Oxoniense* II d5 q2 n6. The source and translation are from Sylwanowicz (1996, p. 90).

⁵⁶On this view, the contribution by the human agent is naturally prior to the action of the divine agent. The alternative of making the act of the divine agent naturally prior to the contribution by the human agent is irrelevant. After all, we are considering the quiescence solution in order to find a way for the agent to do something as a precondition for God giving the disposition for justice.

⁵⁷Anselm seems to have anticipated Scotus’s distinction in *DV* 12: “Indeed, just as the receiving of this rectitude is prior in nature to having or willing it (since neither having nor willing it is the cause of receiving it, but receiving it is the cause of willing and having it) and yet receiving it is temporally simultaneous with having and willing it (since we simultaneously begin to receive it, to have it, and to will it, and no sooner do we receive it than we have it and will it), so also having or willing it, although prior in nature to preserving it, is nonetheless temporally simultaneous with preserving it.”

At first glance, Stump's solution appears attractive for resolving the problem of harmonization in Anselm. It is contingent whether the fallen agent becomes quiescent, and whether she becomes quiescent ultimately depends on her. Since her becoming quiescent is a precondition for God's acting to restore her, she can have at the moment of restoration an ability to be unjust in the all-in sense. She has veto power over whether God restores her by the fact that she can continue to sin rather than be quiescent. Whether she becomes just by God's grace or remains in a state of rejection is not necessitated by anything outside (or inside) of her.

Despite these virtues, Anselm should reject Stump's solution. One reason is that the fallen agent is too passive in changing from a state of rejection to a state of quiescence.⁵⁸ Even though the change from a state of rejection to a state of quiescence is contingent, Stump represents the change as one not involving a choice. Since no choice is involved, the change is not the result of the activity of an agent on the basis of a reason, which Anselm requires for praiseworthiness. It is true that the agent's intellect acquires conflicting reasons for contrary acts, but the agent's will need not be involved in the acquisition of these reasons. Reasons may be acquired by being dragged to the doctor and forced to listen or by a missionary who places his foot in the doorway to continue preaching. Nor is the agent's will involved in forming the higher-order desire for a will to accept the injection or to receive God's grace; it forms merely as a result of the division in the intellect. So although it is true that an agent would not have become quiescent if she did not refrain from sinning, the way in which she refrains is not, by Anselm's lights, the kind of control adequate to make justice and praise attributable to her.

Kane's version of quiescence avoids Stump's problem. He writes, [i.e., the role of the human will in one's restoration] is an exercise of choice between moral alternatives, because under the circumstances in question one is not forced to choose something that is unjust and one is not forced to refrain from choosing something unjust; one may do either as one chooses.⁵⁹

When one chooses to refrain from willing unjustly, God can then restore the unjust agent by giving the disposition for justice. But Kane's version suffers from the problem that Stump explicitly tries to avoid.

If a choice is made to change from a state of rejecting God's grace to not actively rejecting it (i.e. by not doing anything at all), it appears that the agent does something good. But this looks like the second Pelagian error noted in §2.2, where a fallen agent performs some good in order to initiate the reception of God's favor. Kane attempts to avoid Pelagianism when he writes, [i.e. the role of the human will] is essentially a negative role, because it involves not doing something that one could do [i.e. sinning].... It [i.e. this quiescence solution] makes it legitimate to attribute everything in the production of just volitions to God,

⁵⁸This sort of objection appears in Ragland (2006, pp. 356–359) and Rogers (2008, p. 139 n.7).

⁵⁹Kane (1981, p. 165).

for everything positive, i.e. everything that involves a doing rather than a refraining from doing, is done by God.⁶⁰

This response is puzzling. It may make sense to describe refraining from sinning as having a “negative” role, since the agent does not perform any first-order volition in that state. But is not the choice to enter the quiescent state something “positive”? Unfortunately, Kane relies too heavily on the distinctions between *doing* and *refraining*, and *positive* and *negative*, without providing much help in understanding what exactly these notions mean with respect to the agent’s choice. Rather than speculate about his terms, there is a deeper problem that bypasses this difficulty and shows there is nothing an agent can freely do to enable God to make justice attributable to her.

Whether justice and praise are aptly attributable to an agent depends on whether the agent wills something for the right reason (cf. §1.1). So whether justice and praise are aptly attributable to the agent at the moment God restores the disposition for justice depends on the fallen agent’s reason for choosing to enter the quiescent state. There are three options here: the reason is just, unjust, or neutral with respect to justice. None of these reasons are adequate for making justice and praise attributable to the agent at the moment of restoration, as I will now argue.

The reason to refrain from sinning cannot be for a just or an unjust reason. The reason to refrain from sinning cannot be for a just reason because the agent lacks the disposition for justice. The disposition for justice makes possible an agent’s willing or choosing for the sake of justice itself. But the agent’s choosing to enter a state of quiescence is supposed to make possible the agent’s recovering the disposition for justice. Thus this explanation for how the agent recovers justice is circular. Refraining from sin cannot be for an unjust reason because the agent would add new sin while attempting to cease sinning. So the reason the agent chooses to be quiescent must be for a reason that is neither just nor unjust.

Choosing to be quiescent for a morally neutral reason, however, is not the kind of reason sufficient to enable God to change the fallen agent’s will, such that the fallen agent can be restored to justice and be praiseworthy. Suppose that one chooses to be inactive rather than steal a neighbor’s pears for the reason that stealing pears is not interesting enough. Of course, by ‘not interesting enough’ I do not mean ‘not sinful enough’. One can choose to not will an action, where there is no obligation to will that action, for the reason that the act does not strike one as interesting without committing a further wrong, e.g., refraining from counting from 1000 to 0. Likewise one commits no wrong in not stealing pears due to a lack of interest. The problem is that if one refrains from sinning for this reason, it is not clear how God’s changing the will of an agent is any less a form of encroached manipulation than if God acts when one enters a state of inactivity because of simple inattention about a sinful act. In order for

⁶⁰Ibid.

God's causing the disposition for justice to not be a form of manipulation, the agent would have to know, or at least believe, that entering a state of inactivity enables God to cause the disposition to exist in her. However if this knowledge figures in as a reason for entering the state of inactivity, the agent chooses to refrain from sinning because it will bring about something just. And choosing for this reason, we saw, is not available to the unjust agent. Therefore, given that incompatibilism about human freedom and moral responsibility rejects the notion that God can make agents just through manipulating their wills, choosing to refrain from sinning for a morally neutral reason is not adequate to enable God to cause the disposition for justice in her.

It appears that there is no reason for which an agent can choose to be quiescent sufficient to enable God to cause the disposition for justice in her without violating her free agency. So no quiescence solution is available to give the human will an adequate role in receiving the disposition for justice. Thus neither Kane nor Stump's solutions to the problem of harmonization are adequate to complete the ability to do otherwise solution.

2.5. Concluding the Problem of Harmonization

The problem of harmonization is the difficulty of giving an account of how an agent can change from being unjust and condemnable to being just and praiseworthy in a way that preserves a role for the grace of God and human freedom. I have argued that since there is no reason for which an agent can choose to be a particular way sufficient to enable God to change her from being unjust to being just consistent with incompatibilism, it appears that there is no way for Anselm to reconcile human freedom and God's grace.⁶¹ So Anselm was not successful in providing compatible answers to our two opening questions: why is there evil, and how does human freedom harmonize with God's grace?

Where does this leave Anselm and those within his theological tradition who appeal to a similar account of incompatibilist freedom to explain the presence of evil? Notice that the problem of harmonization results from the apparent truth of the following four theses:

- (1) Pelagianism is false with respect to (i) there is no original sin; and (ii) one becomes just by willing some good;
- (2) One cannot become just if being in that state is necessitated by something outside the agent; *Corollary*: One is just only if one is able to be unjust;
- (3) There is nothing a human agent can freely do to enable God to change her will to being just;
- (4) God's grace is necessary to change, and in fact changes, one from being unjust into being just.

Anselm already accepts (1), (2), and (4). But I have argued that (3) is true by canvassing the best available non-Pelagian solutions to the problem of harmonization and finding that none are adequate. The conjunction of (3) and (4)

⁶¹Cross (2005) presents six accounts—some of which resemble those already discussed here—in which God gives a person grace but the grace is resistible. All six accounts, however, suffer from one or more of the difficulties identified with the accounts I have considered.

implies the falsity of (2). For if God's grace is necessary to change an unjust agent into a just agent (i.e. (4)) and it is not true that human freedom can contribute to one's salvation (i.e. (3)), then if one is restored at any time, God's grace is sufficient for one's restoration. So (2) is false. One might try to find some non-Pelagian way of showing (3) is false; but given the arguments above, I believe this route is not very promising.

Rejecting (4) would be the least attractive option for Anselm since it requires rejecting so many other doctrines central to the Christian worldview. That leaves open rejecting (1) or (2). If (1) is rejected by accepting Pelagianism in some respect, then one could show (3) is false. But whereas nearly everyone in Anselm's theological tradition rejects Pelagianism (i.e. accepts (1)), not everyone rejects (2).⁶² This is because rejecting (1) comes with costs considered too great for the tradition to accept. So Anselm and incompatibilists within his theological tradition should become Pelagians and accept the costs of that view, or they should reject (2) by denying incompatibilism about God's causal activity and human freedom and responsibility.

3. ON THE PROBLEM OF JUST CREATION

If the arguments of §2 are sound, then it is possible to provide a quick second argument for the claim that one must either reject incompatibilism about creaturely freedom and responsibility and God's causal activity, or accept Pelagianism in some respect. What creates the difficulty in the previous section is that there is nothing a creaturely agent can freely do as a precondition for God giving the disposition for justice and making her just. So given that God causes her to be just at some time, God acts as a sufficient cause. But that is inconsistent with incompatibilism. This difficulty also arises in the case of God creating an agent and asking about the agent's moral status in the first instant of her existence. Such a case was discussed in the Patristic and High Middle Ages and only recently revived.⁶³

The problem begins with the admission that creaturely agents, whether humans or angels, were created as just in the first instant of their existence prior to the fall of Adam. This admission cuts across the divide between compatibilists and incompatibilists, and in those for whom it is less clear whether freedom and moral responsibility are compatible with necessitation. Here are three prominent examples:

But if it [i.e. a will] was created, was it created at the same time as they were, or did they first exist without it? If it was created with them, then there is no doubt that it was created by Him Who created them. And, as soon as they were created, they clung to Him Who created them with the love He created in them. . . . If, however, the good angels first existed without a good will, and then produced it in themselves without

⁶²Augustine is one possible example; cf. Rogers (2008, pp. 30–54). Also, see van Asselt et al. (2010).

⁶³The recent literature that resembles the scholastic debate is found in Campbell (2007, 2008), Brueckner (2008), Bailey (2012).

God's agency, they thereby made themselves better than He made them. God forbid! ... Hence, we must believe that the holy angels were never without a good will: that is, the love of God. But the angels who, though created good, have nonetheless become evil, became so by their own will... (*Augustine*)⁶⁴

The will as a tool was created good in respect to its being. It was also created just, and able to preserve its received righteousness. However, it became evil by its free choice... (*Anselm*)⁶⁵

Therefore, as all were created in grace, all merited in their first instant. But some of them at once placed an impediment to their beatitude, thereby destroying their preceding merit; and consequently they were deprived of the beatitude which they had merited. (*Aquinas*)⁶⁶

Pelagius, by contrast, denied that rational creatures were created as just. Augustine quotes Pelagius as follows: Everything good, and everything evil, on account of which we are either laudable or blameworthy, is not born with us but done by us: for we are born not fully developed, but with a capacity for either conduct; and we are procreated as without virtue, so also without vice; and previous to the action of our own proper will, that alone is in man which God has formed.⁶⁷

If the non-Pelagian account of creation is that one is created as just, how can justice be attributable to a creaturely agent at the first instant of its existence? I will follow the medievals by limiting my discussion to the creation of angels and asking about the attribution of justice to them in the first instant of their existence.

From our discussion of Anselm's account of freedom in §1 and the attempts to resolve the problem of harmonization in §2, we should expect an incompatibilist about freedom and moral responsibility and God's causal activity to say that justice is attributable in the first instant of an angel's existence only if:

(1) The angel is able in the all-in sense to be unjust in the first instant but is not unjust; or

(2) The angel is in some sense the source of its being just in the first instant because it could do something that ultimately contributes to its being just in the first instant.

If neither (1) nor (2) is true of an angel at the first instant of its existence, and non-Pelagianism about creation is correct, then it follows that one must reject incompatibilism. For if neither (1) nor (2) is true, God's creating the angel as just through giving the disposition for justice in the first instant is necessitated by a source outside of the angel.

Beginning with (1): can an angel be unjust in the first instant of its existence?

⁶⁴Augustine, *City of God*, Bk. 12, Chap. 9.

⁶⁵*DC* 3.13.

⁶⁶*ST* Ia q.63 a.5 re.4

⁶⁷Augustine, *On the Grace of Christ*, and *On Original Sin*, Book 2, Chap. 14, in Schaff (2007, p. 241).

As Aquinas recognizes, opinions vary here.⁶⁸ Aquinas and later Thomists argued that an angel cannot sin in the first instant, but the angel still had a *rectitudo* that was meritorious. Aquinas's argument relies on the premise that an angel's beginning to act at the first instant depends upon the nature of the entity from which the angel drew its existence. God, who cannot cause sin, created the angel. Therefore, the angel cannot sin in the first instant. Scotus, by contrast, denies Aquinas's first premise by offering an account of the will as an instantaneous self-moving process.⁶⁹ For Scotus, one has the power to will $\sim A$ even while willing A . But Scotus, however, is plausibly interpreted as compatibilist who regards the power to will A and the co-present power to will $\sim A$ as consistent with being necessitated by an outside source.⁷⁰ If one has the power to sin at the first instant, this power would only count as a general ability. So if God's will is sufficient for bringing about the state of the world at its first instant, which includes giving the disposition for justice, it appears that an angel cannot be unjust in the all-in sense at the first instant.

What about (2): is there something an angel can do at the first instant, which is naturally prior to, but simultaneous with, God's giving the disposition for justice? The only option I can see is that the angel refrains from sinning in the first instant. But now we encounter the same problem raised against Kane's quiescence solution. If the angel chooses to refrain from sinning, such that God can then give the disposition for justice, justice is attributable only if the angel refrains for the right reason. But no reason is available. If the angel refrains for an unjust reason, the angel is not just at the first instant, which is contrary to the admission by Augustine, Anselm, and Aquinas. The angel cannot refrain for a just reason since the refraining occurs naturally prior to receiving the disposition for justice. Finally, if the angel refrains from sinning for a morally neutral reason, it appears as if God's making the angel with the disposition for justice is a form of manipulation, except that in this case it is creating rather than reshaping the will of an agent. Therefore, the angel cannot contribute to making its will just in the first instant.⁷¹ Hence, given that the angel is just in the first instant, God is a sufficient cause for its being just in that instant.

Of course, I am not in a position to claim that compatibilism is true by the above reflections, for I have not argued that Pelagius is wrong with respect to his view about creation. So I present my conclusion as a disjunction: either one should accept Pelagianism with respect to creation or accept the compatibility of God's causal activity with human freedom and moral responsibility.

⁶⁸Aquinas discusses this question at *ST* Ia Q63 Art. 5.

⁶⁹For more on Thomistic arguments and Scotus's idea of the will, see Sylwanowicz (1996, chapter 4).

⁷⁰See Sylwanowicz (1996, chapter 7).

⁷¹For a striking comparison, see Anselm's thought experiment of God creating an angel piece-by-piece in *DCD* 12.

CONCLUSION

I have argued that Anselm's appeal to an incompatibilist account of freedom in order to explain the presence of evil raises two problems: the problem of harmonization and the problem of just creation. Both problems involve cases where an agent becomes just at some time but is unable to be unjust at that time and in the sense needed for incompatibilism to be true. In addition, these are cases in which there is nothing the agent can do to make justice aptly attributable to her in virtue of her free choice; so given that God acts to make the agent just, God acts as a sufficient cause. This result is inconsistent with incompatibilist accounts of freedom, like Anselm's, that require the agent to be appropriately the source of her willings in the sense incompatible with necessitation. So to the extent that Anselm succeeds in explaining the why there is evil by appealing to freedom of choice, he does not succeed in showing how human freedom is compatible with God's grace in making just creaturely agents. Anselm can resolve the problems of harmonization and just creation by giving concessions to the Pelagian. I have not argued that he should not make such concessions; I only noted that his theological tradition rejects Pelagianism. One might reject that tradition.

If the foregoing arguments are correct, significant consequences follow for contemporary philosophers of religion who fall within Anselm's theological tradition. Some of the most influential Christian philosophers of religion appeal to incompatibilist accounts of freedom in order to explain the presence of evil in the world.⁷² For such philosophers, I believe they face a choice similar to Anselm. They can retain incompatibilism but also endorse Pelagianism in some respects in order to solve the problems of harmonization and just creation. Alternatively, they can endorse compatibilism about God's causal activity and human freedom and responsibility, but find another way to explain the presence of evil. In either case, they must make a significant change.

REFERENCES

- Anselm. 1998. *Anselm of Canterbury: The Major Works* (ed. Brian Davis and G. R. Evans). Oxford: Oxford University Press.
- Anselm. 2002. *Three Philosophical Dialogues: On Truth, On Freedom of Choice, On the Fall of the Devil* (ed. Thomas Williams). Indianapolis: Hackett.
- Augustine. 1998. *The City of God against the Pagans* (ed. R. W. Dyson). New York: Cambridge University Press.
- Bailey, Andrew M. 2012. "Incompatibilism and the Past." *Philosophy and Phenomenological Research* 85 (2): 351–376.

⁷²From the Protestant, Eastern Orthodox and Catholic traditions, respectively, see Plantinga (1974), Swinburne (2004), van Inwagen (2006). Though these philosophers do not believe that the appeal to freedom resolves all puzzles raised by the extent and different kinds of evil, they believe that it does some important work in explaining why there is moral evil.

- Brueckner, Anthony. 2008. "Retooling the Consequence Argument." *Analysis* 68 (297): 10–13.
- Burnell, Peter. 1999. "Concupiscence." pp. . 223–227. Grand Rapids, MI: Eerdmans.
- Campbell, Joseph Keim. 2005. "Compatibilist Alternatives." *Canadian Journal of Philosophy* 35 (3): 387–406.
- Campbell, Joseph Keim. 2007. "Free Will and the Necessity of the Past." *Analysis* 67 (294): 105–111.
- Campbell, Joseph Keim. 2008. "Reply to Brueckner." *Analysis* 68 (299): 264–269.
- Cross, Richard. 1999. *Duns Scotus*. New York: Oxford University Press.
- Cross, Richard. 2005. "Anti-Pelagianism and the Resitability of Grace." *Faith and Philosophy* 22 (2): 199–210.
- Davies, Brian and Brian Leftow. 2004. *The Cambridge Companion to Anselm*. New York: Cambridge University Press.
- Hopkins, Jasper. 1972. *A Companion to the Study of St. Anselm*. Minneapolis: University of Minnesota Press.
- Kane, G. Stanley. 1973. "Anselm's Definition of Freedom." *Religious Studies* 9: 297–306.
- Kane, G. Stanley. 1981. "Anselm's Doctrine of Freedom and the Will." In *Texts and Studies in Religion*. Vol. 10. New York: The Edwin Mellen Press.
- Leyser, Conrad. 1999. "Semi-Pelagianism." pp. . 762–766. Grand Rapids, MI: Eerdmans.
- Plantinga, Alvin. 1974. *The Nature of Necessity*. Oxford: Clarendon Press.
- Ragland, C. P. Scott. 2006. "The Trouble with Quiescence: Stump on Grace and Freedom." *Philosophia Christi* 8(2): 343–362.
- Rigby, Paul. 1999. "Original Sin." pp. . 607–614. Grand Rapids, MI: Eerdmans.
- Rogers, Katherine. 2008. *Anselm on Freedom*. Oxford: Oxford University Press.
- Sadler, Gregory. 2008. "Freedom, Inclinations of the Will, and Virtue in Anselm's Moral Theory." *Proceedings and Addresses of the American Philosophical Association* 81: 91–108.
- Schaff, Philip (ed). 2002. *History of the Christian Church*. Vol. 3. Peabody, Mass: Hendrickson Publishers.
- Schaff, Philip (ed). 2007. *Nicene and Post-Nicene Fathers*. Vol. 5 of *St. Augustine: Anti-Pelagian Writings*. New York: Cosimo, Inc.
- Stump, Eleanore. 2003. *Aquinas*. London: Routledge.
- Stump, Eleonore. 2001. "Augustine on Free Will." pp. . 124–147. Cambridge: Cambridge University Press.
- Swinburne, Richard. 2004. *The Existence of God*. New York: Oxford University Press. 2nd edn.
- Sylwanowicz, Michael. 1996. *Contingent Causality and the Foundations of Duns Scotus' Metaphysics*. New York: Brill.

- TeSelle, Eugene. 1999. "Pelagius, Pelagianism." pp. . 633–640. Grand Rapids, MI: Eerdmans.
- Timpe, Kevin. 2007. "Grace and Controlling What We Do Not Cause." *Faith and Philosophy* 24 (3): 284–299.
- Tyvoll, Stan. 2004. "An Anselmian-Quiescence Approach to the Problem of Grace and Merit." *Budhi* 8: 35–58.
- Tyvoll, Stan. 2006. "Anselm's Definition of Free Will: A Hierarchical Interpretation." *American Catholic Philosophical Quarterly* 80 (2): 155–171.
- van Asselt, William, J. Martin Bac, and Roelf T. te Velde (eds). 2010. *Reformed Thought on Freedom: The Concept of Free Choice in Early Modern Reformed Theology*. Grand Rapids, MI: Baker Academic.
- van Inwagen, Peter. 2006. *The Problem of Evil*. Oxford: Clarendon Press.
- Williams, Thomas. 2009. "Review of Katherin Rogers, *Anselm on Freedom*." In *Notre Dame Philosophical Reviews*. [Http://ndpr.nd.edu/news/23911/?id=15226](http://ndpr.nd.edu/news/23911/?id=15226).
- Williams, Thomas and Sandra Visser. 2001. "Anselm's Account of Freedom." *Canadian Journal of Philosophy* 31 (2): 221–244.

GOD AND MORAL PERFECTION

Shawn Graves

Abstract. We will be hard-pressed to find a morally perfect agent in this world. It's not that there aren't any morally good people. It just takes a lot to be morally perfect. However, theists claim that God is morally perfect. (Atheists claim that if God exists, God is morally perfect.) Perhaps they are mistaken. In this paper, I present an argument for the conclusion that God is not morally perfect. The argument depends upon two things: (1) the nature of the concept of moral perfection, and (2) the modest theistic claim that God is involved in the affairs of the world.

I. INTRODUCTION

We will be hard-pressed to find a morally perfect agent in this world. This is not just a function of the times we live in, either. We would be equally hard-pressed to find a morally perfect agent at any time, past, present, or future. It is not that there haven't been, or aren't, or won't be, any morally good people. It's just that it takes a lot to be morally perfect.

However, theists claim that God is morally perfect. (And atheists claim that if God exists, God is morally perfect.) A look at the literature makes this quite clear.¹ Of course, it is true that theists differ over the details. Some theists

¹According to Michael Peterson, et al., "It is a matter of consensus among theists that God is *perfectly morally good*." (See Michael Peterson et al. (1998, p. 73).) C. Stephen Evans claims that, according to theism, "God is morally perfect; his goodness is unsurpassable." (See Evans (1985, p. 34).) Paul Draper takes it as a necessary condition for a thing's bearing the title 'God' that that thing be a "morally perfect person." (See Draper (1989, pp. 331–350).) Paul Moser follows suit, claiming that "the title *God*, on the proposed usage, signifies a being worthy of worship. . . ." He continues, "Let's say that a being is worthy of worship if and only if that being, having inherent moral perfection, merits worship as unqualified adoration, love, trust and obedience." (See Moser (2009, p. 51).) According to Peter van Inwagen, theism is the proposition that "the universe was created by an omniscient, omnipotent, and morally perfect being." (See van Inwagen (1996, p. 69).) Charles Taliaferro notes both that "Christians imagine God to be completely good, morally perfect and supreme. . ." and that "God is conceived of as limitless in power and knowledge, the principal Creator of all, morally perfect, and unsurpassable in all perfection." (See Taliaferro (2002, p. 64, p. 66).) Erik J. Wielenberg claims that "the standard way" for understanding the term 'theism' is as follows: "It indicates the doctrine that there exists a unique omnipotent, omniscient, morally perfect, necessarily existing creator of the universe." (See Wielenberg (2009, p. 23).) While others use different terminology to describe God's moral status, such as 'wholly good' (see Alvin Plantinga (1988)), 'omnibenevolent' (see Joshua Hoffman and Gary Rosenkrantz (2002)), 'supremely good' (see William Rowe (1979) and Oliver Crisp (2010), or 'perfectly good' (see Eleonore Stump (1993),

make a stronger claim than that if God exists, God just happens, as a matter of contingent fact, to be morally perfect. Rather, they think that if God exists, God is *essentially* morally perfect such that there is no possible world, no other way the world might have been, in which God fails to be morally perfect. Other theists deny this. A number of theists endorse a version of theological voluntarism. Some of these theists think that God is the very source and foundation of morality such that moral facts supervene upon, or are determined by, God's moral nature, will, or commands. Of course, other theists deny that this is so. Regardless of disputes over such details, though, it is accepted among theists that God is morally perfect.

But perhaps theists are mistaken. Perhaps it is false that God is morally perfect.² In this paper, I present an argument for the conclusion that God is *not* morally perfect. The argument against God's moral perfection does not contain a premise that states that God has committed some morally repugnant acts.³ Nor is the argument a statement of the problem of evil. Rather, the argument depends upon two things: (1) the nature of the concept of moral perfection, and (2) the modest theistic claim that God is involved in the affairs of the world.

A few words about the project undertaken here are in order. First, one should

Daniel Howard-Snyder (1996), and William Rowe (2002)), it is clear that philosophers take it that if God exists, God is morally perfect. The references thus far have been to philosophers. But of course theologians also claim that God is morally perfect. Here are two quick examples. Charles Hodge writes, "Holiness, on the one hand, implies entire freedom from moral evil; and, upon the other, absolute moral perfection." (See Hodge (2001, p. 43).) J. I. Packer writes, "When the biblical writers call God *good*, they are thinking in general of all those moral qualities which prompt his people to call him *perfect*. . . . The Bible is constantly ringing the changes on the theme of the moral perfection of God, as declared in his own words and verified in the experience of his people." (See Packer (1993, p. 161).

²An anonymous referee raised the following worry. Some will hear 'God is not morally perfect' as incoherent, on the order of 'some bachelors are married' or 'not all squares are four-sided.' Consequently, why even entertain the question as a sensible one? Here's my response. Well, perhaps it is, in the end, incoherent. Perhaps (as some Anselmian descriptivists would have us think) it is a conceptual truth that God is morally perfect. But (at the very least) that's opaque. It's just not plainly obvious that it is incoherent, even on an Anselmian descriptivist way of thinking about 'God'.

But even adopting an Anselmian descriptivist view need not entail that 'God is morally perfect' expresses a conceptual truth. While, on this view, it is a conceptual truth that God is the greatest possible being, claims about which properties God has need not be taken as conceptual truths (or falsehoods). So even if one were to take the view that 'God is the greatest possible being' expresses a conceptual truth, one need not take the claim 'God is not morally perfect' as expressing the denial of any conceptual truth, resulting in incoherence. (So Anselmian descriptivists may take debates over God's attributes as debates over non-conceptual claims.)

Of course, everything that has been said here in this note has been assuming that we are taking an Anselmian descriptivist view of 'God'. But we need not do that, either. We can take 'God' in 'God is not morally perfect' as a kind of Kripkean proper name picking out the being (actual or not) that looms large in the Abraham narratives. In that case, there is certainly nothing incoherent in claiming that God is not morally perfect.

³Many, many works could be cited here. For example, see Wes Morriston (2009), Elizabeth Anderson (2007), David Lewis (2007), and Richard Fumerton (2001).

not think that I am engaging in a mere terminological dispute. This project is not a matter of disputing over whether or not ‘morally perfect’ is the best phrase to use or whether or not it is a pious thing to say about God.⁴ Rather, this project is a matter of which attributes or properties it is correct to ascribe to God. So, it is a matter of whether it is correct to ascribe the property *being morally perfect* to God. Since I think that theists find it worthwhile to get the attributes of God right, I think that theists will find the issue taken up in this paper worthwhile as well.

Second, and in the same vein, theists may take this project as a friendly attempt to clarify how theists should conceive of God’s moral attributes. Similar projects have been conducted regarding others of God’s attributes, such as omnipotence and omniscience. It is true that most of these other projects do *not* conclude that God fails to have the property under discussion, e.g., that God is not omniscient.⁵ The argument offered here does conclude that God does not have the property *being morally perfect*. Therefore, this project and most other projects on God’s attributes are dissimilar in that respect. But, just as discussions of, say, omnipotence and omniscience are attempts to get clear about God’s attributes with respect to power and knowledge, so is this project an attempt to get clear about God’s attributes with respect to morality.

Finally, a word about methodology. Pretty clearly, if moral perfection is identical to the extent of God’s moral goodness, then any argument to the conclusion that God is not morally perfect is doomed. Obviously, then, in this paper I don’t assume that moral perfection is identical to the extent of God’s moral goodness. (Maybe it will turn out that it *is* identical, but I don’t assume it up front.)⁶ With all of this in mind, let’s turn now to the task of explaining moral perfection.

II. GETTING CLEAR ABOUT MORAL PERFECTION

As Ed Wierenga notes, when we think about being morally perfect, “we have some intuitions about what such goodness consists in, and it clearly involves

⁴This worry was expressed by a well-known Christian philosopher in the audience at my talk at the 2004 Eastern Regional Meeting of the Society of Christian Philosophers.

⁵One notable exception: Peter Geach concludes that God is *not* omnipotent in his discussion of that attribute. Geach opts for *almightiness*, instead. See Geach (2002).

⁶Compare Earl Conee (1991) on omnipotence: “Some philosophers of religion have developed accounts of omnipotence guided by the idea that omnipotence is identical to the extent of God’s power. Often these accounts are interesting and impressive. But this guiding assumption is dangerous, particularly when it is coupled with the assumption that God exists. The principle danger is that omnipotence will be misconstrued in order to insure that it is compatible with God’s other assumed attributes. Such an analysis may explain how powerful it is possible for God to be, instead of explaining omnipotence. This risk would be worth taking if we had no other initial conception of omnipotence except as the extent of God’s power. But we do have an independent conception of omnipotence. . . .” It seems clear, then, that Conee is open conceptually to the possibility of the existence of a less-than-omnipotent God. Peter Geach (2002) was also open to this. Not all are so open. Consider Peter van Inwagen: “I take the phrases “a less than omnipotent God” and “a God who sometimes does wrong” to be self-contradictory, like “a round square” or “a perfectly transparent object that casts a shadow.” See van Inwagen (2005, p. 188).

doing no wrong.”⁷ According to Richard Swinburne, “in claiming that God is morally perfectly good, the theist means that God is so constituted that he never does actions which are morally wrong.”⁸ Moreover, Peter van Inwagen claims: “To say that God is morally perfect is to say that he never does anything morally wrong—that he could not possibly do anything morally wrong.”⁹ So, following Swinburne and van Inwagen, one might think that the following is a good analysis of moral perfection.

MP1 A subject *S* is morally perfect if and only if *S* never does anything morally wrong.¹⁰

It is surely true that a morally perfect subject never does anything morally wrong. But MP1 cannot be correct. Newborn infants are subjects that have not done anything morally wrong, but it is a mistake to say that they are morally perfect. While newborn infants have done nothing morally wrong, they have done nothing morally right, either. From the moral point of view, newborn babies have not done anything at all. So, perhaps adding a condition to the analysis to fix this defect will give us an adequate analysis of moral perfection. Consider the following analysis from Colin McGinn.

MP2 A subject *S* is morally perfect if and only if (i) *S* always does what is morally right and (ii) *S* never does what is morally wrong.¹¹

MP2 has some intuitive plausibility.¹² A subject whose actions are always morally right and never morally wrong is doing extremely well from the moral point of view. It would surely be an amazing feat for the likes of us.

But doing everything right morally and nothing wrong morally still falls short of moral perfection. Here’s why. In order for something to be perfect in a given area, it must be that there can be no improvements made in that area. But it is possible for one always to do what is morally right and never do what is morally wrong, and still be improved from the moral point of view.

To see this, consider the following. Suppose some moral subject, call him ‘Okay’, always performs morally right acts and never performs a morally wrong act. In addition, Okay is of comparatively upstanding moral character. However, he does have some character flaws. Every now and then he is impatient, cowardly, prideful, or unforgiving. Consequently, Okay will sometimes act rightly, but from

⁷See Wierenga (1989, p. 203).

⁸See Swinburne (1977, p. 179).

⁹See van Inwagen (2005, p. 191).

¹⁰MP1 is what Erik Wielenberg has called ‘impeccability’. He claims: “*x* is impeccable = *σf*. *x* never performs a morally wrong action.” See Wielenberg (2004, p. 44). It seems that a straightforward consequence of Wielenberg’s view is that aardvarks, bats, chickens, dogs, even food, grass, and hills count as morally impeccable. I wouldn’t have thought that those things had *that* moral status.

¹¹See McGinn (1992, pp. 32–34). Michael Clark seems to endorse MP2 (or something very close to it) in his 1993. A couple things: First, ‘morally right’ here (and in the remainder of the paper) means the same thing as “morally permissible”. Second, condition (ii) seems redundant given condition (i), yet I leave it here (and in subsequent analyses) because (1) it is explicitly the analysis Colin McGinn offers and (2) it makes explicit the progression from MP1 to MP2 (and to other analyses).

¹²See McGinn (1992).

a less than commendable moral character. For example, some of his morally right acts are performed impatiently.

Okay satisfies the conditions in MP2. He is always doing what is morally right and never doing anything morally wrong.¹³ But it is also clear that Okay can be improved upon from the moral point of view. Okay has some character flaws. His intentions and motivations are *not* thoroughly virtuous. Surely Okay is not a morally perfect subject, even though he satisfies MP2. MP2 is false.

Though the conditions specified in MP2 are insufficient for moral perfection, it is very plausible to suppose that they are necessary. The problem was that Okay lacked a thoroughly virtuous and upright character. Perhaps adding a condition to MP2 to correct for this particular moral deficit would yield a satisfactory analysis. Consider MP3.

MP3 A subject S is morally perfect if and only if (i) S always does what is morally right, (ii) S never does what is morally wrong, and (iii) *S always acts from thoroughly virtuous motives and intentions.*

MP3 has considerable intuitive plausibility.¹⁴ MP3 considers both the acts performed by the agent and the internal states of the agent. As a result of taking into account the internal states of the agent, MP3 avoids cases like Okay's. Surely one who meets the conditions specified in MP3 is a moral model and worthy of high esteem and admiration from the moral point of view. However, one would *not* thereby be morally perfect. I suggest that satisfying the conditions in MP3 is still *insufficient* for moral perfection. The problem with MP3 is with condition (i). Let's consider why.¹⁵

It seems possible that, among a set of alternative actions for an agent at a time, there are multiple morally right actions available to that agent. Among

¹³Virtue theorists would disagree here. Since virtue theorists place internal psychological constraints on right actions such that the right action is always done from virtuous internal psychological states, they would claim that Okay's occasional morally vicious mental states prevents some of his actions from being morally right. While there is some plausibility to this notion, it seems mistaken. It seems that one can do the morally right thing for morally bad reasons or from a morally flawed psychological state. So, one might do the morally right thing by telling the truth in a situation, but still be morally blameworthy because one did so out of pride or cruelty. Regardless, those who remain committed virtue theorists and read MP2 in that way may regard the criticisms of MP3 that follow as criticisms of a virtue theorist's reading of MP2.

¹⁴Something like MP3 seems to be what L. L. Garcia has in mind. She writes, "...let us stipulate that the definition of moral perfection is as follows:

(P4) An agent is morally perfect in a given possible world if and only if the agent never does anything morally blameworthy in that world (where 'does' is read in a broad enough sense to include sins of omission as well as sins of commission)."

See Garcia (1987, p. 139). Garcia asserts that "the property of being morally perfect is one which depends upon the agent's behavior" such that "it refers...to an absence...of any immoral actions on the part of the agent" (pp.139-140). This seems like an endorsement of something like MP2. However, it seems to me that MP3 better captures the idea of being free from moral blame. Regardless, Garcia's definition is false (though she is free to stipulate as she pleases). Her proposed condition is insufficient for moral perfection. The case of God that follows shows this.

¹⁵In what follows I am indebted to Conee (1994, pp. 815-816).

the alternative actions that are right, some are morally better than others.¹⁶ These morally better alternative actions are sometimes called *supererogatory acts*. Given that there are supererogatory acts, it seems possible that one could satisfy the conditions of MP3 yet still make moral improvements.

Consider a moral agent, call her ‘Good’, who is a stunning example of virtuous character and is sure to act in accordance with that virtuous character. Good satisfies conditions (i)-(iii) of MP3. Importantly, though, Good does not always take the best alternative from among her options. Sometimes an alternative requires much more effort and sacrifice than another alternative, which still requires significant effort and sacrifice, and the difference is that the alternative requiring much more effort and sacrifice produces only very slightly more good overall. In those cases, Good takes the option with far less, though not insignificant, effort and sacrifice.

As a result, Good is not morally perfect. Improvements can be made in her moral behavior. She could perform the supererogatory acts. Since Good’s actions are not the morally better acts of supererogation, it follows that Good is not a morally perfect person.¹⁷ MP3 says otherwise. Therefore, MP3 is false.

But here’s an objection.¹⁸ By hypothesis, Good satisfies the conditions specified in MP3. Consequently, by hypothesis, Good always acts from thoroughly virtuous motives and intentions. But if that is so, then we have no reason to think that Good fails to perform supererogatory acts. This is because acting from *thoroughly* virtuous motives and intentions entails that one does more than what duty requires. That is, acting from *thoroughly* virtuous motives and intentions entails performing acts of supererogation. But if all of this is correct, then it follows that the above (alleged) counterexample where Good satisfies MP3’s conditions yet sometimes fails to perform supererogatory acts is impossible. So this objection to MP3 fails.

There is a plausible reply to this objection. Quite simply, acting from thoroughly virtuous motives and intentions does *not* entail performing acts of supererogation. It is possible that one act from thoroughly virtuous motives and intentions yet fail to perform a supererogatory act. Here’s just one reason why. One might not be aware of, or consider, all of one’s genuine alternatives at a given time and so fail to be aware of, or consider, any of one’s supererogatory alternatives at that time. This need not be due to some culpable negligence. It need not be due to some vicious internal state, motive, intention, or disposition.

¹⁶Some consequentialists would deny this. These consequentialists endorse a maximizing form of consequentialism such that only the alternative (or set of alternatives) that maximizes the intrinsic good in the world is the action that is morally right. All other non-maximizing alternative actions are morally wrong. On this view, it is not true that there are some alternative actions that are morally inferior to other alternatives yet are still morally right actions. Those who are committed maximizers and read MP3 that way should regard the criticisms of MP4 that follow as criticisms of a maximizing consequentialist reading of MP3.

¹⁷Both Roy Sorenson (2007) and Susan Wolf (1982) seem to agree that failing to perform available acts of supererogation disqualifies one from being morally perfect.

¹⁸Thanks to an anonymous referee for raising this objection.

Plausibly, one's supererogatory alternatives simply did not occur to one. Therefore, one might very well fail to perform any of those actions. And so one who is acting from thoroughly virtuous motives and intentions might fail to perform a supererogatory act in some situation. This failure is not the result of some character problem. Rather, one might suffer from a kind of *epistemic* problem. This is sufficient to rebut the objection.

So, acts of supererogation cause problems for MP3. A morally perfect agent will not settle for what is morally right. A morally perfect agent will always perform acts of supererogation.¹⁹

It would be a mistake, though, to think that some simple revision of the analysis with a reference to supererogatory acts will take care of the issue. This is because some supererogatory acts bring about more good than other supererogatory acts, and so are better from the moral point of view. Perhaps an example will help here.²⁰ Suppose Better is a mailman on his route. While on his route he happens upon a burning building. The fire seems especially fierce. A woman is out front screaming that her small children are trapped inside. There are no rescue workers anywhere, and apparently no one else is around. Better drops his bag and rushes into the burning building and comes out with a small baby in his arms. Better rescues the baby from the fire.

Clearly Better has done something heroic in rescuing the baby. He has performed a supererogatory act. And while this is surely laudable, it remains possible to improve the act morally. For suppose Better rushed into the burning building and came out with the small baby and another one of the woman's children in his arms. In this alternative version, Better brought about more good than in the prior version of the case. The second version constitutes a moral improvement over the first version, even though both acts are supererogatory. A revision to MP3 will need to take this into account. Consider MP4.

MP4 A subject S is morally perfect if and only if (i) S always does what is *maximally morally good*, (ii) S never does what is morally wrong, and (iii) S always acts from thoroughly virtuous motives and intentions.

The revised condition (i) of MP4 does not, as MP3 does, yield the verdict that Good is morally perfect. In addition, it takes into account that even among supererogatory acts, there are morally better and worse acts.

It is surely true that MP4 specifies incredibly exacting standards for a moral agent.²¹ To satisfy conditions (i)-(iii) of MP4 is something that very few, if any,

¹⁹Oliver Crisp seems to neglect this point in his 2010. For example, in arguing that God could create a world where no human person is saved, and that such a world is consistent with God's goodness, he writes, "God is not obligated to save any human sinner who exists in the world he does create, provided the sinner has freely sinned against God—for salvation is a matter of divine grace, which is gratuitous." On the supposition that God is morally perfect, *merely* acting in compliance with moral obligations is not consistent with God's perfect goodness. To be consistent with moral perfection, God must do more. God must go beyond mere obligations.

²⁰The case that follows is based upon a case presented by Fred Feldman (1978, p. 48).

²¹Indeed, some might say *too* exacting. Some might object that satisfying condition (i) of MP4 is not necessary for moral perfection since it seems possible that in some cases there

people have done. I'm extremely confident that even St. Francis of Assisi, Mother Theresa, and Gandhi—individuals that Robert Adams identifies as actual moral saints—failed to achieve such a status.²² One who did satisfy such conditions would be worthy of great admiration, extensive praise, and tremendous respect and honor from the moral point of view. Unfortunately, it would *not* make one a morally perfect agent. Let's consider why this is so.²³

It is plausible to suppose that moral agents sometimes find themselves in situations in which it is necessary to do a regrettable thing in order to do what is morally right. The easiest cases are those involving agents that bring about or knowingly and intentionally allow pain and suffering. Consider the following sort of case.²⁴

Suppose some highly regarded pediatrician, call him 'Great', has a new patient that is suffering from a deadly disease. Great knows that there is only one drug currently available that can cure the ailing boy. Unfortunately, Great also knows that the drug will produce tremendous pain and suffering for the boy for approximately three weeks. Great is sad that the drug has such a side effect, but knows that there is no other way. He administers the drug and the boy, after three weeks of agony, is cured. The boy lives a long and happy life.

It seems correct to say that Great did the morally right thing in this case by administering the drug and saving the boy's life. However, Great did the morally right thing by knowingly bringing about the boy's tremendous pain and suffering. That seems unfortunate and regrettable, even from the moral point of view.

Now consider what we might think of as Great's *moral résumé*, that is, a comprehensive, exceedingly detailed account of every morally relevant feature about Great, concerning both his actions and his internal states, throughout his life. Among the things listed on this moral résumé will be Great's performance of an act that brought about the boy's miserable pain and suffering. Let us

simply is no alternative available to the agent that is maximally morally good. That is, it seems possible that in some cases there simply is no best act because one is confronted with an infinite array of better and better options from the moral point of view. Were an agent confronted with such an array of options, the objection continues, the mere fact that the agent performed an act that fails to be maximally morally good, or best, would not count as a strike against that agent's candidacy for moral perfection. For an objection along these very lines, see the Howard-Snyders' Jove case and subsequent commentary found in their 1994. See also Thomas Morris (1993) and William Hasker (2008). There is also the vast and varied literature spawned by Robert Adams' oft-reprinted 1972. Now, I think that there are some plausible things to say in reply to this sort of worry. But due to space constraints I simply can't get into all that here. Furthermore, it's important to note that the fate of the central argument presented in this paper—that God is not morally perfect—does *not* depend on whether condition (i) of MP4 is necessary for moral perfection.

²²See Adams (1984, p. 392).

²³In what follows I am indebted to Conee (1994, pp. 816–817).

²⁴The case and commentary that follows focuses upon bringing about pain and suffering. It is easy enough to revise the case and commentary so that it focuses upon knowingly and intentionally allowing pain and suffering. Nothing of substance changes.

suppose that Great has always done what is maximally morally good, never done what is morally wrong, and has always acted from morally virtuous motives and intentions. His moral résumé reflects these facts about him. Nonetheless, it seems clear that at least one moral improvement could be made to his moral résumé: it could have been the case that Great never once performed an act that brought about any pain and suffering.

To see that this would be an improvement, imagine an otherwise identical moral résumé placed next to his, with the difference being that on this résumé there are no reports of acts that brought about any pain and suffering. There is nothing in this moral résumé at all to regret, nothing whatsoever that requires any moral justification or excuse of any sort. This résumé is spotless. From the moral point of view, it seems correct to say that this résumé is better than Great's. Great's résumé is not spotless. Since that résumé is better than Great's from the moral point of view, and nothing can be better than perfection, it follows that Great's moral résumé is *not* perfect. Thus, even though Great has satisfied conditions (i)-(iii) of MP4, Great fails to be morally perfect. The conditions of MP4 are *insufficient* for moral perfection.

At this point we need to consider a potential objection.²⁵ Take an otherwise identical moral résumé. Call the owner of that résumé 'Bill'. That otherwise identical résumé includes the circumstances of action that Great faced. Did Bill save the boy's life by administering the drug? If so, then Bill's résumé is no different than Great's résumé, so not better than Great's résumé, and so there's no basis for claiming that Great is not morally perfect. On the other hand, if Bill faced those same circumstances as Great, and Bill did not save the boy by administering the drug, then Bill did what was wrong. In that case, Bill's résumé is *worse* than Great's, and so there's no basis for claiming that Great is not morally perfect. Either way, Bill's résumé is not better than Great's résumé, and so either way there's no basis for claiming that Great is not morally perfect.

I think there's a decent reply to this objection. The objection goes wrong in assuming that an otherwise identical résumé includes the same *need-to-administer-the-drug-to-save-a-boy's-life* circumstances as Great. Think about it this way: the otherwise identical résumé is such that it represents the résumé Great *would* have in the closest possible world where the following obtains: (1) no pain and suffering is brought about, and (2) just as much good is brought about as is found on Great's résumé. And that *is* a morally superior résumé to Great's résumé. The objection fails.

Now, it is important to be clear about the moral evaluation being made here in this case of Great. I am claiming that Great fails to be morally perfect because he performed an act that brought about an intrinsically bad state, in this case some pain and suffering. The claim is *not* that Great has done something morally wrong by administering the drug. Nor am I claiming that Great is to be reprimanded, scolded, or otherwise rebuked for administering the drug. On the

²⁵Thanks to an anonymous referee for raising this objection.

contrary, it seems that Great is to be thoroughly praised for doing what he did. He did the best that he could given the regrettable circumstances that he found himself in through no fault of his own.

Some might find this latter consideration to be some reason for resisting the charge that Great fails to be morally perfect. That Great found himself in these regrettable circumstances is not something that he could control. It was simply a matter of bad luck, one might say, that Great found himself in such regrettable circumstances that demanded he bring about some pain and suffering. It's not as though Great was negligent or otherwise culpable for the unfortunate state of affairs coming about. Consequently, one might reason, it is no strike against Great's moral résumé that he brought about some pain and suffering. The trouble, one might put the point, is to be located in the circumstances, not in the agent. So, the objection continues, we have been too hasty in charging Great with moral imperfection.²⁶

There's reason to find this objection unconvincing. It will be instructive to see why the objection fails. I grant the premises in the objection. Nevertheless, it remains true that Great is morally imperfect. Note that all of the premises found in the objection are consistent with the above moral résumé argument. All of those moral claims made about Great in the objection are listed in Great's moral résumé. Nevertheless, it still seems that an otherwise identical moral résumé containing no such report of an act that brought about pain and suffering is better than Great's résumé. We would prefer to have *that* résumé over Great's résumé, presumably because we think that that résumé is *better* than Great's. There is nothing morally regrettable about *that* résumé; there is something morally regrettable about Great's résumé. All other things being equal, we would prefer to never do anything regrettable or that stands in need of justification of any sort.

Perhaps an extended example from epistemology would be helpful here. Suppose Brain is an exemplary epistemic agent. So, for example, for any proposition that Brain considers, he takes the justified doxastic attitude toward that proposition. That is, when belief is called for, he believes, when disbelief is called for, he disbelieves, and when suspension of judgment is called for, he suspends judgment. Not only does Brain always take the justified doxastic attitude, but he always does so for just the right reasons. We might say that all of his beliefs are completely *well-founded* or, if you prefer, *well-formed*. Moreover, Brain always takes the justified doxastic attitude in a thoroughly epistemically virtuous way. Brain always engages in thoroughly epistemically responsible inquiry. Whatever else is of positive epistemic value or counts as cognitive success, Brain maximizes it or experiences it. Brain is a top-notch epistemic agent.

Now suppose Brain finds himself, through no fault of his own, in a set of circumstances where there is a kind of unavoidable epistemic value conflict. There are lots of ways this could be. For example, suppose in these circumstances belief is the epistemically justified attitude to take toward some proposition *p*. Further

²⁶Thanks to William Hasker and David Basinger for pressing this objection.

suppose that justifiably believing that p will lead to a lot of other epistemic goods, say, the acquisition of a lot of true beliefs. However, the circumstances are also such that, unavoidably, justifiably believing p (i.e., acquiring an epistemic *good*) will bring about something else that is otherwise epistemically *bad*, say, the acquisition of a few false beliefs that q , r , and s .²⁷ Now suppose Brain does the expected and justifiably believes p . That's epistemically optimal in the circumstances. But justifiably believing p results in the false beliefs that q , r , and s . That's epistemically regrettable.

What are we to make of this? Well, we certainly wouldn't *blame* Brain in this case. After all, he is believing what he ought to be believing, his belief is completely well-founded and thoroughly epistemically virtuous, and he is doing the best he can do from the epistemic point of view. In addition, it is true that he is not culpable for being in the circumstances he is in. Nevertheless, it seems clear that there is an epistemic imperfection here. If we were to compare Brain's epistemic résumé to an otherwise identical epistemic résumé that did not have this feature (i.e., the acquisition of a few false beliefs), we would surely prefer that other résumé to Brain's. That other résumé is better than Brain's résumé, from the epistemic point of view. Quite simply, Brain's résumé is not spotless; it can be improved upon.

There is a lesson to be learned here. Circumstances can ruin perfection. Perfection requires the world to cooperate. What circumstances we find ourselves in can be a matter of luck. That seems true in the epistemic realm. The moral realm is no different. Brain, our epistemic subject, fails to be epistemically perfect as a result of the circumstances that he finds himself in through no fault of his own. Our moral subject, Great, is morally imperfect as a result of the circumstances that he finds himself in through no fault of his own. Since MP4 says otherwise, MP4 is mistaken.

There's a final objection we should consider here.²⁸ This objection concedes that Great's moral résumé is flawed, but resists the claim that Great is thereby morally imperfect. That is, the above reasoning involving Great seems to rely crucially upon the following principle:

Perfect résumé Principle (PRP). If an agent A's moral résumé is not perfect, then A herself is not morally perfect.

But perhaps that principle is false. Perhaps, the objector continues, we can consistently grant that Great's *moral résumé* is imperfect, but insist that Great *himself* remains morally perfect. If that's so, then it looks like the above reasoning about Great is no good.

²⁷Tell the story however you wish. Here's one convenient way to do it. Suppose a powerful angel has vowed to reward Brain should Brain form a justified doxastic attitude toward proposition p . This reward will consist in the bringing about of a lot of true beliefs for Brain. Yet further suppose that a powerful demon has also made a vow. He has vowed to deceive Brain over the truth of propositions q , r , and s (through powerful perceptual illusions) should Brain form a justified doxastic attitude toward p . Now, as expected, Brain entertains p and promptly forms the justified belief that p . Both the angel and the demon keep their vows.

²⁸Thanks to an anonymous referee for raising this objection.

There's a decent reply to this objection. It seems plausible to suppose that morally perfect agents have nothing short of perfect *résumés*. Their records are squeaky clean, utterly spotless from the moral point of view. There's nothing to regret, nothing that requires justification, nothing that requires excuse. There's no record of violating a prima facie moral duty, no record of harms brought about, nothing of that sort at all. Morally perfect agents have *résumés* that represent the moral ideal. Denying (PRP) goes against these claims, permitting morally perfect agents to have spotty *résumés*, with records of actions that violate prima facie moral duties, bring about harm, require justification or excuse, and are regrettable. That seems objectionable.

Morally perfect agents are marked by a number of things. One mark of a morally perfect agent is this: morally perfect agents only perform actions that are *perfectly okay* from the moral point of view. But acts that bring about harms, violate prima facie duties, require justification or excuse, or are regrettable are *not* perfectly okay from the moral point of view. There are genuine moral considerations against committing such acts. Moral *résumés* reflect whether such acts have been committed. If they have been committed, then the *résumé* is not perfect. So an imperfect *résumé* entails an imperfect moral agent.

Finally, I suspect that this objection stems from the conviction that moral perfection really amounts to "proper responsiveness to value", to "responding to the opportunities to promote and respect what is good" regardless of what's going on outside the agent.²⁹ So, even if one's moral *résumé* records actions that violate prima facie duties, bring about harms, demand justification or excuse, or are regrettable, it doesn't follow that the agent has failed to respond properly to value or respond to the opportunities to promote and respect what is good (as it was put above). The agent may have responded in superior fashion, but the spots on the *résumé* were unavoidable given the circumstances.

In reply, it's worth pointing out that there may be *other* concepts besides moral perfection that get at an agent's properly responding to value or responding to the opportunities to promote and respect what is good regardless of what's going on outside the agent. Consider *perfect righteousness*, *perfect uprightness*, *perfect moral rectitude*, or *perfect holiness*. Perhaps those concepts pick out agents that respond the best they can (from the moral point of view) in the best way they can (from the moral point of view) given the circumstances. But none of those concepts are equivalent to moral perfection. While those concepts consider *some* things of moral relevance (e.g., internal states of the agent), moral perfection considers *everything* of moral relevance. Moral perfection is a summary notion. That an act (unavoidably) violates a prima facie duty, brings

²⁹As an anonymous referee put it to me in written comments on a previous draft. Indeed, this referee suggested that this is the notion of goodness "presupposed by those who hold that God is a morally perfect agent". Maybe that's so. I take no stand in this paper on what most people have in mind or presuppose when they ascribe moral perfection to God. However, if this is what some people do have in mind or presuppose when they ascribe "moral perfection" to God, then it seems that they are either misspeaking or are mistaken about moral perfection.

about harm, needs justification or excuse, or is regrettable is morally relevant. None of those things count as perfections. Morally perfect agents are marked only by perfections. So, a morally perfect agent has nothing short of a perfect rsum.

Let's remind ourselves where we're at in this discussion. The claim is that satisfying conditions (i)–(iii) in MP4 is insufficient for being morally perfect. The case of Great seems to show this. Some objections were raised to that conclusion. There's reason to think that the objections fail. MP4 is false.

An obvious fix to make to MP4 in light of the above considerations is to add a condition to MP4 that precludes bringing about or knowingly and intentionally allowing any intrinsically bad state at all. The result is MP5.

MP5 A subject S is morally perfect if and only if (i) S always does what is maximally morally good, (ii) S never does what is morally wrong, (iii) S always acts from thoroughly virtuous motives and intentions, and (iv) *S never brings about or knowingly and intentionally allows any intrinsically bad state.*

I think that with MP5 we may have arrived at a plausible analysis of moral perfection. I put this a bit tentatively because there is some reason to think that even this analysis fails. The reason has to do with the nature of improvement.

As I have already indicated above, one cannot improve what is perfect. Now consider two agents, S and S*, who inhabit two different possible worlds, w and w*, respectively. S and S* are duplicates in that they perform the exact same acts from the exact same morally virtuous motives and intentions. They are internal duplicates from the moral point of view. Indeed, both S and S* satisfy the conditions of MP5. However, w and w* are not so similar. The conditions of w and w* are such that for every action A and its counterpart A* that S and S* perform, respectively, if the amount of intrinsic good produced by A in w measures n total units, then the amount of intrinsic good produced by A* in w* is n+1 total units.

Now suppose we evaluated S and S* from the moral point of view. Suppose we wanted to determine whether S or S* was the better moral agent overall. Given that there are no moral differences internal to S and S*, we have no basis there for deciding the better moral agent between the two. However, when we look at their acts, we see that S*'s acts always produce 1 more unit of moral good than S's acts do. Thus, S*'s acts seem to be of greater moral value.³⁰ They are an improvement over S's acts. If we wished to improve S's acts morally, we would make them like S*'s in moral value. On this basis, we may evaluate S* as the better moral agent.

Supposing that we are inclined to judge S* as the better moral agent, we now have a basis for objecting to MP5. Recall that both S and S* satisfy the

³⁰There are complications here. I have been assuming that S has produced only a *finite* amount of intrinsic good. If n is an *infinite* amount of intrinsic good, though, then S*'s acts having produced n+1 amount of intrinsic good will not result in a *greater* total amount of intrinsic good, since $n + 1 = n$ in that case. For discussion on cases dealing with an infinite amount of intrinsic goods, see Hoffman and Rosenkrantz (2002).

conditions specified in MP5. We are inclined to evaluate S as morally imperfect because S's acts fail to produce the greatest *possible* amount of intrinsic good. This is shown by the possibility of S*. Since this is so, there remains the possibility, demonstrated by S*, of improving S's moral rsum. Therefore, S is not morally perfect.

As a result of this alleged difficulty with MP5, we may seek a better analysis of moral perfection. Consider MP6.

MP6 A subject S is morally perfect if and only if (i) S always does what is maximally morally good, (ii) S never does what is morally wrong, (iii) S always acts from thoroughly virtuous motives and intentions, (iv) S never brings about or knowingly and intentionally allows any intrinsically bad state, and (v) *S's acts always result in the greatest possible amount of intrinsic good such that there is no possible moral agent S* such that S*'s acts result in a greater total amount of intrinsic good than S's acts.*

One who thinks that MP5 does have the sort of problem identified above will presumably adopt MP6 over MP5.

Both MP5 and MP6 have a plausible claim to being the correct analysis of moral perfection. Of course, they can't both be correct. At least one is mistaken. Fortunately, I do not need to say which one is mistaken. As noted above, it does not matter whether one prefers MP5 or MP6. Indeed, so long as one takes condition (iv) of MP5 and MP6 to be *necessary* for moral perfection, it would not even matter to the main argument of this paper if one prefers some further analysis MP7. Since the main argument of this paper asserts that God does not satisfy that necessary condition for moral perfection, there's no need to proceed any further with the task of analysis. Let's proceed now to the argument.

III. THE ARGUMENT AGAINST GOD'S MORAL PERFECTION

The argument against God's moral perfection is a straightforward two premise argument. According to both MP5 and MP6, a necessary condition for a subject S's being morally perfect is that S never brings about or knowingly and intentionally allows any intrinsically bad state. Taking God to be our moral subject S, this condition provides us with our first premise in the argument.

P1 God is morally perfect only if God never brings about or knowingly and intentionally allows any intrinsically bad state.

Given the discussion in the previous section, P1 seems very plausible.

It is not hard to predict what the next premise in the argument will be. It denies that God satisfies the necessary condition specified in P1.

P2 It is not the case that God never brings about or knowingly and intentionally allows any intrinsically bad state.

Equivalently, God has at some time brought about or knowingly and intentionally allowed some intrinsically bad state. To establish the intended conclusion, we now need some reason for thinking that P2 is true. The reason has to do with the modest theistic claim that God is involved in the affairs of the world.

According to theists, God is involved in the affairs of the world. That is, God brings about some states of affairs in the world, knowingly and intentionally

allows others to bring about some states of affairs in the world, and both approves and disapproves of some states of affairs in the world. By being involved in the affairs of the world, God has either performed an act that caused something to experience pain and suffering, or God has knowingly and intentionally allowed others to perform an act that caused something to experience pain and suffering. Therefore, God has brought about or knowingly and intentionally allowed some pain and suffering. Since the states of pain and suffering are intrinsically bad states, it follows that God has brought about or knowingly and intentionally allowed an intrinsically bad state. This is sufficient to establish P2. P2 seems true.

We now have a straightforward argument against God's being morally perfect. Here is the argument.

The Argument Against God's Moral Perfection

P1 God is morally perfect only if God never brings about or knowingly and intentionally allows any intrinsically bad state.

P2 It is not the case that God never brings about or knowingly and intentionally allows any intrinsically bad state.

C Therefore, it is not the case that God is morally perfect.

We have seen that it is quite plausible to think that both P1 and P2 are true. Since the argument is valid, it is quite plausible to think that the conclusion is true. So, theists are now faced with the plausible conclusion that God is *not* morally perfect.

IV. SOME POTENTIAL MISUNDERSTANDINGS OF THE ARGUMENT

There are some ways in which one might misunderstand the argument. Here I present two of these ways and respond to them.

'But God never does anything immoral.'

That's right. It is no part of the premises to claim that God sometimes does something morally wrong. Indeed, it is consistent with the premises that God *cannot* ever do anything morally wrong. That is, one can affirm that there is no possible world such that God does something morally wrong and still consistently accept both premises of the argument. This is because P2 merely claims that God has brought about or knowingly and intentionally allowed an intrinsically bad state, *not* that God has ever done anything morally wrong. Those are different claims. As noted above, it is consistent to claim that God has brought about or knowingly and intentionally allowed such bad states and yet never done anything morally wrong. It is no part of the premises to claim that God has ever done anything morally wrong.

'Isn't the argument really just another statement of the problem of evil?'

Ultimately, my answer is 'no'. But this depends upon what one means by 'the problem of evil'. Since there are a plenty of arguments that purport to qualify as statements of the problem of evil, it is hard to say whether the argument I present above is best regarded as a member of that host.

With that noted, there are still two points to make here. First, insofar as arguments from evil may be taken as arguments purporting to provide reasons for thinking that a morally perfect God does not exist, then the argument offered here at least shares the same purpose as arguments from evil. Of course, most people then infer from the conclusion that a morally perfect God does not exist the further claim that God does not exist. It's not at all clear that such an inference is warranted. So long as some other superior, exalted moral status—one that preserves God's intrinsic worthiness of worship—may be predicated of God, then theists need not conceive of God as being morally perfect.

Second, as I understand the literature on arguments from evil, a central point of contention is whether there is, or even could be, a morally justifying reason that we know of, or could know of, for God's permitting some evils to obtain.³¹ The argument offered here takes absolutely no stand on this issue. The premises of the argument are consistent with either a positive or negative response. This is so because the argument relies merely upon the claim that God brought about or knowingly and intentionally allowed some evils, and that this is a violation of a necessary condition for moral perfection. The argument does not entail that bringing about or knowingly and intentionally allowing such evils is morally unjustifiable, nor does it entail that bringing about or knowingly and intentionally allowing such evils is morally justifiable. It remains neutral on this point. For this reason, I do not take the argument to be a member of the host of arguments that fall under the heading 'the problem of evil'.³²

V. CONCLUSION

In this paper I offered an argument to the conclusion that God is not morally perfect. I began by analyzing the concept of moral perfection. I then gave reasons for thinking that God fails to satisfy one of the necessary conditions of moral perfection. Finally, I identified and responded to two ways that one might misunderstand the argument.

Theists have three alternatives: they may deny one or more premises in the argument, they may accept the argument and abandon their theism, or they may accept the argument, continue to affirm that God exists, but adopt the view that God is not morally perfect, opting instead for some other moral status that preserves worthiness of worship. Presumably, theists who wish to avoid the conclusion that God is not morally perfect would reject the premise of the argument that follows directly from the analytical work on the concept of moral

³¹See Howard-Snyder (1996).

³²Incidentally, Peter van Inwagen comments that it would be very difficult to show that Plantinga's ontological argument is unsound. However, he claims, "The most promising line of attack would be to try to show that the set omniscience, omnipotence, moral perfection is not instantiated, which it would be if [Plantinga's argument] were sound. And the only way I can see to show this would be to employ some variant or other on the argument from evil" (See van Inwagen (1995, p. 36)). The main argument of this paper, if sound, shows that van Inwagen is mistaken. It accomplishes the task of showing that Plantinga's ontological argument is unsound, but, as I just noted, it is not a variant on the argument from evil.

perfection. That is, they would reject P1. But if one were to make this move, it'd be nice to hear an alternative account of the concept of moral perfection; in other words, it'd be nice to hear exactly what such theists mean when they say that God is morally perfect.³³

REFERENCES

- Adams, Robert M. 1972. "Must God Create the Best?" *The Philosophical Review* 81.3: 317–332.
- Adams, Robert M. 1984. "Saints." *Journal of Philosophy* 81: 392–401.
- Anderson, Elizabeth. 2007. "If God is Dead, Is Everything Permitted?" pp. . 215–230. Oxford: Oxford University Press.
- Clark, Michael. 1993. "On Wanting to Be Morally Perfect." *Analysis* 53: 54–56.
- Conee, Earl. 1991. "The Possibility of Power Beyond Possibility." *Philosophical Perspectives* 5: 447–473.
- Conee, Earl. 1994. "The Nature and the Impossibility of Moral Perfection." *Philosophy and Phenomenological Research* 54: 815–825.
- Crisp, Oliver. 2010. "Is Universalism a Problem for Particularists?" *Scottish Journal of Theology* 63: 1–23.
- Draper, Paul. 1989. "Pain and Pleasure: An Evidential Problem for Theists." *Noûs* 23: 331–350.
- Evans, C. Stephen. 1985. *Philosophy of Religion: Thinking about Faith*. Downers Grove: InterVarsity Press.
- Feldman, Fred. 1978. *Introductory Ethics*. Englewood Cliffs: Prentice-Hall.
- Fumerton, Richard. 2001. "Plantinga, Warrant, and Christian Belief." *Philosophia Christi* 3: 341–351.
- Garcia, Laura L. 1987. "The Essential Moral Perfection of God." *Religious Studies* 23: 137–144.
- Geach, Peter. 2002. "Is God Omnipotent?" In *Questions about God*, edited by Steven M. Cahn and David Shatz. Oxford: Oxford University Press.
- Hasker, William. 2008. *The Triumph of God over Evil*. Downers Grove, IL: InterVarsity Press.
- Hodge, Charles. 2001. *Systematic Theology*. Peabody, Mass: Hendrickson Publishers.
- Hoffman, Joshua and Gary Rosenkrantz. 2002. *The Divine Attributes*. Oxford: Blackwell Press.
- Howard-Snyder, Daniel. 1996. "The Argument from Inscrutable Evil." In *The Evidential Argument from Evil*, edited by Daniel Howard-Snyder. Bloomington, IN: Indiana University Press.

³³Special thanks to Earl Conee for his extensive comments on multiple drafts of this paper. In addition, thanks to Ed Wierenga, David Basinger, William Hasker, and several anonymous referees for helpful comments on this paper. An earlier version of this paper was presented both at the 2004 Eastern Regional Meeting of the Society of Christian Philosophers and at Houghton College. Thanks to those present at both for the helpful discussion.

- Howard-Snyder, Daniel and Frances Howard-Snyder. 1994. "How an Unsurpassable Being Can Create a Surpassable World." *Faith and Philosophy* 11: 260–268.
- Lewis, David. 2007. "Divine Evil." pp. . 231–242. Oxford: Oxford University Press.
- McGinn, Colin. 1992. "Must I Be Morally Perfect?" *Analysis* 52: 32–34.
- Morris, Thomas. 1993. "Perfection and Creation." pp. . 234–247. Ithaca: Cornell University Press.
- Morrison, Wes. 2009. "Did God Command Genocide? A Challenge to the Biblical Inerrantist." *Philosophia Christi* 11: 7–26.
- Moser, Paul. 2009. "Evidence of a Morally Perfect God." In *God is Great, God is Good: Why Believing in God is Reasonable and Responsible*, edited by William Lane Craig. Downers Grove: InterVarsity Press.
- Packer, James I. 1993. *Knowing God*. Downers Grove, IL: InterVarsity Press.
- Peterson, Michael, William Hasker, David Basinger, and Bruce Reichenbach. 1998. *Reason and Religious Belief: An Introduction to the Philosophy of Religion*. New York: Oxford University Press. 2nd edn.
- Plantinga, Alvin. 1988. "Epistemic Probability and Evil." *L'Archivio di Filosofia* 56: 557–584.
- Rowe, William. 1979. "The Problem of Evil and Some Varieties of Atheism." *American Philosophical Quarterly* 16: 335–341.
- Rowe, William. 2002. "Can God be Free?" *Faith and Philosophy* 19: 405–424.
- Sorenson, Roy. 2007. "Permission to Cheat." *Analysis* 67: 205–214.
- Stump, Eleonore. 1993. "Aquinas on the Sufferings of Job." pp. . 328–357. Ithaca, NY: Cornell University Press.
- Swinburne, Richard. 1977. *The Coherence of Theism*. Oxford: Oxford University Press.
- Taliaferro, Charles. 2002. "Is God Vain?" In *Questions about God*, edited by Steven M. Cahn and David Shatz. Oxford: Oxford University Press.
- van Inwagen, Peter. 1995. "Ontological Arguments." In *God, Knowledge, and Mystery: Essays in Philosophical Theology*. Ithaca: Cornell University Press.
- van Inwagen, Peter. 1996. "The Problem of Evil, the Problem of Air, and the Problem of Silence." In *God, Knowledge, and Mystery: Essays in Philosophical Theology*, edited by Daniel Howard-Snyder. Ithaca: Cornell University Press.
- van Inwagen, Peter. 2005. "The Problem of Evil." In *The Oxford Handbook of Philosophy of Religion*, edited by William J. Wainwright. Oxford: Oxford University Press.
- Wielenberg, Erik. 2004. "A Morally Unsurpassable God Must Create the Best." *Religious Studies* 40: 43–62.
- Wielenberg, Erik. 2009. "In Defense of Non-Natural, Non-Theistic Moral Realism." *Faith and Philosophy* 26: 23–41.

- Wierenga, Edward. 1989. *The Nature of God*. Ithaca, NY: Cornell University Press.
- Wolf, Susan. 1982. "Moral Saints." *Journal of Philosophy* 79: 419–439.

THE FATHER OF LIES?

Hud Hudson

A severe and underappreciated problem confronts anyone who holds a certain popular combination of theses—namely, that there is such a thing as knowledge by revelation alone and that a defensive maneuver known as skeptical theism is sufficient to undermine a variety of popular arguments from the magnitude, intensity, duration, and distribution of evil to the nonexistence of God. After briefly characterizing and commenting on these two positions, I identify and explore the puzzle generated by their combination, and I critically examine a variety of proposals for responding to that puzzle.

§1 The Elmer Fudd Problem

My garden isn't doing very well. I suspect there is a rabbit in the garden. I'd like to know. I call the A1-Rabbit Identification and Removal Service. They come out, and after 20 minutes of sitting on the edge of the garden, having some lunch, and stamping a bit here and there, they knock at the door ready to be paid:

"No rabbits here." "Really? How do you know?" "Look, if there were a rabbit in the garden, then we would be aware of it and we would recognize it as a rabbit, but we aren't aware of any rabbits in the garden." "Well . . . why should I accept the conditional that takes you from admitted ignorance to confident verdict?" "Because even though we haven't searched every possible rabbit-hiding place in the garden, we have searched a couple of them and we came up empty."

I don't pay. I like the move, though. The A1-Team took a sample of the garden, found it empty, engaged in inductive generalization, and adopted their no-rabbit stance. But their sample was unrepresentative—in this case because it was abysmally small—and that fact adequately undermines the strength of the inductive inference they have offered me.

I'd still like to know the answer to my question. I call the A2-Rabbit Identification and Removal Service. They come out, and after 20 hours of exhausting rabbit-searching, they knock at the door ready to be paid:

"No rabbits here." "Really? How do you know?" "Look, if there were a rabbit in the garden, then we would be aware of it and we would recognize it as a rabbit, but we aren't aware of any rabbits in the garden." "Well . . . why should I accept the conditional that takes you from admitted ignorance to confident verdict?" "Because we have searched every possible rabbit-hiding place in the garden and have done so in such a way that a rabbit cannot hide in any once-searched spot, and we came up empty."

I pay. Unlike before, I'm not offered a defense of the crucial conditional by way of inductive generalization, but rather by way of exhaustive and complete

examination—no turnip unmolested, no lettuce unturned. (And—*take note*—I simply spotted the A2-Team the thesis that they know a rabbit when they see it.)

The relevance of our two warm-up cases:

The theist and atheist can often come to agreement about a claim of this sort: “Event E is evil, and given God’s essential omnipotence, essential omniscience, and essential perfect goodness, God would have prevented the occurrence of E unless there were a compensating good or some other morally justifying reason for the permission of E.” Theist and atheist then part company on the question of whether there is such a compensating good or morally justifying reason.

Moreover, the theist and atheist can often come to agreement about a further claim, as well: “We are not aware of any compensating good or morally justifying reason for this or that (admittedly) evil state of affairs.” A popular atheistic move at this juncture is to introduce a bridge premise that exploits some such inscrutable evil, E, a bridge premise of the following form: “If there were a compensating good or some other morally justifying reason for E, then we would be aware of it and able to recognize it as such.” Once this bridge premise is defended, the joint admission of ignorance can be used to infer that there is no compensating good or other morally justifying reason for E. Accordingly—given the first point of agreement noted above—the inference to atheism is secured.

It is worth noting that with respect to compensating goods or morally justifying reasons we can claim (at the very least) to have matched the efforts of the A1-Team. That is, we have stamped around a bit in value theory and in a theory of permissions, and we have disqualified a few candidates for being the morally justifying reason for the permission of world’s horrific evils. But if our sample should turn out to be unrepresentative (perhaps by being abysmally small)—we shouldn’t pay. We’ll need a strong inductive argument for that bridge premise, and that depends on just how much more we’ve accomplished in our search than the A1-Team accomplished in theirs.

It is also worth noting that we certainly cannot claim to have matched the efforts of the A2-Team. The garden of abstracta (where hide the potential compensating goods and other morally justifying reasons) is infinitely large. Whereas we might be inclined to pay if we had surveyed the lot—we quite simply haven’t; in fact we’ve barely begun.

A far more interesting case:

The following year (when the garden is ailing once again) I’d like myself another rabbit report . . . but I don’t want to pay the exorbitant prices of the A2-Team. I call the A3-Rabbit Identification and Removal Service. They come out, and after 12 hours of rigorous and systematic searching, they knock at the door ready to be paid:

“No rabbits here.” “Really? How do you know?” “Look, if there were a rabbit in the garden, then we would be aware of it and we would recognize it as a rabbit, but we aren’t aware of any rabbits in the garden.” “Well . . . why should I accept the conditional that takes you from admitted ignorance to confident verdict?” “Because

even though we haven't searched every possible rabbit-hiding place in the garden, we have searched a very large and representative portion of it—some 60% in fact, and we came up empty. That's good enough for us."

It's good enough for me, too, and I start to write that check. As did the hopeless A1-Team, the A3-Team has offered me an inductive generalization, but unlike their undistinguished predecessors, they have produced an inductive generalization allegedly supported by a representative sample of the whole. But then, in the middle of writing that check, I find myself worrying about a couple of issues.

First, a small point: Whether or not the sample is representative will not simply be a function of its size, even if it is very large. And, of course, the analogue I care about—the garden of abstracta—is so infinitely vast that sheer sample-size can't be reported with phrases like "we have surveyed 60% of the candidates for morally justifying reasons."

Second, a somewhat larger point: We are interested in whether we have good reason to believe the bridge premise that promises to take us from an admission that we do not know of a morally justifying reason to the conclusion that there isn't one. Suppose we wanted to undermine the strength of the inductive generalization described above which purports to furnish us with that good reason.

Accordingly, consider a continuation of the conversation at the door in our A3-Rabbit Identification and Removal Service Case:

"I know you boys searched a very large portion of the garden, but you'll note that a significant portion of the garden is covered by a tarp held in place by rocks and earth. Was any of your sample taken from under the tarp?" "No, sir, we gave it a try, but that tarp was just too heavy to lift." "But given the business you're in, I'm sure you'll agree that it's likely that if there were a rabbit in the garden, it would be hiding under the tarp." "Yes, sir." "Then, despite its being quite large, I've discovered a reason to deny the representativeness of your sample and a reason to suspect the strength of your inductive generalization." "Yes, sir."

At this point, I'm inclined to stop payment on that check.

Now what is supposed to correspond to the tarp-in-the-garden image? Just as the tarp covers a part of the garden that was not examined (because it was too heavy to lift), so too, some portions of abstracta are unsearchable (because—to take one example—we cannot penetrate their complexity given our crude cognitive capacities and tools). Let us inquire, then: Isn't it likely that if there were a morally justifying reason for permitting the world's horrific evils, it would be located in a portion of the garden of abstracta that is impenetrable to us?

Several affirmative answers have been presented for evaluation. Here are two for starters: (i) It wouldn't be at all surprising if it is likely that the magnitude, intensity, duration, and distribution of the evils to be accounted for are themselves so exceedingly complex that a morally justifying reason would exhibit the same feature, a feature that would place it beyond our ken. We are, after all, frequently reminded just how unfathomable is the full and unadulterated history of evil; (ii) It wouldn't be at all surprising if it is likely that if, owing to a desire to cultivate a certain kind of attitude in his creatures (a desire that would

move God if he exists), all the morally justifying reasons are masked by divine intervention and are thus kept safely in obscurity.

Just to be clear—if either of those considerations provides us with a good reason to believe the sample is unrepresentative, the inductive argument for the bridge premise is in real trouble. But there is no need to aim so high and defend such speculative answers; let's ratchet down—even if we aim lower and claim only that we have no good reason to believe that the sample is representative or that we are in the dark or in doubt about whether the sample is representative, the inductive argument is still in jeopardy. So called skeptical theists' are in the business of denying the bridge premise we have been discussing in just this manner. In general, they argue either that we have no good reason to believe (or else that we are in the dark about whether) the goods we are aware of are representative of the goods that there are (see Wykstra (1984), Alston (1991), Howard-Snyder (1996), Bergmann (2001, 2009), and McBrayer (2010)). And our situation gets even worse.

The Elmer Fudd Problem: Bugs Bunny was a cartoon hero of my childhood. Bugs' nemesis, Elmer Fudd, was forever shotgunning for Bugs but was also astonishingly stupid. All Bugs had to do was throw on a dress and hide his ears, and Elmer immediately mistook him for anything but a rabbit and (disturbingly—for a child's cartoon) usually for a prospective love-interest. Let us not forget that we have here pretended that (as with rabbits) we know morally justifying reasons when we see em. But we should really drop this conceit. We may well be the Elmer Fudds of value theory. Once again, we have no good reason to believe (or else we are in the dark about whether) the entailment relations we know of between goods and permitted evils are representative of the entailments relations there are. Consequently, we may well have already discovered an exceedingly valuable good that would justify God in permitting this or that horrific evil and then (after failing to recognize its necessary connection to the evil in question) mistakenly rejected its candidacy on the grounds that it doesn't require permission of the evil.

And our situation gets even worse, yet again, for we have no good reason to believe (or else we are in the dark about whether) the degree of value we recognize in those goods we are aware of is representative of the total degree of value those goods actually manifest. Consequently, we may well have already discovered a good that would justify God in permitting this or that horrific evil and also have already discovered its necessary connection to the evil in question and then (after failing to recognize its full range of goodness) mistakenly rejected its candidacy on the grounds that it was not sufficiently compensatory.

Consequently, argue our skeptical theists, we have no good reason to think that if there were such a compensating good or some other morally justifying reason for the world's horrific evils, we would be aware of it, or—if it were somehow an object of mere awareness—that we would be able to recognize its full degree of value or its function as a compensating good or morally justifying reason.

Such a realization of our epistemic position does not, of course, amount to a reason to turn theist, but absent any other way of demonstrating the lack of a compensating good or morally justifying reason, it would (if successful) eliminate many of the most powerful arguments for atheism.

§2 For The Bible Tells Me So

A number of religious traditions—including Judaism, Christianity, and Islam—have maintained that among the sources of knowledge available to human beings we should recognize divine revelation. The details differ but in ways that don't matter much for our purposes. I shall focus on the case of Christianity in the following discussion, a religion in which (with a few dissenters here and there) the historical and contemporary view appears to be that God has revealed certain truths touching on matters of consequence to all human persons near and far, past and future, and that these truths are not ones we could have fully discovered left to our own devices (see Swinburne (2007) and Davis (2009)).

Several models for divine revelation of truths (as opposed to the revelation of God himself) are on offer: Sometimes the proposed mechanism is causal, effected perhaps by dreams, or visions, or some sort of direct neurophysiological tinkering. Sometimes the truth is manifested in some person, some bit of behavior, some miracle, or some other magnificent chain of events. Sometimes the revelation is portrayed as a kind of divine testimony—addressed and spoken to an individual or to a people and communicated directly in God's own voice, or by prophet, or by inspired scripture. (For discussion of the models, see Mavrodes (1988)). However it eventually gets transferred, such testimony has propositional content (Davis (2009, p. 35)), and the force of the term alone' in the phrase knowledge by revelation alone' is simply once again to signal that human powers of cognition—reason, understanding, imagination, sensation, introspection, memory—are not up to the task of discovering the truth values of these propositions on their own. Apart from revelation (if there is such a thing) we do not have any independent means of verification or any significant epistemic access to the relevant subject matter.

Although particular examples are always contested, candidates for bits of knowledge by revelation alone with respect to the Christian tradition include claims regarding the fall of humanity, the trinity, the incarnation, the atonement, and the general resurrection and life of the world to come. Although specific traditions and creeds aren't really the focus here and although any doctrine whose credentials are restricted to the testimony of an omniscient and perfectly good being will suffice to generate the conflict I intend to examine, I will need a placeholder in our discussion.

“We look for the resurrection of the dead and the life of the world to come.” So ends the Nicene Creed with a reference to one of the most central teachings of Christianity. Similarly, the Apostles' and Athanasian Creeds explicitly and prominently call attention to the resurrection of the body. In the discussion to follow I will take this position to imply (at the very least) what I will

call *the general resurrection thesis*—the view that every human person who has ever died will rise again from the dead—and in what follows I will invoke this (controversial) thesis as our example of a candidate for knowledge by revelation alone.

§3 Notrustem Inferences

It is commonly agreed that being deceived is a bad state of affairs and that lying is morally wrong. A more careful pronouncement, however, is that there are some features of being deceived that are of disvalue (e.g., a mismatch between one's cognitive states and the world) and that lying is *prima facie* morally wrong (i.e., that there is moral reason not to deceive others). Nearly everyone, however, recognizes both that deception can be extrinsically valuable (e.g., it can lead to substantial goods such as saving the deceived from making a life-ruining mistake in a moment of passion) and that the moral presumption against lying can be overridden by even stronger moral reasons in favor of deception on a particular occasion. As is well known, *prima facie* obligations can come into conflict with one another, and when they do the directive against lying quite obviously is not always the most stringent, overriding consideration in moral decision making.

The problem, then, is straightforward. Consider any piece of alleged knowledge by revelation alone, *K*. Moreover—and *this is the surprising part*—simply grant without a fuss that the testimony inviting us to adopt *K* as a truth comes from someone whom we know to be both essentially omniscient and essentially perfectly good. (Of course, traditions may differ on those two aspects of the deity, but I am interested in the problem that emerges on even the most generous list of known, divine attributes.)

Here, then, is a first pass at characterizing our puzzling predicament: *Do we thereby have knowledge of K?* Well, not if *K* is false—if *K* is false, then we are deceived by someone who knows that *K* is false.

Do we have a way of verifying the truth of K and exposing a deception, if deception it be? Well, not if *K* is a genuine candidate for knowledge by revelation alone, for earlier we specified we do not have any independent means of verification or any significant epistemic access to the status of such candidates.

Do we have a guarantee that God would not deceive us about whether K is true? Well, not if our being deceived about *K* is the kind of bad state of affairs for which there exists a compensating good or morally justifying reason. If there is a compensating good or morally justifying reason for such deception, God's essential perfect goodness is not in any way impugned by the deception—on the contrary, it may be morally obligatory to so deceive us.

Do we know that there is no such compensating good or morally justifying reason in this case? There's the rub . . . apparently not, if we are among those who adhere to the skeptical theist's defensive maneuver for undercutting arguments from evil to atheism. It would seem that consistency would require us to claim ignorance here as before and for more or less the same reasons, too.

Consequently, our claim to knowledge seems to be threatened: We cannot without reservation trust such divine pronouncements—even if we simply help ourselves to the background assumptions that God exists, that God is essentially omniscient and

essentially perfectly good, that God has provided us with his testimony, and that we have interpreted that testimony aright. And once we have lost this particular kind of trust in the testimony, it cannot be the source of testimonial knowledge.

Given the centrality and importance of the sorts of theses identified in Section 2 above as candidates for bits of knowledge by revelation alone, such a “notrustem inference” points to a severe and underappreciated problem confronting anyone who holds this popular combination of views. (Not entirely underappreciated, however. See Wielenberg (2010), Maitzen (2013), Wilks (2013).) Finally, it is worth remarking that this problem is significantly different from other critiques of skeptical theism, even those which are explicitly grounded in claims about moral deliberation and moral knowledge (for such critiques see Almeida and Oppy (2003), Pereboom (2005), and Jordan (2006); for an evaluation of such critiques, see Bergmann and Rea (2005) and Howard-Snyder (2010)).

In the remaining sections, I will critically examine and evaluate three proposals for responding to this problem.

§4 The George-Washington Defense

Perhaps we are in the clear on the grounds that God simply cannot tell a lie, since doing so either betrays a kind of weakness or else is always morally wrong and since God’s perfection is incompatible with both defects and moral wrongdoing. After all, the Commandment doesn’t say “Thou Shalt Not Lie, Unless, You Know, You Can Wring Some Advantage Out Of It.” Moreover, scripture contains a number of passages in which lying and deceiving are subjected to heavy criticism and condemnation. And historical champions such as Augustine, Gregory the Great, and Aquinas were all on board with an absolute prohibition on lying as is the *Catechism of the Catholic Church* (see Tollefsen and Pruss (2011)).

Such support is frequently hedged, though, and there are considerations that make the prohibition less severe than it seems at first blush. Whereas Kant, in his ? essay, “On a Supposed Right to Lie from Altruistic Motives,” might actually have advanced the view that—“we should let the sky fall before committing the certain sin of telling a lie or practicing deception”—it is harder to believe that the sources just cited could have had that view in mind. Four quick comments on this theme:

First (and significantly), the prohibition does not usually extend to all cases of deception but rather to the special case of deception known as lying. As noted above, however, knowledge by revelation could come in many different flavors, only some of which involve the sort of direct assertion that is subject to the charge of lying (as opposed to some other form of deception).

Second, as many authors have noted and sometimes lamented, scripture is also a source of evidence both for the claim that God himself has perpetrated (whether directly or indirectly) an intriguing assortment of lies and cunning deceptions as well as for the claim that such behavior was laudable and morally permissible (see Smith (2011a)).

Third, the *Catechism of the Catholic Church* and the support of certain influential individuals are often qualified by the presence of excepting-clauses of some kind or other. For example, in the case of the *Catechism's* first version, the rule against lying was softened a bit by an accompanying definition of lying' which insisted upon "having the right to know the truth" as a necessary condition of being lied to. The prohibition, then, would be silent on cases in which the right had never been present to begin with or else had been either waived or forfeited. In a similar fashion, by imposing restrictions on what counts as a lie, our historical figures are able to consistently advocate an absolute ban on lying while acknowledging the permissibility of some instances of deliberate deception by way of explicit assertion known by the speaker to be false. Compare the fact that Aquinas or the Church could also stand in favor of the prohibition against murder and yet endorse the permissibility of deliberate killings in certain cases of self-defense—not because such actions are permissible murders, but because they do not fall under the scope of the proper analysis of murder' at all (see Smith (2011b)).

It would be an odd morality that maintains that deliberate killing in self-defense cases when innocent lives are at stake is acceptable moral behavior but that lying to achieve the same ends in those very cases would be inexcusable, unqualified moral wrongdoing.

Fourth, it seems dialectically inappropriate to mount a defense of an *ultima facie* duty to refrain from lying by appealing to the very sources whose veracity has been called into question. If I am (in fact) being lied to by Dean Zimmerman (and for my own good), and I come to suspect that this might be the case, I doubt I'll make much useful headway in my attempt to sleuth out the facts by asking Dean if he is lying to me. I'd do better to ask Meghan. If the context in which the question arises of the permissibility of engaging in direct and intentional deception by way of false assertion is one in which scripture and the authority of certain figures is presupposed as reliable and trustworthy ground, then we primarily face the problem of arriving at the correct interpretation of those sources. But that is decidedly not the context in which the question is being raised here.

The George-Washington defense against notrustem inferences would presumably have to be founded on independent moral reasoning rather than any appeal to knowledge by revelation alone (or else be subject to a question-begging critique). And when turning to and inspecting that independent moral reasoning, it seems clear that the widespread and considered opinion remains that despite the disvalue of being misled and the *prima facie* wrongness of lying, the moral presumption against lying can be overridden by even stronger moral reasons in favor of deception on a particular occasion. Once again, then, *prima facie* obligations can come into conflict with one another, and when they do the directive against lying quite obviously is not always the most stringent, overriding consideration in moral decision making.

§5 The Greatest (Tall) Tale Ever Told

Perhaps we are in the clear on the grounds that whereas God can tell the occasional lie, he simply cannot deceive us about something as important as the fall of humanity, the trinity, the incarnation, the atonement, or the general resurrection thesis.

In other words, the prohibition need not be on the telling of lies as such, but rather on the telling of certain lies rather than others. Well, why? What, exactly, is the problem? Apparently, the idea is that the consequences of being deceived on such momentous topics would be so severe, so unfathomable, so unutterably bad, that despite the intellectual humility ordinarily manifested by the skeptical theist, he can quite clearly discern that an atrocity of this magnitude just could not be justified.

This ploy strikes me as wrongheaded twice over:

First, it is a half-hearted (and vulnerable) skeptical theism which professes ignorance about how much we know about just which things are good, about just how good they are, and about the necessary conditions of their realization in a wide variety of scenarios, but changes its tune in a few special cases to declare that something like *that* just couldn't be tolerated. Whence this confidence? I suspect it might have its origins in the plausible intuition that we know (for example) that there could be no morally justifying reason to permit a world consisting of nothing but sentient beings in devastating pain at every moment of their existence. But even if we do know such a thing and even if such knowledge is more or less consistent with the general attitudes of skeptical theism elsewhere, note that the extreme scenario before us is not at all relevantly similar to the case of divine deception at issue. A lot gets packed into the nothing but' qualification in the phrase "nothing but sentient beings in devastating pain at every moment of their existence." In particular, it has the effect of stipulating a case in which no other factors (apart from those entailed by that description) are relevant to determining its overall value or justifiability. Perhaps, under the guidance of that ideal stipulation, we are in a special position to see that such a world is inconsistent with God's nature, but there is no guarantee of the absence of other relevant factors in a case of divine deception about the general resurrection thesis, and absent the absence, we have no business temporarily departing from the recommendations of skeptical theism to make confident pronouncements about the intolerability of divine deception on the matters of revelation.

Second, the proposal strikes me as a bit of overexcited hyperbole, a kind of reverse-Panglossianism—"No Worse, There is None / Pitched Past Pitch of Grief." Bad? Yes. But this bad? Suppose that—as several Christians have argued—we human persons, are material beings, and suppose further that despite the half-dozen or so proposals for reconciling a materialism for human persons with the doctrine of the general resurrection, this union is simply metaphysically impossible (see Hudson (2001)). Yet perhaps even creatures essentially barred from a certain kind of afterlife are worth creating anyway, and moreover, ought to be deceived on exactly that point. To be clear, that's not my view. I think I am a material object and that materialism for human persons is consistent with the general resurrection thesis, and so I look for the general resurrection

and the life of the world to come. But if I'm wrong about those things, and God has deceived me about the prospects for things of my kind, I just can't see that I'm in a position to say that the deception-component of such a state of affairs is as bad as the world-of-unrelenting-and-devastating-pain or anything else that convinces me that it could not be permitted by a morally justifying reason.

§6 In God We Trust

My colleague, Michael Rea, has always been a divine friend to me. Of course, that adjective has to be taken in context. Mike isn't God, but he isn't wholly unlike God, either. Mike and God share *some* properties relevant to recognizing him as the fine friend I know him to be. Perhaps Mike and God share some other properties, as well, properties that enable him to impart knowledge to me by way of testimony, despite the fact that I cannot completely trust what he has to say either. It is important to note, however, that my lack of complete trust need not be based on any failing of Mike's; indeed (as just suggested) it may be predicated on some respect in which Mike resembles God—namely, on his knowledge and goodness.

A brief story: I first gave this paper as an after-dinner talk at a conference organized by Mike at the University of Notre Dame. Mike knows I'm fond of rum and coke, that I tend to get a touch nervous before giving talks, and that if there is rum and coke at hand before I give a talk—it's not at hand for long. Steeling myself against the onslaught of questions soon to be coming from people who had thought more about skeptical theism than I had, I asked Mike if the drink he had just fixed and handed over was indeed a rum and coke. He *knew* the answer. And he had my interests in mind. And he had other goals, too; in particular, he wanted the talk and discussion to go well. He fully understood that my being deceived would have disvalue for me and that his lying would be *prima facie* morally wrong, and yet that he nevertheless may have had other *prima facie* duties in favor of misleading me that were even more stringent in those circumstances. Moreover, as it turns out, on this occasion he did both what he ought to have done and what he knew he ought to have done.

Now it is easy to understand how our story could have concluded in either of two ways: *The first (merely possible) ending:* Mike lied to me (quite justifiably) and handed me a coke. *The second (and actual) ending:* Mike told me the truth (hoped for the best) and handed me a rum and coke. Was my claim to testimonial knowledge that I was then downing a rum and coke imperiled by my concession that I could see how Mike (while knowing the facts of the matter) may have lied to me as a manifestation of his own goodness?

Just to be clear—of course I wouldn't have known, if I had in fact been successfully deceived (for then the relevant belief would have been false). The question is rather, given that Mike spoke honestly, did I then have knowledge by way of receiving his (true and known) testimony, despite my realization that there may well have been morally justifying reasons for my being a victim of deception on that occasion? Initially, perhaps, one may think the answer is “of

course,” and furthermore, “if I can gain testimonial knowledge from Mike in the face of such uncertainty and doubt, surely I can receive it from God.”

But how persuasive, on reflection, is transferring that reaction in the Mike-scenario to the case involving God? I have to admit, I’m secretly hoping it furnishes an adequate response, since I’d like some decent way to reconcile my views and this seems to me the most promising. But allow me to at least state the case for the opposition.

I suspect the problem with this strategy will be found in a feature in which Mike is unlike and inferior to God but which surprisingly makes it all that much easier to acquire knowledge by way of Mike’s testimony. Accordingly, let us grant for the sake of argument that I acquired the knowledge that I was drinking a rum and coke when Mike told me so and then explore the difficulties with generalizing on this admission.

Testimonial knowledge is a controversial and hotly-debated topic. The non-reductionists see testimony as a basic source of justification (on a par with perception and memory) and as requiring very little effort from the recipient—it’s an innocent-until-given-grounds-for-guilt view; in the absence of relevant defeaters, one may justifiably accept a piece of testimony upon hearing it (see Lackey (2006) and Lackey (2011)). By contrast, in the (to my mind) more plausible reductionist view, testimony requires something more than the lack of undefeated defeaters to impart justification, and different species of reductionists are divided by their different answers to the what-else’ question. One popular proposal for the additional ingredient is a non-testimonial good reason for thinking that testimony is generally reliable or trustworthy, while another popular proposal advocates the weaker constraint of a non-testimonial good reason for thinking that a particular instance of testimony is reliable or trustworthy (Lackey (2006)).

Focusing on the weaker version of reductionism (featuring the local rather than global requirement on reliability), I realize that I have a plausible chance of meeting that requirement in my interactions with Mike. After all, he and I are relevantly similar on a wide variety of measures, and I am thereby well-positioned to get good evidence from perception, memory, induction and the like relevant to a judgment of reliability on this occasion. Besides, I’m not altogether a hopeless judge of the likely consequences of my imbibing before presenting, and I share much of Mike’s information about whether there are such morally justifying reasons to deceive me in the scenario just related.

In fact, it is precisely because I resemble Mike in a way that neither of us resembles God that I have any prayer of meeting this constraint. But how shall I—with my skeptical-theist hands securely tied behind my back—provide myself with analogous assurance in the case involving God? I am not relevantly similar to the source of revelation. I am not well-positioned to get good evidence from perception, memory, induction, and the like relevant to a judgment of reliability on this occasion (*more on this crucial point in the following section*). And I am a hopeless judge of what is at stake—cosmically speaking—if I and my fellows are not deceived in some comprehensive, irresistible, and undetectable fashion.

Indeed, I am quite utterly in the dark on that matter, as we have been saying.

§7 How Bad Can It Get?

So what if the replies are not adequate? As committed skeptical theists could we just take our lumps? Could we learn to live with the possibility of divine deception about the most important of matters? Could we—*wait for it*—could we *bear this false witness*?

Of course we have recourse to the speech with which Descartes once taunted the evil demon: “Let him deceive me about whatever he can, he will never bring it about that I am not so long as I am conscious that I am.” But, then again, each of us securing his own existence isn’t all that much comfort.

On the assumptions that we do not know there is an absolute moral prohibition on lying and that the characteristic tenets of skeptical theism are true, it is epistemically possible that an essentially omnipotent, essentially omniscient, perfectly good God deceives us about a *tremendous* number of topics—from whether the world is billions of years old to whether the incarnation occurred to whether I myself am embodied to whether I had eggs for breakfast. That should be disconcerting enough; but even worse, on those assumptions, we should apparently be utterly in the dark about whether that is exactly what is happening. Or should we? Can we somehow contain the skeptical threat?

In one case, the answer seems to be clearly in the negative. One unfortunate cost of invoking skeptical theism to combat certain atheistic arguments is to suffer its undermining certain theistic arguments, as well. The fine-tuning argument, for example, contains a premise which asserts that given the fine-tuning thesis, the existence of life-permitting, cosmic conditions is very probable under the hypothesis that God exists (see Collins (1999)). The defense of this premise turns on a line of reasoning linking God’s perfect goodness and omnipotence to the high likelihood of creating a world in which free creatures (like us) can interact with one another. Why so likely? Because, *everything told*, such a world would be so wonderfully valuable. The skeptical theist, however, is in the business of cautioning us against drawing inferences of exactly that kind. We are simply in the dark, she explains, about whether the existence of creatures (like us) freely interacting with one another ensures a valuable world *all things considered*, for we simply have no idea whether our presence is inconsistent with a much more magnificent good or else requires the permission of something exceedingly evil. Perhaps we should simply concede this unfortunate consequence of skeptical theism (as Bergmann (2009) does and for just these reasons) and hope that the collateral damage comes to an end there.

After all, why should I think that my perfectly reasonable belief that the cosmos is several billion years old or that I am embodied or that I had eggs for breakfast are even remotely jeopardized? As Michael Bergmann has argued, my knowledge of these truths is not acquired by reflecting on possible goods, possible evils, and the necessary conditions of their realization, but rather in some other—*some independent*—way (Bergmann (2009)). And wasn’t it exactly

this lack of independent access to verifying truth that made putative cases of knowledge by revelation alone so vulnerable to the commitments of skeptical theism? Surely, then, we can draw a line between the genuine skeptical threats concerning moral matters or divine testimony and the merely apparent skeptical consequences elsewhere.

I'm not so sure; if we get this far, there may be no turning back.

Suppose I temporarily breathe a sigh of relief, reminding myself that many of my beliefs are backed independently by appeal to astronomy, biology, chemistry, geology, physics, intuition, memory, perception, introspection, the natural light of reason, or even Reidian common sense (the last appeal being the focus of Bergmann (2012)). But the relief is not long-lasting. Ian Wilks reintroduces anxiety—for the theist in particular—by remarking: “Of course, radical skepticism does not yield so easily. For all we know God has invested us with a delusional tendency of common sense, and we accept on its basis as true, beliefs that are actually false. For all we know there is a greater good served, or evil avoided, by this deception” (Wilks (2013)).

It seems it won't do to simply inquire—“Is my belief that P grounded in some judgment about possible goods, possible evils, and the necessary conditions of their realization?”—and then to take skeptical theism to be a humility-inducing corrective to Yes' replies, but altogether silent in the case of No' replies. For no matter what alleged independent evidence is marshaled in favor of the thesis at hand, skeptical theism can lead our theist to the view that we should be utterly in the dark about whether investing us with the means to obtain that misleading evidence is the very mechanism that our perfectly-good-and-deceiving creator selected for our arriving at the confident but false judgment that P.

Here's one way of seeing the difficulty: If there is a morally obligating reason for God to deceive me, then I am deceived. If there is no morally justifying reason for God to deceive me, then I am not deceived. If there is a morally justifying reason for God to deceive me, then either I am or am not deceived depending on God's other purposes. Skeptical theists would remind me that I am utterly in the dark about which of those three antecedents is satisfied. And thus the darkness expands so that I am also utterly in the dark about whether I am deceived in the most comprehensive, irresistible, and undetectable fashion.

In short, the skepticism in question threatens to *explode* for our theist—i.e., for anyone who accepts that there is a being possessed of the power to so deceive—quickly moving from a well-contained strategy for opposing a popular style of atheistic reasoning to a near global catastrophe, threatening to undermine the reasonableness of our views in nearly all matters, great and small. (For further discussion of the threats of such explosion and strategies for containment, see Russell (1996), Beaudoin (2005), Bergmann (2009, 2012), Roeber (2009), Wilks (2009) and Wilks (2013)).

§8 Some Disconcerting Options

So—what to do?

Option I—Skepticism Run Amok: Skeptical theism is non-negotiable, its skepticism spreads uncontrollably (for the theist) to other issues, and we simply have to get used to trading in our confidence in alleged divine revelation for comfort in the thought that even if we are the victims of systematic and far-reaching lies, at least God ain't misbehavin'.

Option II—Revisiting Arguments from Evil: Knowledge by revelation alone is non-negotiable, divine testimony can be and has been bestowed upon us, and we simply have to get used to resisting arguments from evil in some way other than bemoaning how little we know about goods, their full range of value, and the necessary conditions of their realization.

Option III—Demonstrating Consistency: Skeptical theism is (despite appearances) thoroughly consistent with knowledge by revelation alone. But—*how*—exactly? Here's a hopeful parting thought.

If I were picked up (as if by some unseen hand) and placed in the middle of a labyrinth, I might, after hours or days of unsuccessful attempts to leave, come to the skeptical position that there just might be no way out of its interior. On the other hand, if I found myself unable to escape a maze only after I had wandered in on my own, I would be far less likely to embrace that skeptical hypothesis—after all, if there's a way in, one would think there's a way out.

It seems to me that my current perplexity is more similar to the latter case than to the former. I am unsure which way to turn next, because I accept both skeptical theism and knowledge by revelation alone, and each move toward reconciliation thus far seems to me something of a dead end. Still, I wandered into this mess only after accepting certain theses—that God exists and is both essentially perfectly good and essentially omniscient, that the characteristic tenets of skeptical theism are true, that I have some revelation-based knowledge, that lying is only *prima facie* morally wrong, and so forth. The tension I have been exploring arises only for someone who is inclined, as I am, to take on that particular collection of theses, and so I'm hoping that by unraveling whatever epistemological story successfully explains how I reasonably arrived at that combination of views, I may thereby discover a guide to extricate me from this puzzle.

But I have to say that this particular labyrinth is a dark and uncomfortable place, and I would be delighted if anyone would like to play Ariadne to my Theseus and just get me the hell outta here.¹

REFERENCES

Almeida, Michael and Graham Oppy. 2003. "Skeptical Theism and Evidential Arguments from Evil." *Australasian Journal of Philosophy* 81: 496–516.

¹For comments and criticism on earlier versions of this paper I thank Michael Bergmann, Godehard Brntrup, Trent Dougherty, Peter Forrest, John Hawthorne, Daniel Howard-Snyder, Ned Markosian, Trenton Merricks, Sam Newlands, Guo Peng, John Pittard, Michael Rea, Blake Roeber, Jeffrey Russell, Thomas Senor, Amy Seymour, Ge Siyou, Eleonore Stump, Jason Turner, Dennis Whitcomb, and audiences at Logos 2012 (the University of Notre Dame) and at Meiguo Metaphysical Mayhem (Peking University).

- Alston, William. 1991. "The Inductive Argument from Evil and the Human Cognitive Condition." *Philosophical Perspectives* 5: 29–67.
- Beaudoin, John. 2005. "Skepticism and the Skeptical Theist." *Faith and Philosophy* 22: 42–56.
- Bergmann, Michael. 2001. "Skeptical Theism and Rowe's New Evidential Argument from Evil." *Noûs* 35: 278–296.
- Bergmann, Michael. 2009. "Skeptical Theism and the Problem of Evil." pp. . 374–399. Oxford: Oxford University Press.
- Bergmann, Michael. 2012. "Commonsense Skeptical Theism." In *Science, Religion, and Metaphysics: New Essays on the Philosophy of Alvin Plantinga*, edited by Kelly Clark and Michael Rea. Oxford: Oxford University Press.
- Bergmann, Michael and Michael Rea. 2005. "In Defense of Skeptical Theism." *Australasian Journal of Philosophy* 83: 241–251.
- Collins, Robin. 1999. "The Fine-Tuning Design Argument: A Scientific Argument for the Existence of God." pp. . 47–75. Grand Rapids, MI: Eerdmans.
- Davis, Stephen. 2009. "Revelation and Inspiration." pp. . 30–53. Oxford: Oxford University Press.
- Howard-Snyder, Daniel. 1996. "The Argument from Inscrutable Evil." In *The Evidential Argument from Evil*, edited by Daniel Howard-Snyder. Bloomington, IN: Indiana University Press.
- Howard-Snyder, Daniel. 2010. "Epistemic Humility, Arguments from Evil, and Moral Skepticism." In *Oxford Studies in the Philosophy of Religion*, edited by J. L. Kvanvig. Oxford: Oxford University Press.
- Hudson, Hud. 2001. *A Materialist Metaphysics of the Human Person*. Ithaca: Cornell University Press.
- Jordan, Jeff. 2006. "Does Skeptical Theism Imply Moral Skepticism?" *Philosophy and Phenomenological Research* 72: 403–417.
- Lackey, Jennifer. 2006. "Introduction." pp. . 1–21. Oxford: Oxford University Press.
- Lackey, Jennifer. 2011. "Testimony: Acquiring Knowledge from Others." pp. . 71–91. Oxford: Oxford University Press.
- Maitzen, Stephen. 2013. "The Moral Skepticism Objection to Skeptical Theism." In *A Companion to the Problem of Evil*, edited by Daniel Howard-Snyder and Justin McBrayer. Malden, MA: Blackwell.
- Mavrodes, George. 1988. *Revelation in Religious Belief*. Philadelphia: Temple University Press.
- McBrayer, Justin. 2010. "Skeptical Theism." *Philosophy Compass* 5 (7): 611–623.
- Pereboom, Derk. 2005. "The Problem of Evil." pp. . 148–170. Malden, MA: Blackwell.

- Roeber, D. Blake. 2009. "Does the Theist Have an Epistemic Advantage Over the Atheist? Plantinga and Descartes on Theism, Atheism, and Skepticism." *Journal of Philosophical Research* 34: 305–328.
- Russell, Bruce. 1996. "Defenseless." pp. . 193–205. Bloomington, IN: Indiana University Press.
- Smith, Hanet E. 2011a. "Why Tollefsen and Pruss are Wrong about Lying." *First Things* (December).
- Smith, Janet E. 2011b. "Fig Leaves and Falsehoods." *First Things* (June/July).
- Swinburne, Richard. 2007. *Revelation: From Metaphor to Analogy*. Oxford: Oxford University Press.
- Tollefsen, Christopher and Alexander Pruss. 2011. "The Case Against False Assertions." *First Things* (September).
- Wielenberg, Erik. 2010. "Sceptical Theism and Divine Lies." *Religious Studies* 46: 509–523.
- Wilks, Ian. 2009. "Skeptical Theism and Empirical Unfalsifiability." *Faith and Philosophy* 26: 64–76.
- Wilks, Ian. 2013. "The Global Skepticism Objection." In *A Companion to the Problem of Evil*. Malden, MA: Blackwell.
- Wykstra, Stephen. 1984. "The Humean Obstacle to Evidential Arguments from Suffering: On Avoiding the Evils of Appearance." *International Journal for Philosophy of Religion* 16: 73–93.

GOD ACTS IN THE QUANTUM WORLD

*Bradley Monton***1. Introduction**

Let's suppose, for the purposes of this paper, that God exists, as a personal, omnipotent being, and as the creator and sustainer of the physical universe. Let's also suppose, for the purposes of this paper, that the theologians who believe God would not intervene in the world are correct. Specifically, we're supposing that God would not violate the laws of nature that he created for the world, since that would involve God dealing in two different manners with his creation (McMullin (1993, p. 324)). (The worry is: why would God create laws of nature that govern the world, and then violate them?) Even though these theologians believe that God doesn't violate the laws of nature, they do believe that God can act in the world, as long as he does so in ways that don't violate the laws. Nancey Murphy, Robert Russell, and Thomas Tracy are prominent theologians who are proponents of this position, called "noninterventionist special divine action". (Murphy (1995), Russell (2002), and Tracy (2002) are three essays which are representative of an extensive literature on this topic.)

So how could God act in the world without intervening? One way for God to do this is by acting at the indeterministic quantum level. For example, if there's some quantum process that has a 10% chance of yielding outcome A, and a 90% chance of yielding outcome B, God can, in a particular instance of this process, decide which outcome will result, without violating any laws.

It has sometimes been maintained that God's actions in the world are quite limited, if all he can do is intervene at the quantum level. My first goal is to show that, on some ways of understanding quantum mechanics, that is false: God's actions are almost unlimited. This gives God the almost unlimited freedom to bring about any effect in the physical world, including (for example) parting the sea, changing water into wine, resurrecting the dead, and producing fish and loaves of bread.

Moreover, it has sometimes been maintained that God's actions in the world, within this quantum-mechanical framework, are problematically *episodic*, in the sense that God can't continuously act in the world. I will show that, on some ways of understanding quantum mechanics, this is false; God's actions can be continuous (though they need not be).

Quantum mechanics is probably a false theory—it can't accommodate the empirical evidence that supports general relativity, and that's one reason that physicists are working on a theory of quantum gravity, to supplant both quantum

mechanics and general relativity. But still, this discussion presents a useful model of how it could be the case that God can have almost unlimited freedom to continuously act in the world, in a noninterventionist way.

2. Quantum Mechanics

A key part of my paper is about how to understand quantum mechanics, but I recognize that not all readers will want to deal with the details. So feel free to skip ahead to the next section, where I'll summarize the results of this section and show that, on some ways of understanding quantum mechanics, even if all God can do is intervene at the quantum level, his actions in the world are almost unlimited. Those who read on will see that my discussion in this section is necessarily brief. There's a lot one could say about how to understand quantum mechanics; I'll just provide the highlights to situate my discussion of God's actions.

Let's start with a crucial point (to reward those readers who are sticking around). Some versions of quantum mechanics are indeterministic, while others are deterministic. The main deterministic version is David Bohm's pilot wave theory. According to this interpretation, particles always have definite positions, and outcomes of processes that look indeterministic are actually determined by the precise locations of the particles (locations which are only imprecisely accessible to us). If the laws of nature are governing propositions established by God, and the laws of nature are Bohmian, then God won't be able to act in the world without contravening the laws he established.

The many worlds interpretation is another deterministic version of quantum mechanics. Not only does this involve the same restrictions on God's actions as Bohm's theory, it also allows different branches of the universe (colloquially called different "worlds") to go in many different ways, leading to the arguably theologically unhappy consequence that every possibility allowed by the initial conditions of the universe (and the laws of quantum mechanics) is actually instantiated in some branch. This arguably makes the problem of evil even more of a problem than one would have thought, since there will be branches of the universe where the branch is filled with evil events. (For related theological issues, see Monton (2010).)

While there are deterministic versions of quantum mechanics, most all physicists favor indeterministic versions. According to (one version of) the Copenhagen interpretation, quantum systems evolve deterministically, except when a measurement occurs, in which case the quantum state of the system collapses to a particular state, ensuring that the measurement has a definite outcome. Philosophers of physics nowadays universally reject this interpretation as at least being incomplete (and most likely false). The problem is that what count as a measurement is not specified by the interpretation, and anyways, measurements probably aren't fundamentally different physical processes from all the other physical processes that occur.

There are two popular indeterministic versions of quantum mechanics amongst philosophers of physics, modal interpretations and the GRW theory. According to modal interpretations, the quantum state of a system never collapses, but in addition to the quantum state, a system has a *value state*, which specifies which properties the system actually has. The quantum state evolves deterministically, but on most modal interpretations, the value state evolves indeterministically, and that's where the fundamental indeterminism arises in the world. There are many different versions of modal interpretations, corresponding to different specifications of which properties are determinate, and how the properties evolve. (For more see Bub (1997).)

Unlike the modal interpretations, the GRW theory (named after its proponents, G. C. Ghirardi, A. Rimini, and T. Weber (1986)) is a collapse theory, where the quantum state (also known as the wave function) of the system collapses down to a state that ensures that measurements have outcomes. (There have been various improvements made to the GRW theory; the resulting theories are collectively called Continuous Spontaneous Localization (CSL) theories. I'll stick with the GRW theory for simplicity, and because all the CSL theories have the same consequences for how to model God's actions.)

And those are all the major versions of quantum mechanics. Since we're interested in God's acts in an indeterministic quantum world, I'll focus on modal interpretations and the GRW theory. The results for God's acts would be the same on some versions of the modal interpretation as on the GRW theory, and modal interpretations are more complicated to understand, so let's focus on the GRW theory.

According to the GRW theory, the wave function for a N -particle system always has a chance of spontaneously undergoing a collapse. A collapse (called a "GRW hit") on a particular particle has an 1 in 10^{15} chance of happening per second. What happens when a system undergoes a GRW hit is that the wave function is multiplied by a localized Gaussian (bell-curve-shaped) function, with probability of the Gaussian being localized in any region being given by the value of the wave function in that region. (Technically, where the wave function is Ψ , the probability of the Gaussian being localized is a function of $|\Psi|^2$.) The GRW hit happens on the part of the wave function associated with a single particle, but for a typical macroscopic object, the positions of the particles are correlated which each other, so a GRW hit on one particle affects the state of the whole system. The net effect of a GRW hit is to localize the wave function for the N -particle system in the region where the GRW hit happened.

But on the GRW theory, wave functions are never completely localized—they always have tails that go to spatial infinity. (That wave functions have tails is true according to other versions of quantum mechanics too. For example, on the Copenhagen interpretation, a wave function may be localized to a particular region when the measurement collapse happens, but by the standard dynamical equation of quantum mechanics, Schrödinger's equation, the wave function will only be localized at an instant; after that instant it will once again have tails that

go to infinity.) By the dynamics of the GRW theory, the wave function can have value zero in some regions of space, but it has non-zero value in an unbounded region of space. Any place the wave function has non-zero value, there's a non-zero probability that the GRW hit will happen in that region. This issue of the wave function having tails that go to infinity will be key for the discussion of God's acts below. (If one is having trouble mentally picturing the tails, it helps to think about a Bell curve distribution—it's concentrated in one region, but there are tails that go to infinity.)

One complication I haven't yet introduced to the discussion is that the wave function is mathematically defined over $3N$ -dimensional space (where N is the number of particles in the system in question). Some philosophers of physics (such as Albert (1996) have argued that the wave function, evolving in the high-dimensional space, is what's real, and the experience we have of living in a three-dimensional space is illusory. I maintain that this is an incorrect way to understand the ontology of quantum mechanics (for argument see for example Monton (2006) and Monton (2012)). There's no need for quantum mechanics to be that radically revisionary, with respect to our common-sense understanding of the world as involving things evolving in three-dimensional space. And there is a way to make sense of that on the GRW theory. While there has been much debate about the ontology of the GRW theory, the standard view now is that the mass density ontology provides the (or at least: a reasonable) way of understanding the ontology of the GRW theory. (See Monton (2004) for the best defense of this ontology, and see also Ghirardi (2011).) On the mass density ontology, the wave function for a particle represents how the mass of the particle is spread out throughout space. Since wave functions have tails, the mass is spread throughout an unbounded region of space, but the vast majority of it is sometimes concentrated in a small region of space; when the mass is concentrated in this way, the particle behaves more classically. Macroscopic objects will almost always be such that the vast majority of the masses of their particles are concentrated in small regions of space, since there are very likely to be frequent GRW hits on particles in the macroscopic object, and these GRW hits are very likely to happen where wave function is already concentrated. Single particles, or small-particle systems, are very unlikely to undergo a GRW hit (recall that the chance of its happening, for a single particle, is just 1 in 10^{15} per second). This is how the GRW theory reproduces the results that we experience—we experience the world behaving classically at the macroscopic level, but quantum-mechanically at the microscopic level.

(For the cognoscenti: in addition to the wave function and mass density ontologies, the other main ontology for the GRW theory is the flash ontology. I find that ontology to be implausible (given that an individual particle is only in existence for an occasional instant of time), but none of the claims below about God's action hinge on whether the flash ontology or the mass density ontology is the correct ontology for the theory (though the presentation would be quite different on the different ontologies).)

3. How God Can Act Without Intervening

Recall that we're following the strictures of the proponents of noninterventionist special divine action: God doesn't intervene in the world (in that God doesn't violate the laws of nature), but God can act in the world (by for example determining the results of indeterministic physical processes). On the GRW theory, and some modal interpretations, God has an effective, wide-ranging means of doing this. I'll focus my argument on the GRW theory.

For those who skipped the previous section: the GRW theory is an indeterministic version of quantum mechanics that allows for indeterministic "GRW hits" to happen on the wave function of a particle, thus localizing the wave function. This means that a majority of the mass density of the particle is in a small region of space, but the wave function has tails that go to infinity, so the mass of the particle is also spread out throughout this infinite region of space. The GRW hit can happen anywhere that the wave function is non-zero, so the GRW hits can happen most anywhere in space, concentrating most of the mass density for the particle in that region where the GRW hit happened.

Now we come to the crucial argumentative move of my paper. Within the constraints of the laws of the GRW theory, God can make a GRW hit happen anywhere, on any particle, or collection of particles. This gives God the power to move particles around, anywhere in the universe. And moreover, God can do so arbitrarily quickly, just by making the GRW hits happen in an arbitrarily small amount of time.

(One slightly technical point: what if the wave function for a particle is zero in a particular region of the universe (for example, because another particle is there)? In that situation, God could do a GRW hit on the blocking particle to move it out of the way, thus changing the wave function of the particle God wants to move, and then God can do a GRW hit on the particle God wants to move, to get it in the desired location. That's how God can move particles anywhere.)

The standard (purported) miracles of God can all be accounted for in this way, or so I'll now argue. Let's start with an easy case: God parting the sea. Here, particles in the sea just need to shift a fixed distance to a new location. God could do this by simultaneously executing a GRW hit on the wave function for each of the particles, with the GRW hits all centered where God wants the sea to be. (It may take multiple GRW hits, but God could cause those to happen arbitrarily quickly.) Note that the sea parting is allowed by the laws of the GRW theory; it's just incredibly improbable that that would happen. (It's akin to the (empirically well-demonstrated) quantum tunneling phenomenon, where a particle suddenly appears in a new location, except instead of the single particle tunneling, every particle in the relevant portion of the sea.) If we were to witness the sea part in this way, we should recognize that no laws of nature need have been violated. (Whether such a sea-parting event would nevertheless provide evidence for the existence of God is an epistemological question outside the scope of this paper.)

Let's turn to more difficult cases of how God might act without intervening.

Alvin Plantinga talked with experts in quantum mechanics, and on that basis wrote the following:

According to the expert opinion to which I have had access, some of these (parting the Red Sea, miraculous healings) are pretty unproblematically compatible with QM. On other miracles, however—for example, raising someone from the dead, and transmutation, as with changing water into wine—there seems to be a substantial difference of opinion among the experts. Little analysis of these kinds of cases has been published; but some of the experts I’ve talked with (Katherine Brading, Craig Lent, Bas van Fraassen) think it implausible that QM be compatible with these miracles. Others, for example John Earman and Bradley Monton, think QM is compatible with them. (Plantinga (2008, p. 382), Plantinga (2011, pp. 94–95))

I’ll now argue that Earman and Monton are correct. Plantinga goes on to cite personal correspondence with Monton, discussing the case of changing water into wine:

you’ve got a bunch of individual particles (electrons, protons, etc.) that are composing the water, and they can all have GRW hits such that their positions are redistributed to the locations that would be appropriate for them to compose wine. (Monton, from Plantinga (2008, p. 382), Plantinga (2011, p. 95))

Why would Brading, Lent, and van Fraassen think otherwise? Plantinga doesn’t say, but the worry, as I understand it, is that the appropriate bonds wouldn’t be able to be established, so that the particles that make up the water actually form bonds into the chemical molecules that make up wine. I’ll now argue that this worry is unfounded.

Imagine that we start with a glass of water, which is fundamentally composed of various particles of a few different types: electrons, quarks, and so on. Imagine that a large number of GRW hits occur on the wave function of the water, in such a way that the particles are moved into a particular configuration of positions that they would be in if they constituted wine. (More precisely, the masses for the particles are highly localized in this configuration of positions.) Would this rearrangement of positions be enough for the liquid in the glass to be wine? The answer is “yes”: once the particles are in the appropriate positions, they will automatically have the appropriate bonds such that they have the chemical structure of wine. Here’s why.

As we currently understand it, there are four forces in nature, gravity, the electromagnetic force, and the weak and strong nuclear forces. The force of gravity acts between every particle with mass, so once the particles are in the positions they would be in if they constituted wine, the force of gravity will act between the particles just as if it had been a glass of wine from the start. The electromagnetic force is the same way; it is based on the positions of the particles. For example, an electron and a proton that are a certain distance apart from one another will feel a certain attractive force, even if the electron and proton spontaneously appeared in those positions just an moment before.

What about the strong and weak nuclear forces? The strong nuclear force (responsible for binding in the atomic nucleus) has a very short range—its effects aren’t experienced when particles are more than 10^{-15} meters apart. The range of

the weak nuclear force is even shorter—about 1/1000th the diameter of a proton. Thus, for both these forces, particles must be close together to feel its effects. But once particles are put close together, the effects will be felt. God's actions in the form of GRW hits will be enough to get the particles in the appropriate positions, and the fundamental forces (as governed by the laws of physics) then automatically come into play.

Thus, God can act by causing GRW hits in such a way that water is turned into wine. Raising someone from the dead is similar—the GRW hits move the particles around into the configuration of a living person, and as a result, the appropriate forces come into play to establish the appropriate chemical and atomic bonds. (With regard to the personal identity question, of whether this amount of physical and psychological connectedness is enough for the resurrected person to count as the same person as the person who died, my firm opinion is that it would be the same person, but to argue for that would take us too far afield.)

What about, starting from a few loaves of bread and a few fish, producing enough to feed 5000 people? This is an interestingly different case than the water-into-wine and resurrection cases, because it seemingly involves more particles coming into play that one started with. Indeed, Peter (Hodgson 2000, p. 514) claims that “the feeding of the five thousand is contrary to the law of the conservation of matter.” Where could the extra particles come from? A limitation of my view of how God can act in the quantum world is that God can't create new particles *ex nihilo*—that isn't allowed by the laws of the GRW theory. But, nothing stops God from bringing in other particles from far away, and using those particles to make fish and bread. For example, God could take particles making up nitrogen and oxygen high up in the atmosphere, or particles making up rocks on Mars, and do GRW hits on them in such a way that those particles become bread and wine here on the surface of the Earth. Nothing in the Biblical story of the fish and loaves of bread rules out God acting in that way.

I could give more examples, but you get the picture. All the standard purported miracles can be accounted for by God acting in the world via GRW hits, without violating the laws of nature. (The one obvious exception is a miracle that essentially involve a divine element, such the incarnation of God in the form of Jesus.)

Moreover, the GRW theory isn't the only version of quantum mechanics that yields these happy results of God being able to perform miracles in a noninterventionist way. Some modal interpretations, such as Bas van Fraassen's (1991) Copenhagen Variant of the Modal Interpretation (CVMI), have a highly stochastic dynamics for the value states. Given such a version of quantum mechanics, God could pick a dynamics for the value states that is compatible with the laws, in such a way that that dynamics yields the miracle God desires. (Note that when I say “is compatible with the laws”, that can be interpreted to mean “is allowed by a model of the theory”; van Fraassen famously does not believe in laws of nature (see for example van Fraassen (1989).)

4. God's Actions Need Not Be Episodic

After Plantinga quotes the personal correspondence with Monton, regarding the case of changing water into wine (discussed above), Plantinga writes:

Monton is speaking of the GRW approach to quantum mechanics; presumably a similar point would apply to the classical Copenhagen interpretation. (Plantinga (2008, p. 382), Plantinga (2011, p. 96))

Plantinga is mistaken here. On the Copenhagen interpretation, a quantum system evolves deterministically, until a measurement process occurs; only then is there a stochastic process in which God can act without violating the laws. Proponents of quantum special divine action typically present their theory using the Copenhagen interpretation, and so from that standpoint God can only act when a quantum measurement occurs.

This leads John Polkinghorne to raise the following important criticism of the standard, Copenhagen-interpretation-based view of quantum special divine action:

There is a particular difficulty in using quantum indeterminacy to describe divine action. . . . Occasions of measurement only occur from time to time and a God who acted through being their determinator would also only be acting from time to time. Such an episodic account of providential agency does not seem altogether satisfactory theologically. (Polkinghorne 1995, p. 152)

Polkinghorne holds that it is not satisfactory for God to be able to act only when measurements occur. But why is this not satisfactory? The answer I would give is: because then God does not have the freedom to generate a miracle whenever he wants (at least, he does not have that freedom, given that he does not want to deal in two different manners with his creation). But Polkinghorne gives a different answer. He writes, regarding measurement events:

their strictly episodic nature does not obviously fit them to describe agency, which must surely be assumed to have a more free-flowing character. (Polkinghorne 2002, pp. 188–189)

Polkinghorne doesn't give an argument here, but nevertheless I have some sympathies with his position. But first: what is his position? As I'll now explain, there are three plausible interpretations of his view. This is worth going into, not just for the sake of Polkinghorne exegesis, but because it will help us to explore to what extent my versions of quantum special divine action are better than the standard Copenhagen one.

First, Polkinghorne could mean that God's actions must take place over intervals of time. But this view strikes me as implausible—presumably God is capable of acting instantaneously. But perhaps there are situations where God *wants* to act over an interval of time. This leads to the second interpretation: Polkinghorne could mean that God should be free to act over an interval of time; God's agency shouldn't be constrained to instants. And finally, the third interpretation is to take "free-flowing" seriously: it's not just that God's actions should take place over an interval of time, but there should be some sort of tight connection between the actions at instants within that interval.

On the first interpretation, both the standard Copenhagen view of special

divine action and the GRW version fall prey to Polkinghorne's critique: the collapse of the wave function as a result of measurement happens at an instant, as do GRW hits. But on the second interpretation, that God should be free to act over an interval of time. Polkinghorne's critique only impacts the Copenhagen view. On the Copenhagen view, God's actions are limited to the instants where wave function collapse occurs as a result of measurement. But on my view, God can act at every instant in an interval of time, by causing GRW hits at each of those instants. (Or, under the assumption that a modal interpretation like van Fraassen's is true, God could pick a dynamics for the value states over some interval of time that is compatible with the laws, in such a way that that dynamics yields the result God desires.)

It's unclear to me exactly how to spell out the third interpretation, that there should be a tight connection between God's actions at instants within the interval, but I see that there is something to it. On the third interpretation, the standard Copenhagen view of special divine action is definitely maligned, because on the Copenhagen view God just acts at an instant. On view that God acts via GRW hits, I see the concern that each GRW hit is its own event, without there being a free-flowing connection between the different GRW hits. Nevertheless, if the different GRW hits over some interval of time are all for some cohesive end (turning water into wine, for example), then what is free-flowing is the process of the water gradually turning into wine, as the GRW hits happen. (Given modal interpretations like van Fraassen's, God's actions could be more free-flowing in a natural way, by having God control the dynamics for the value states over some interval of time.)

In sum, Polkinghorne's critique is open to interpretation, but regardless of how one interprets it, the standard Copenhagen view of quantum divine action falls prey to the critique, while on some reasonable interpretations my versions of quantum special divine action do not.

Robert Russell attempts to argue that, in fact, his Copenhagen view of quantum divine action does not fall prey to Polkinghorne's critique. If Russell is right about this, then that removes an important reason to favor my versions of quantum special divine action over Russell's. But as I'll now show, Russell is not right about this.

Sometimes, Russell concedes the point, to an extent at least. For example, he writes: "It is ... true that indeterminism arises only during an irreversible interaction with more complex objects" (Russell 2003, p. 362). I hold that this is already granting Polkinghorne's critique: God cannot act at any time he wants, but only when there is an "irreversible" (more on that below) interaction with, not just any object, but a more complex object. But after the above quoted sentence, Russell goes on to say:

My point, however, is that these interactions are not limited to physical measurements in the lab; instead they occur throughout the universe wherever elementary particles are irreversibly absorbed by objects ranging from complex molecules and interstellar dust to those of the ordinary macroscopic world. To me this suggests a God who acts

throughout innumerable occasions in the universe, and thus a much more comprehensive view of divine action than the term “episodic” suggests. (Russell 2003, p. 362)

I’m not convinced by Russell’s defense of nonepisodicness here. For a measurement to occur, there has to be the sort of interaction Russell specifies, where an object interacts with a more complex objects. When such an interaction does not occur, God cannot act in the world (subject to the assumption we’re making, regarding God’s not wanting to violate the laws). This strikes me as a clear limitation on God’s actions. Moreover, this is a limitation that the GRW version of quantum special divine action doesn’t share: a GRW hit can happen at any time, anywhere that the wave function is non-zero, so God is always free to cause a GRW hit to occur.

Russell puts the point in a slightly different way in another essay. He writes: What about the “episodic” nature of such interactions? In fact, such interactions can occur at any time and place in the universe where the deterministic time-development of the quantum phenomena governed by the Schrödinger equation is disrupted by an irreversible interaction (measurement)... (Russell (2002, p. 310); Russell (2009, p. 375))

There’s an ambiguity here that I want to resolve. One might think that the laws of quantum mechanics determine which interactions are irreversible, but that’s not the case. In principle, all interactions could be governed by the Schrödinger equation, and as long as that’s the case, all interactions are reversible. The problem with the view that all interactions are governed by the Schrödinger equation is that, on that view, measurements typically don’t have results, so something about quantum mechanics needs to be modified, to ensure that quantum mechanics captures our experiences of measurements having results. This is why the Copenhagen interpretation specifies that, when a measurement happens, the wave function collapses—this collapse process is an irreversible process. The main problem with the Copenhagen interpretation is that it doesn’t give a physical specification of what it takes for a measurement process to happen. Russell is giving a partial characterization of what it takes for a measurement to happen (in terms of interaction with a more complex system), but ideally one would want a more precise characterization (and I’m confident that Russell would agree). The key point is that, according to the Russell’s Copenhagen version of special divine action, God can only act when this measurement process occurs—and, as Russell points out, such measurement processes occur only when an object interacts with a more complex object. Absent such an interaction, God cannot act. I conclude that God’s actions on Russell’s view are episodic, and thus fall prey to Polkinghorne’s critique. God’s actions according to my versions of quantum special divine action need not be episodic, and this is an important reason to prefer my versions over Russell’s.

5. Saunders’ Critiques

But are there other reasons to be unhappy with my versions of quantum special divine action? As far as I can tell, modal interpretations have never been discussed in this literature on quantum special divine action, so no critiques of

those interpretations have been raised in this context. The GRW theory, though, has been discussed in this context, and Nicholas (Saunders 2002, pp. 156–159) has argued against the GRW theory as a way of understanding quantum special divine action. I'll now show that Saunders' critiques are ill-founded.

The first of Saunders's arguments against the GRW theory which I'll focus on is that the GRW theory appears basically contrived:

for a microscopic quantum system with a small number of particles N , the multiplication of the Gaussian [i.e., the GRW hit] becomes so infrequent as to be practically undetectable; moreover the width of the Gaussian appears to be chosen to be sufficiently big to ensure that any energy conservation violations which arise as a result of its multiplication would be very small and are thus experimentally undetectable. (Saunders 2002, pp. 158–159)

It's true that the two new fundamental constants of the GRW theory (the probability of a GRW hit happening, and the width of the Gaussian) have been picked to fit the extant data, but there's nothing wrong with that. It's a typical process in science that sometimes one uses empirical evidence to determine how to formulate one's theory; if the process of theory-formation didn't rely on empirical evidence there would be something horribly misguided about the process. So I wouldn't call the GRW theory "contrived"; it was formulated under empirical constraints, which is a standard part of science.

It is true that the GRW theory violates the principle that energy is conserved, but that principle isn't sacrosanct; according to the GRW theory that principle is simply false. (For those who think giving up that principle is egregious, note that it's not even clear how to formulate the principle of conservation of energy in the context of general relativity. It's not at all clear that that is a true principle of physics.)

The second argument of Saunders against the GRW theory (and other Continuous Spontaneous Localization theories) is that they "bypass rather than fully address the problem of measurement in quantum mechanics", because "the highly localized state which results from Gaussian multiplication is *not the same as a completely localized state*". Saunders goes on to explain the problem of wave function tails—that wave functions aren't highly localized; they go out to infinity. But it's this feature of the GRW theory that I'm exploiting to explain how God is so free to act in the quantum world: God can do GRW hits to move particles anywhere. Moreover, there's now a well-worked out and agreed-upon ontology for the GRW theory which accommodates these wave function tails, the mass density ontology (Monton 2004, Ghirardi 2011).

The third and final argument of Saunders focuses on God's actions given the ontology of the GRW theory. But the argument is only plausible because Saunders has a different view of how God would act than I do. Saunders writes: under this connection between quantum special divine action and the [GRW] approach to measurement, God still does no more than determine when interventionistically to toss the quantum dice' and get a probabilistic result. The theologian is still left with the burden of describing how it is that God might obtain a

purposive result under this scheme... (Saunders 2002, p. 159)

As I interpret him, Saunders is saying that on the GRW-based account of quantum special divine action, God determines *when* a GRW hit happens, but not *where* it happens. If that were the account, then indeed, God wouldn't be able to control the evolutions of systems, and hence God wouldn't be able to get purposive results. But as a result, that account is a non-starter. On my GRW-based account of quantum special divine action, God determines when and where GRW hits happen, and that is how God has the almost unlimited freedom to bring about any effect in the physical world.

6. Two Concluding Thoughts

I'll close with two concluding thoughts. First, this ability of God to act in the physical world, via quantum effects, comes with a price. Specifically, God *has* to act in the world, or at least be ready to act in the world, to ensure that it does not go vastly awry. God cannot be a hands-off God who starts the universe in motion but then no longer intervenes—or at least, God can't ensure from the start that he will be such a God, under the assumptions that God is unwilling to let the world go certain horrible ways, and that God does not determine the outcomes of indeterministic processes at the moment of creation. Why? Because, for example, it's possible that, if nature is left to its own devices, in the next second GRW hits happen in such a way that all humans cease to exist. This is incredibly unlikely, but it's compatible with the GRW theory that that happens (because the GRW hits could happen in such a way that each human's particles are scattered as if there were an explosion, leading to at least the physical death of that human). God presumably wouldn't want that to happen, and that's why he has to be ready to act. (Note that I haven't taken up the debate over whether God is the cause of all quantum indeterministic events or just some; my point here is that God has to be willing to be the cause of at least some.)

Here's my second concluding thought. I have for the most part been talking as if the GRW theory could be a true theory of the world, but in fact, it (like every other version of quantum mechanics) is probably a false theory. As I noted at the start, quantum mechanics cannot accommodate the evidence that supports general relativity, and that is one reason physicists are searching for a theory of quantum gravity, which will supplant both quantum mechanics and general relativity. It could turn out that some theory of quantum gravity is the true theory of the world, or it could turn out that after coming up with that theory, we will realize that we still have further to go. What I've defended in this paper is a model for how noninterventionist special divine action could work. I've defended the details of the GRW-based model, not to argue for the truth of the GRW theory, but to show that there is a plausible extant model for how noninterventionist special divine action could happen. (I've also suggested that some modal interpretations provide another plausible model.) We will have to see whether future theories of physics have the level of indeterminism built in to them such that my model for noninterventionist special divine action still

applies. Given the current state of physics, we just don't know, and it's hard to even make a well-informed guess.

Polkinghorne makes a similar point:

In our present state of knowledge, no proposal relating to the conceivable causal joint of divine providential interaction can be more than what the physicists would call a "zero-order approximation," a crude starting point from which one may hope that better developments might eventually spring." (Polkinghorne 2002, p. 190)

Philosophers sometimes aren't happy with such inconclusive results, and there has been a trend of metaphysicians appealing to physics as if physics is giving us conclusive answers regarding the nature of the world. (For a critique of such metaphysicians, see Monton (2011).) If I wanted to be such a metaphysician, I would present the GRW theory as if it were a plausibly true theory of the world, and conclude that the problem of noninterventionist special divine action has been solved. But our current physics is far from conclusive, and as a result, our physics-based hypotheses regarding how God could act in the world must be far from conclusive as well.

REFERENCES

- Albert, David. 1996. "Elementary Quantum Metaphysics." pp. . 277–284. Dordrecht: Kluwer.
- Bub, Jeffrey. 1997. *Interpreting the Quantum World*. Cambridge: Cambridge University Press.
- Ghirardi, GianCarlo. 2011. "Collapse Theories." *The Stanford Encyclopedia of Philosophy*. [Http://plato.stanford.edu/archives/win2011/entries/qm-collapse/](http://plato.stanford.edu/archives/win2011/entries/qm-collapse/).
- Ghirardi, Giancarlo, Alberto Rimini, and Tullio Weber. 1986. "Unified Dynamics for Microscopic and Macroscopic Systems." *Physical Review D* 34: 470–491.
- Hodgson, Peter E. 2000. "God's Action in the World: The Relevance of Quantum Mechanics." *Zygon* 35: 505–516.
- McMullin, Ernan. 1993. "Evolution and Special Creation." *Zygon* 28: 299–335.
- Monton, Bradley. 2004. "The Problem of Ontology for Spontaneous Collapse Theories." *Studies in History and Philosophy of Modern Physics* 35: 407–421.
- Monton, Bradley. 2006. "Quantum Mechanics and 3N-Dimensional Space." *Philosophy of Science* 73: 778–789.
- Monton, Bradley. 2010. "Against Multiverse Theodicies." *Philo* 13: 113–135.
- Monton, Bradley. 2011. "Prolegomena to Any Future Physics-Based Metaphysics." pp. . 142–165. Oxford: Oxford University Press.
- Monton, Bradley. 2012. "Against 3N-Dimensional Space." In *The Wave Function: Essays in the Metaphysics of Quantum Mechanics*, edited by David Albert and Alyssa Ney. Oxford: Oxford University Press.
- Murphy, Nancey. 1995. "Divine Action in the Natural Order: Buridan's Ass and Schrödinger's Cat." pp. . 325–356. Vatican City: Vatican Observatory Publications and Berkeley: Center for Theology and the Natural Sciences.

- Plantinga, Alvin. 2008. "What is 'Intervention'?" *Theology and Science* 6: 369–401.
- Plantinga, Alvin. 2011. *Where the Conflict Really Lies: Science, Religion, & Naturalism*. Oxford: Oxford University Press.
- Polkinghorne, John. 1995. "The Metaphysics of Divine Action." pp. . 147–156. Vatican City: Vatican Observatory Publications and Berkeley, Center for Theology and the Natural Sciences.
- Polkinghorne, John. 2002. "Physical Process, Quantum Events, and Divine Agency." pp. . 181–190. Vatican City: Vatican Observatory Publications and Berkeley: Center for Theology and the Natural Sciences.
- Russell, Robert John. 2002. "Divine Action and Quantum Mechanics: A Fresh Assessment." pp. . 293–328. Vatican City: Vatican Observatory Publications and Berkeley: Center for Theology and the Natural Sciences.
- Russell, Robert John. 2003. "Special Providence and Genetic Mutation: A New Defense of Theistic Evolution." pp. . 335–369. Grand Rapids: William B Eerdmans Publishing.
- Russell, Robert John. 2009. "Divine Action and Quantum Mechanics: A Fresh Assessment." pp. . 351–404. Leiden: Brill.
- Saunders, Nicholas. 2002. *Divine Action and Modern Science*. Cambridge: Cambridge University Press.
- Tracy, Thomas. 2002. "Creation, Providence, and Quantum Chance." pp. . 259–292. Vatican City: Vatican Observatory Publications and Berkeley: Center for Theology and the Natural Sciences.
- van Fraassen, Bas. 1989. *Quantum Mechanics: An Empiricist View*. Oxford: Oxford University Press.
- van Fraassen, Bas. 1991. *Laws and Symmetry*. Oxford: Oxford University Press.

NO (NEW) TROUBLES WITH OCKHAMISM

Garrett Pendergraft and D. Justin Coates

Abstract: The Ockhamist claims that our ability to do otherwise is not endangered by God’s foreknowledge because facts about God’s past beliefs regarding future contingents are soft facts about the past—i.e., temporally relational facts that depend in some sense on what happens in the future. But if our freedom, given God’s foreknowledge, requires altering some fact about the past that is clearly a hard fact, then Ockhamism fails *even if* facts about God’s past beliefs are soft. Recent opponents of Ockhamism, including David Widerker and Peter van Inwagen, have argued along precisely these lines. Their arguments, if successful, would undermine Ockhamism while avoiding the controversy over the alleged softness of facts about God’s past beliefs. But these arguments do not succeed. The past facts they rely on must be clear and uncontroversial examples of hard facts about the past, and these facts must be such that an ability to refrain from the relevant future action implies an ability to alter the relevant hard fact. We demonstrate the flaw in these arguments by showing how they rely on past facts that do not satisfy these criteria. The Ockhamist may have troubles, but this type of argument is not one of them.

§1 Introduction

There is a well-known tension between divine foreknowledge and the freedom to do otherwise—a tension that has led many philosophers of religion to declare the two incompatible. Typically, the dialectic between these theological incompatibilists and their compatibilist opponents begins with some version of the standard argument for incompatibilism, which was first formalized by Nelson Pike in 1965.¹ According to the standard argument, if it is true that Jones (for example) will mow his lawn tomorrow, then it is a fact about the past that God

¹Pike (1965). Pike’s argument has been poked, prodded, and tweaked by a host of writers. William Hasker’s reconstruction, as found in his Hasker (1989) is particularly good (although we have changed the characters somewhat):

- (1) It is now true that Jones will mow his lawn tomorrow. (premise)
- (2) It is impossible that God should at any time believe what is false, or fail to believe anything that is true. (premise: divine omniscience)
- (3) Therefore, God has always believed that Jones will mow his lawn tomorrow. (1, 2)
- (4) If God has always believed a certain thing, it is not in anyone’s power to bring it about that God has not always believed that thing. (premise: the fixity of the past)
- (5) Therefore, it is not in Jones’s power to bring it about that God has not always believed that he would mow the lawn. (3, 4)

believed that Jones would mow his lawn tomorrow.² Thus, if Jones were to refrain from mowing his lawn, then he would have to change that fact about the past. But the past is fixed, which means that nobody can change it; so Jones is not free to refrain from mowing his lawn. According to the Ockhamist, this traditional incompatibilist argument fails because God's beliefs are "soft facts" about the past, and thus in some sense alterable.³

David Widerker has challenged this Ockhamist move,⁴ offering a "general objection to any sort of Ockhamist attempt to reconcile divine foreknowledge with human freedom by treating facts about God's foreknowledge of future contingent events as soft facts about the past over which agents may have power."⁵ Widerker presents a scenario in which God's past knowledge of some event that will occur in the future sets off a "future-contingent causal chain" leading up to the occurrence of the event. Each of the events in this chain is connected to the next event in virtue of being a causally necessary condition for its occurrence. Moreover, most of the past events in this future-contingent causal chain appear to constitute hard facts about the past. And if such facts truly are hard facts, then this is a problem for the Ockhamist—because it implies that an agent has the ability to alter the fact that God knows that a future event will occur only if she has an ability to alter a hard fact about the past (e.g., one of the events that together constitute the future-contingent causal chain). Whether or not some of God's beliefs are in some sense alterable, it seems that no agent has the ability to alter a *hard* fact about the past.

We take this argument to de-emphasize the question of whether God's beliefs are soft, and focus instead on past facts that are more widely accepted as hard. For if our ability to do otherwise requires altering a fact about the past that is clearly a hard fact, then Ockhamism fails *even if* God's past beliefs are soft. In this paper, we claim that Widerker's argument from future-contingent causal chains, thus construed, is unsuccessful. More generally, our claim is that the incompatibilist who takes Widerker's line cannot avoid the debates over the softness of God's beliefs after all, and hence has not succeeded in providing a general objection to the Ockhamist move.

As we attempt to establish this thesis, we will first say a little more (in §2) about the distinction between hard and soft facts. We will then consider (§3) Widerker's argument. We think that his argument falls prey to a dilemma,

(6) It is not possible for it to be true both that God has always believed that Jones would mow the lawn, and that he does not in fact mow the lawn. (from 2)

(7) Therefore, it is not in Jones's power to refrain from mowing the lawn tomorrow. (5, 6) So Jones's mowing the lawn tomorrow is not an act of free choice.

²Of course, from God's perspective at the relevant time in the past, the content of his belief would be future-tensed: "Jones will mow his lawn tomorrow."

³Although it is more precise to say that "facts about God's beliefs" are soft (or hard) facts, we will occasionally just say that God's beliefs themselves are soft (or hard) facts.

⁴Widerker (1990).

⁵Widerker (1990, pp. 474–475).

which we develop in §4. We then turn (§5) to a related argument from Peter van Inwagen, and criticize this argument (§6) by applying the dilemma developed in response to Widerker.⁶ In §7 we tie up some loose ends regarding the connection between soft facthood and entailment. We conclude (§8) by offering a summary of the dialectic.

§2 The distinction between hard and soft facts

The Ockhamist contends that our ability to do otherwise is not endangered by God's foreknowledge because facts about God's past beliefs regarding future contingents are soft facts about the past. Two assumptions should help to clarify this distinction between hard and soft facts. First, hard facts are *temporally non-relational*, whereas soft facts are *temporally relational*. A fact is temporally relational when it consists of two other facts—each of which is about a different time—and a relation between those constituent facts. (Informally, we might say that a soft fact has two “parts,” whereas a hard fact has only one part.⁷)

The following, then, are hard facts about the past (relative to the writing of this paper in 2010):

- (1) King John signed the Magna Carta in 1215.
- (2) Barack Obama was inaugurated as President of the United States on January 20th, 2009.

Contrast (1) and (2) with

- (3) King John signed the Magna Carta 794 years before Barack Obama was inaugurated as President.

This fact (3) is temporally relational because it comprises the fact that King John signed the Magna Carta (which is about 1215), the fact that Barack Obama was inaugurated as President (which is about 2009), and the “before” relation between the first fact and the second fact.

The second clarifying assumption is that whereas all hard facts are unalterable, some—but not all—soft facts are alterable. (Or we can put the point in terms of fixity: Whereas all hard facts are fixed, soft facts may or may not be fixed.) There are at least two reasons why a soft fact might be fixed. The first reason is that both of the constituent facts might be about the past. On January 1st, 2009, someone might be able to do something such that (3) would not be a fact; but on January 21st, that particular soft fact is no longer alterable. The second reason why some facts are soft but nonetheless fixed is that past facts can be temporally related to future events over which nobody has any power.⁸ For example, the fact that Barack Obama was inaugurated on the

⁶van Inwagen (2007).

⁷We might also put it this way: soft facts, while *genuinely* about the past, are not *solely* about the past.

⁸Fischer (1997) , emphasizes the importance of temporal relationality in drawing the distinction between hard and soft facts. In Fischer's language (chapter 6), soft facts that are nevertheless fixed are “hard-core soft facts” or “hard-type soft facts.” In this paper we will not consider the question of whether any of the putatively soft facts discussed below might be hard-core or hard-type soft facts.

morning of January 20th, 2009, prior to the sun's rising on January 21st, is a soft fact relative to the evening of January 20th—but it is also fixed, because no one at that time has any power over whether the sun rises on January 21st.⁹ Thus, if we want to establish that some fact is alterable, we must show that it is temporally relational (i.e., soft), and that one of the constituent facts is about a future event, and moreover that someone has power over that future event. It is therefore open to the opponent of Ockhamism to claim that theological incompatibilism is true because God's beliefs, though temporally relational, are nevertheless fixed.¹⁰ This, however, is not the line that Widerker takes, so we will not attempt to respond to it here.¹¹

It has proven exceedingly difficult to provide a satisfying general account of the distinction between hard and soft facts.¹² Nevertheless, Ockhamists often claim that God's beliefs should turn out soft on any satisfactory analysis of the distinction.¹³ Thus the Ockhamist would maintain that the following fact:

(4) God believed that Jones would mow his lawn tomorrow

is a soft fact about the past, relative to today. Although it isn't immediately obvious that (4) is temporally relational, we can recast it in a way that reveals its softness. Given the assumption that God is essentially omniscient, (4) is logically equivalent to

(5) God *correctly* believed that Jones would mow his lawn tomorrow.

And now we can say that (5) consists of the fact that God believed that Jones would mow his lawn tomorrow (which is about the past), the fact that Jones will mow his lawn tomorrow (which is about the future), and the relation of correspondence between those two facts. Thus, (5) is soft and—on the further assumption that soft facthood is closed under logical equivalence—so is (4).¹⁴

Ockhamists also claim that in addition to the temporal relation between God's past beliefs about our future free actions and the actions themselves, there is also a relation of *counterfactual dependence*: we sometimes have the ability to

⁹Strictly speaking, of course, this isn't true: for presumably God, on the evening of January 20th, had power over whether the sun rose on January 21st. But since the debate were concerned with is the debate over whether *human* freedom is compatible with divine foreknowledge, we will restrict ourselves to the domain of human agents.

¹⁰Fischer (1997, chapter 6) does take roughly this line, arguing that even though facts involving God's past beliefs are soft facts, certain elements of their internal structure are hard (temporally non-relational), which means that those facts, though soft, are nonetheless fixed.

¹¹Widerker discusses (and rejects) this approach in §V of his 1990.

¹²Valiant attempts have been made, however, starting with Adams (1989); and continuing with (among others) Hoffman and Rosenkrantz (1989); and Freddoso (1989).

¹³This is roughly the position that Alvin Plantinga takes in his Plantinga (1989).

¹⁴Thanks to Carl Ginet for suggesting (in personal correspondence) this construal of facts such as (4). Ginet also proposes that the facts constituting a soft fact must themselves be *hard* facts—but we don't see any reason to rule out the possibility that a soft fact has another soft fact as a constituent. For example, the following seems to be a soft fact: "God believed that Jones would mow his lawn tomorrow, 794 years before Barack Obama was inaugurated as President." The first constituent fact is about two times (which might complicate the analysis), but it remains true that the two constituent facts are about different times.

do otherwise, and if we were to do otherwise, God would not have held a belief that he actually held. Ockhamists acknowledge that an ability to falsify a hard fact would amount to an ability to change the past; but they also maintain that falsifying a soft fact does not require such extravagant power. If it is in one's power to falsify one of the parts of a soft fact, then it is in one's power to falsify the soft fact itself.

The most obvious point of attack for opponents of Ockhamism is the claim that facts about God's beliefs are alterable soft facts about the past. Widerker does indeed attack this claim, arguing that it is not enough merely to point out superficial similarities between a temporally relational fact about God's past beliefs, such as (4), and temporally relational facts such as (3). The Ockhamist must also identify a property that these two facts share, and moreover "he must give us a good reason to think that it is *in virtue* of having this property or feature" that a fact like (4) is alterable.¹⁵ Widerker considers several attempts to satisfy this requirement, but finds them wanting. He recognizes, however, that there is at least one way for the Ockhamist to fend off his attack (albeit a way he considers implausible), and so, in order to fortify his position, he also develops a more general objection to Ockhamism: the argument from future-contingent causal chains.¹⁶

§3 Widerker's argument from future-contingent causal chains

Widerker begins by considering a scenario in which God's past knowledge (and hence, we will assume, his past belief) regarding some future free action is, given the circumstances, a causally necessary condition for the occurrence of some other event—an event that occurs after God's past belief, but prior to the future action. For example, God's past belief that Jones would mow his lawn tomorrow might prompt God to warn Smith prior to Jones's mowing his lawn. (Imagine that Smith's new puppy—an incredibly sound sleeper—is having a snooze in the tall grasses of Jones's backyard.) In this scenario, God's past belief that Jones would mow his lawn is a causally necessary condition for his warning Smith prior to Jones's mowing. Thus, on the Ockhamist supposition that Jones is able to refrain from mowing his lawn, it would appear that Jones is able to prevent the occurrence of a causally necessary condition for God's warning Smith (i.e., God's belief that Jones would mow his lawn). And since it is plausible to think that if Jones has such control over a causally necessary condition of some event, then he also has control over the event itself, we can conclude from Jones's ability

¹⁵Widerker (1990, p. 465 (emphasis in original)).

¹⁶Widerker (1990, p. 474). We should note that the Ockhamist response (which Widerker thinks is available but implausible) suggests the eternity (or atemporal) solution to the problem of divine foreknowledge and human freedom. He criticizes the eternity solution in his Widerker (1991). His argument against the eternity solution is structurally the same as his argument against the Ockhamist solution, and we think our criticisms apply to his argument against the eternity solution with equal force. But we will not develop that point in any detail here.

to do otherwise (and hence his ability to affect God's belief in some way) that he is also able to prevent the occurrence of

(W1) God warned Smith that Jones would mow his lawn tomorrow (and thus that his puppy was in grave danger) prior to Jones's mowing his lawn tomorrow.

Moreover, since (W1) will serve as a causally necessary condition for various subsequent facts, it would appear that Jones is also able to prevent the occurrence of facts such as

(W2) Smith heard a voice telling him that (Jones would mow his lawn tomorrow and thus that) his puppy was in grave danger

and

(W3) Smith came to believe that his puppy was in grave danger.

In short, Widerker argues that if Jones is able to do otherwise, then he is able to falsify past facts like (W1)–(W3).¹⁷ But (W1)–(W3) are hard facts about the past, and nobody can alter hard facts. So Jones must not be able to refrain from mowing his lawn after all. Therefore we cannot have power over God's past beliefs, since such power would imply that we could, in some circumstances, have the power to falsify some hard facts about the past. The Ockhamist appears to be on the ropes.

Despite the apparent force of this argument, we contend that the Ockhamist need not be overly concerned. In particular, we contend that Widerker faces a dilemma—which arises because the past facts (W1)–(W3) are not the uncontroversially hard facts that they need to be in order for his anti-Ockhamist argument to do its work. Before we elaborate on that claim, however, we need to say a little more about hard and soft facts.

Recall that if a fact about the past is soft, then it is *temporally relational*. So, for example, the fact that it was true in 1900 that we would write a paper on Ockhamism in 2010, if a fact,¹⁸ is temporally relational—i.e., it is a soft fact about the past (relative to January 1st, 2010). Had we failed to write this paper in 2010, then that soft fact about 1900 would not have been a fact. And because we had the ability to falsify the temporal relatum, “we write a paper on Ockhamism in 2010,” we had the ability to falsify the soft fact itself. On the other hand, the fact that the competition for the Davis Cup was established in 1900 is temporally non-relational (i.e., hard) and hence fixed. Nobody can in 2010 render that fact false.

¹⁷In addition to being able to render (W1)–(W3) false, we have good reason to suspect that Jones also has the ability to render (W4)...(Wn) false as well, where (W4)–(Wn) are putatively hard facts about some time between God's belief and Jones's mowing.

¹⁸We say “if a fact” because some have expressed doubts about there being such a thing as truth at a time. Pike (1965, p. 36), for example, says: “I share the misgivings of those contemporary philosophers who have wondered what (if any) sense can be attached to a statement of the form ‘It was true at T1 that E would occur at T2.’” Peter van Inwagen also discusses (and expresses skepticism about) the notion of truth at a time in his van Inwagen (1983). We will not consider such worries here, although a complete defense of Ockhamism would clearly require addressing them.

We will now recast the dispute between the incompatibilist and the Ockhamist in a way that will clarify the dilemma that we present below. First consider some putatively free action X . If some agent S is able at some time t_2 to do X at a later time t_3 , then the following ability claim (or “can-claim”) will be true:

(A) S , at t_2 , can do X at t_3 .

Suppose that (A) is true. Now take some fact F about the past. The relevant question—the one to which incompatibilists and Ockhamists will provide conflicting answers—is whether the truth of (A) is consistent with the truth of a backtracking counterfactual involving both X and F . The backtracking counterfactual (or just “backtracker”) will look like this:

(B) If S had done X at t_3 , then F , which obtained at t_1 , could not have obtained at t_1 .

Both parties to this dispute will agree that when F is a hard fact about t_1 , the relevant can-claim and the paired backtracker cannot both be true; in particular, they will agree that if the backtracker (B) is true (when F is a hard fact), then the can-claim (A) must be false.¹⁹ What they will not agree on is what happens when we substitute one of God’s past beliefs for F . The theological incompatibilist will claim that substituting one of God’s past beliefs for F changes nothing: F is still a hard fact, and (A) and (B) are still inconsistent. The Ockhamist, however, claims that when the fact F in the relevant counterfactual is (a fact about) one of God’s past beliefs, then that fact is both soft (temporally relational) and counterfactually dependent (on the event to which it is temporally related). Moreover, the Ockhamist maintains that there is no inconsistency in claiming that both (A) and (B) are true. If this is right, then certain of God’s beliefs are such that S can act in a certain way, and if she acts in that way, it would then be the case that God would not have held a belief that he actually held.

Once we construe the dialectic in this way, we can see that what the incompatibilist needs to show is that (A) and (B) are indeed incompatible (where the X stands for some putatively free action that is allegedly foreknown by God). One way to do this, of course, is to argue directly that (facts about) God’s past beliefs, Ockhamist claims notwithstanding, are hard facts after all. The other way to do this is to argue *indirectly*—by substituting for F a less controversial hard fact, and so sidestepping the question of whether God’s beliefs are soft facts

¹⁹We are construing the Ockhamist as acknowledging that nobody can so act that a hard fact would not have been a fact, but maintaining that there are some soft facts such that someone can so act that a soft fact would not have been a fact. In other words, we are following Widerker (who is himself following Fischer) in construing the set of past facts that are alterable as a subset of the past facts that are soft facts. Someone could question this framework, however, and argue that the Ockhamist should *not* agree that nobody can so act that a hard fact would not have been a fact. On this view, the question of whether someone can so act that a certain fact would not have been a fact is unrelated (or at least orthogonal) to the question of whether that fact is hard or soft. Nonetheless, since it is at the very least an open question as to which framework better captures the Ockhamist commitment, we will continue to construe the Ockhamist as denying that hard facts are alterable. (Thanks to Patrick Todd for emphasizing this point to us in personal conversation.)

(and whether, even if they are, that allows the Ockhamist to affirm the compatibility of (A) and (B)). Widerker's argument can plausibly be construed as an implementation of this indirect strategy because he substitutes for F facts about the past, such as (W1), that do not explicitly involve any of God's beliefs. (For the sake of concision, in what follows we will refer to the indirect anti-Ockhamist strategy implemented by Widerker [and, as we will see below, van Inwagen] as the "indirect strategy," or the "indirect argument.") In order for such a strategy to represent dialectical progress, the relevant fact F about the past must be a clear case of a hard fact—for only when F is indisputably a hard fact is it clear that if the backtracker (B) is true, then the can-claim (A) must be false. Otherwise, it is open to the Ockhamist to maintain that the F in question is alterable—i.e., that F is relevantly similar to a fact about God's beliefs insofar as it exhibits both temporal relationality and counterfactual dependence.

§4 A dilemma for Widerker

In the previous section we identified two desiderata that must be satisfied by proponents of the indirect strategy if they are to succeed in refuting the Ockhamist. First, they must avoid scenarios in which the fact that is supposed to call into question some future free action is a fact about God's past beliefs. (Embracing such scenarios is not necessarily a hopeless strategy; it is just a different, more direct strategy.) Second, they must provide a clear case of a hard (i.e., temporally non-relational) fact that calls into question the freedom of some future action. If this can be done, then the relevant backtracking counterfactual will have a hard fact in the consequent, implying that it can be plausibly considered inconsistent with its corresponding can-claim. Widerker's argument against the Ockhamist avoids explicitly appealing to facts about God's past beliefs to challenge our freedom to do otherwise. However, as we will now argue, the past facts he *does* appeal to do not threaten our freedom either. As a result, his argument should not unsettle the committed Ockhamist.

Consider an abridged version of (W1):

(W1) God warned Smith that his puppy was in grave danger prior to Jones's mowing his lawn tomorrow.

Since the most salient feature of the indirect strategy is that it avoids engaging the debate over whether or not facts about God's past states or activities (in particular, his past beliefs) are soft, it seems inappropriate to claim that a different sort of divine activity—warning Smith, as in (W1)—can be part of an uncontroversially hard fact about the past. In support of this point, compare (W1) with (4):

(4) God believed that Jones would mow his lawn tomorrow.

One plausible reason why God's past beliefs are taken to be soft is that those beliefs *entail*, and are thus necessarily connected to, the occurrence of the future event that they are about. The problems that beset a fully general entailment criterion (according to which a fact is soft if and only if it entails

some fact about the future) are well-documented.²⁰ But it remains an open question whether some *restricted* entailment criterion—such as the one proposed by Plantinga²¹—will turn out to be viable. Thus, for now at least, the Ockhamist can follow Plantinga in claiming that nothing that entails that Jones will mow his lawn tomorrow can be a hard fact about the past.²² And since (W1) also entails that Jones will mow his lawn tomorrow,²³ the Ockhamist can plausibly say that the warning in (W1) is no less a soft fact than the divine belief in (4) that prompts the warning. More generally, it seems plausible to claim that any past fact about t_2 that entails that Jones mows his lawn at some future time t_3 is at least a candidate for soft facthood about t_2 .

Above we construed the Ockhamist as arguing that Jones can refrain from mowing his lawn at t_3 , and maintaining that this can-claim is consistent with the claim that if Jones were to refrain from mowing his lawn at t_3 , then God would not have believed that Jones would mow his lawn at t_3 (i.e., consistent with the claim that God's past belief is counterfactually dependent on Jones's mowing his lawn). Given the crucial similarity between (4) and (W1), the Ockhamist can arguably maintain that if Jones were to refrain from mowing his lawn at t_3 then God's warning to Smith would not have been issued. Thus, concludes the Ockhamist, the warning in (W1) does not endanger Jones's freedom.

Perhaps the real mistake here is the supposition that the truth of (W1) guarantees the occurrence of the future event. If the incompatibilist rejects this supposition, then there is no entailment relation between the warning and the occurrence of the future action, and hence no temptation to think that the warning's occurrence is a soft fact. It should be clear, however, that this response will not generate any reason to think that the ability claim (A) is inconsistent with the backtracking counterfactual (B). To allow that the future action predicted by the warning may not occur is to give up on the claim that the act is not free. For if there is no guarantee of the act's occurrence, e.g., of Jones mowing his lawn at t_3 , then it is open to him to refrain from mowing—thus rendering, e.g., the warning false. And if Jones can render the warning false then there is no problematic counterfactual dependence between his action and a hard fact about the past. There is no need, that is, for him to be able to act such that the

²⁰See, for example, Fischer (1983).

²¹Plantinga (1989).

²²We too are following Plantinga (1989, p. 193), at least insofar as we are endorsing a restricted entailment criterion that begins with a certain sort of immediate or basic fact such as "Jones mows his lawn." The restricted entailment criterion says that any past fact that entails that the basic fact in question obtains at some future time is a soft fact about the past. so we join Plantinga (along with pretty much everyone else) in rejecting a fully general entailment criterion.

²³Here one might object that (W1), unlike (4), does not entail that Jones mows his lawn because although God's beliefs are necessarily true, it is possible that some of his warnings are conditional. But we have formulated (W1) in a way that forestalls this objection. Even if some of God's warnings are conditional, a warning that Jones will mow his lawn that is issued *prior to his mowing* cannot be a conditional warning.

warning would not have been issued.

In other words, we are suggesting a dilemma for Widerker. For an incompatibilist argument based on his strategy to do the necessary dialectical work in the debate against the Ockhamist, the relevant backtracking counterfactual, e.g., (B1) If Jones had refrained from mowing his lawn at t_3 , then (W1), which is a fact about t_2 , could not have been a fact about t_2

must be true. Moreover, the fact (W1) in the consequent of the backtracker (B1) needs to be a hard fact, so that the truth of (B1) clearly rules out the truth of the relevant can-claim:

(A1) Jones can refrain from mowing his lawn at t_3 .

The dilemma is that the truth of (W1) either entails that Jones mows his lawn at t_3 , or it does not—and in neither case does its truth call into question the relevant can-claim. If the truth of (W1) does entail that Jones mows his lawn, then the backtracker is obviously true. But in this case it is not at all obvious that (W1) is indeed a hard fact. Thus the Ockhamist can plausibly maintain that (W1) is a soft fact, in which case there is little reason for him to doubt the truth of the relevant can-claim (because, recall, Ockhamists are happy to affirm the conjunction of a can-claim and its corresponding backtracker when the fact in the consequent of the backtracker is a soft fact). If, on the other hand, the truth of (W1) does not entail that Jones mows his lawn, then (W1) seems to be an obvious case of a hard fact. But in this case it is not at all obvious that the backtracker (B1) is true—for there is little reason to think that (W1) could not have been a fact if Jones had refrained from mowing his lawn. (For example, as mentioned above, the warning could have been mistaken, which case (W1) would still have been a fact.) And of course if the backtracker is false, then there is no reason *stemming from the backtracker* to conclude that the can-claim is false. Thus, whether or not the truth of (W1) entails Jones's mowing of his lawn (at the relevant time), the Ockhamist should feel no pressure to doubt the truth of the claim that Jones can refrain from mowing at that time.

What about (W2)?

(W2) Smith heard a voice telling him that his puppy is in grave danger.

(W2) is farther removed from divine activity, and so not evidently susceptible to the Ockhamist move. Nevertheless, the same response is available. If the truth of (W2) entails the occurrence of the action that it predicts, then it is temporally relational; on this assumption, the voice in question is assumed to be God's voice. Hence, it is dialectically open to the Ockhamist to insist that it also exhibits counterfactual dependence—which is to say that the relevant backtracker

(B2) If Jones had refrained from mowing his lawn, then (W2), which is a fact about t_2 , could not have been a fact about t_2

is consistent with the claim (A1) that Jones can refrain from mowing his lawn. On the other hand, if Smith's hearing a voice does not entail that Jones will mow his lawn (because here it is not assumed that the voice in question is God's voice), then there is little reason to believe that his refraining from mowing will require that the voice would have said something different (or nothing at

all). In this case, the voice that Smith heard could have simply been wrong. Thus, if the truth of (W2) does not entail Jones's mowing, then (B2) seems to be false and as a result does not threaten the truth of (A1). As with (W1), it seems that without the guarantee produced by entailment, there is little reason to believe that the relevant can-claim is false.

Not surprisingly, we think that

(W3) Smith came to believe that his puppy was in grave danger

can be dealt with in a similar fashion.²⁴ We think that the Ockhamist could say the following: Given that we are talking only about a belief of Smith's, it seems quite clear that this fact is temporally non-relational, and thus fixed. Unfortunately for the incompatibilist, however, there is very little pressure toward thinking that Jones's doing otherwise will *require* altering the fact that Smith came to believe a certain thing. And if there is little reason to think that this backtracker is true, then there is little reason to think that the can-claim is false. The incompatibilist could of course remind us here that God's past belief is a causally necessary condition for this particular belief of Smith's—but this amounts to abandoning the indirect strategy, which is supposed to be neutral on the question of whether God's past beliefs are soft facts. (Recall that this strategy, as we see it, grants that God's past beliefs might be such that they can be changed by, e.g., Jones's actions, and tries to generate a problem by showing how that ability would entail an ability to falsify some obviously hard fact about the past.)

Widerker's argument uses a past fact F for which God's past belief is a causally necessary condition. Our general contention is that any indirect argument that uses a fact having this feature is going to fall prey to our dilemma. For the more directly F represents some activity of God, the more it is going to seem soft (at least by Ockhamist lights). And if it is soft, then there is at least a case to be made that the relevant backtracker is consistent with the relevant can-claim. Conversely, as F becomes more remote from divine activity, it will seem more plausible to claim that S need not falsify that fact in order to do otherwise—i.e., it will seem more plausible to claim that the relevant backtracker is false. And a false backtracker cannot be used to challenge the truth of any can-claims. In support of our general contention, we will now examine van Inwagen's implementation of the indirect anti-Ockhamist strategy.

§5 Van Inwagen's argument

van Inwagen begins with a variation of the standard argument for theological incompatibilism that he takes to be irrefutable. Assuming that God is temporal, he points out that if God believed in the past that he (van Inwagen) would tell a lie in the future (at a specified time), then in order for him to tell the truth at that time, either God's past belief would have to be false, or it would have to be the case that God did not have the past belief that he in fact had. But,

²⁴And we would argue that the same holds for (W4), . . . , (Wn).

claims van Inwagen, both of those are impossible; so he is not free to tell the truth if God believed in the past that he would tell a lie, and more generally, we are never free to do otherwise if God has exhaustive foreknowledge of future contingents.²⁵

Van Inwagen recognizes that this argument does not apply to a God who is atemporal, or outside of time—but he reformulates his argument in a way that is intended to apply with equal (or close to equal) force to the atemporalist about God.

The key aspect of the reformulated argument is a move from God's past belief about what one will do to the past existence of some temporal effect of God's timeless activity. (This makes van Inwagen's argument importantly similar to Widerker's, as we will explain in further detail below.) As the modified story goes, God causes, *ex nihilo*, a monument to come into existence in 1900. On this monument is etched a statement to the effect that van Inwagen will tell a lie at some specified point in the future (11:46am EST, 23 December 2006, as it so happens). Thus van Inwagen:

Suppose God has done this thing he is able to have done. Can it be that my lying . . . was a free act? That is, was I able, on that occasion, to tell the truth? Well, was there, just before that moment, a possible continuation of the (then) present state of affairs in which I told the truth? Let us consider all the possible continuations of that state of affairs. It is true in every one of them that [an inscribed] monument . . . came into existence *ex nihilo* in 1900—and true that its coming to be was caused by God's extra-temporal act of creation. Is it true in any of the possible continuations of the then-present state of affairs that the words inscribed on the monument did not express a true proposition? No, for in that case God would either have been mistaken or have been a deceiver, and both are impossible.²⁶

Van Inwagen concludes that his telling a lie is not a free act, since there is no possible continuation of the present of the sort that would be required for his telling a lie to be a free act—i.e., of the sort in which (1) the inscribed monument does not exist or (2) the inscription expresses a false proposition.

This inscribed monument falls under the category of what van Inwagen calls "Freedom-Denying Prophetic Objects"—a category which includes any object (animate or not) that infallibly expresses a true proposition about some future act of some putatively free creature. This notion of a Freedom-Denying Prophetic Object (FDPO) is intuitive and intriguing, albeit rather underspecified. The only general specification that van Inwagen gives us is that the concept of an FDPO is "a very abstract one."²⁷ Nevertheless, he does give some examples of other prophetic objects that could be freedom-denying—for example, an actual human prophet who foretells some human action with the help of divine inspiration would count as an FDPO. In any case, facts about the past existence of a monument appear to be hard facts, and if doing otherwise requires altering

²⁵van Inwagen (2007, pp. 217–218).

²⁶van Inwagen (2007, p. 219).

²⁷Ibid.

such facts, then the Ockhamist is back on the ropes.

Before we address van Inwagen's argument, notice the parallel with Widerker's argument. In particular, notice that van Inwagen makes no attempt to address the question of whether or not God's past beliefs about the future actions of free creatures are hard or soft facts about the past. Instead, he appeals to a fact the hardness of which (presumably) no one would challenge. Moreover, crucial to both arguments is the supposition that God, were he to have exhaustive foreknowledge, would bring about some state of affairs—facts about which would clearly be hard facts about the past (at any time after God brings about the state of affairs in question). For Widerker, the relevant state of affairs initiates a future-contingent causal chain, the links of which appear to constitute hard facts about the past. For van Inwagen, the relevant state of affairs is the existence of some object that infallibly foretells some future event. And whether we are talking about Smith hearing a voice in the past or the past existence of an inscribed monument, both facts seem plausibly to be hard facts about the past, and thus provide the grounds for establishing the incompatibility of divine foreknowledge and human freedom. Thus, although van Inwagen has a different target in mind (the atemporalist, rather than the Ockhamist), his approach can plausibly be classified as a development of the anti-Ockhamist strategy introduced by Widerker.²⁸

We have seen that when Widerker's facts (W1)(W3) are examined more closely, they do not give the Ockhamist sufficient reason to surrender his view. We will now argue that the same can be said for certain facts about (allegedly) freedom-denying prophetic objects.

§6 A dilemma for van Inwagen

The indirect strategy, recall, requires a fact F that does not involve any of God's beliefs and is clearly and uncontroversially a hard fact about the past—thus making it evident that the relevant claims,

(A) S , at t_2 , can do X at t_3

(B) If S had done X at t_3 , F could not have been a fact about t_1 ,

are inconsistent. If van Inwagen's argument does not meet these requirements, then it will not convince the Ockhamist, who can respond in a way that resembles the response developed in §4. He can, i.e., maintain that God's beliefs are soft facts, and that counterfactuals such as (B), which contain a fact about one of God's beliefs in the consequent, are consistent with corresponding can-claims. Thus, unless van Inwagen can give the Ockhamist a reason why a soft fact about an FDPO is different from a soft fact about one of God's beliefs, the Ockhamist

²⁸Another reason for considering these arguments together is that, as we pointed out above (note 16), although Widerker first developed this strategy against Ockhamism, he later (1991) employed it against the eternity (atemporalist) solution. So we are, as it were, completing the circle by bringing the argument back around to apply to the Ockhamist once again. But, as also noted above, even though we are confident that our defense of Ockhamism could be adjusted so as to constitute a defense of the eternity solution, we will not make those adjustments here.

seems within his rights in claiming first that counterfactuals such as (B), which contain a fact about an FDPO in the consequent, are also consistent with corresponding can-claims; and second, that if the fact about the FDPO is indeed a hard fact, then the appropriate backtracker will be false.

In van Inwagen's argument, the relevant claims are as follows:

(C) Van Inwagen, in 2005, can tell the truth at the specified time in 2006.

(D) If van Inwagen had told the truth at a certain time in 2006, then the (fact of the) existence of an inscribed monument testifying to his telling a lie at a certain time in 2006 could not have been a fact about 1900.

The relevant fact F is the existence of a particular FDPO: an inscribed monument testifying to van Inwagen's telling a lie at a certain time in 2006. The question (for the proponent of the indirect anti-Ockhamist strategy) is whether the truth of (D) rules out the truth of (C). And only if F is clearly a hard fact can the claim that the truth of (D) rules out the truth of (C) get any traction. As it turns out, however, F is not obviously a hard fact about the past. At first blush, it may *seem* to be a hard fact; after all, what could be harder than the past existence of an inscribed monument? But the truth (perhaps obscured by the lack of detail in van Inwagen's characterization of FDPOs) is that we have good reason—so says the Ockhamist—to believe that facts about the existence of FDPOs are actually soft.

The proposed line of attack here runs parallel to the line of attack against Widerker. It seems natural to think that the presence of the FDPO in question (i.e., the inscribed monument) entails that the future act (the lie to which the monument testifies) will occur. This seems true because, after all, it was God who caused the monument to come into existence. As van Inwagen points out, God cannot be mistaken, and he cannot be a deceiver; so it must be true that the existence of the monument entails the future telling of the lie. But, again, this is precisely the feature of a past fact that is supposed to make it a soft fact. It seems plausible to claim, relative to 2005 at least, that any fact about 1900 that entails that Peter van Inwagen will tell a lie at 11:46am EST on 23 December 2006 is a soft fact about 1900. Given the softness of this fact, the Ockhamist can maintain that van Inwagen can tell the truth on 23 December 2006, and further that the can-claim (C) is consistent with the truth of (D)—i.e., consistent with the claim that if he does tell the truth on 23 December 2006, then the inscribed monument would not have existed in 1900 (or would not have been inscribed to foretell a lie occurring in 2006). Thus, continues the Ockhamist, the possible presence of the inscribed monument (or any other alleged FDPO) does not endanger van Inwagen's (or anyone else's) freedom.

Moreover, giving up on the entailment relation helps van Inwagen no more than it helps Widerker. To suggest that the presence of the FDPO does not entail the occurrence of the future action it predicts is to allow that the future action may not occur, and hence to give up on the claim that the FDPO presents a threat to the relevant agent's freedom. For if there is no guarantee of the act's occurrence (e.g., of van Inwagen's lying in December 2006), then it is open to

him to refrain from lying—thus rendering, e.g., the inscription on the monument false. In other words, giving up on the entailment means that the backtracker (D) is false, and hence that it cannot call into question the truth of the can-claim (C). Or, to put the point in a different way, if the inscription on the monument is a hard fact about 1900, then van Inwagen should view it as no more threatening to his ability to tell the truth than the following sort of case: It is November of 2006. Wishing to expose van Inwagen in a lie, his friend hires a cloud writer to write across the sky, “Van Inwagen will tell a lie at 11:46 EST on 23 December 2006.” Seeing the cloud-writing wouldn’t give van Inwagen any reason to doubt his ability to tell the truth at that time, and—given the lack of entailment—neither should the inscription.

In short, the dilemma faced by van Inwagen is that the existence of the monument either entails that he tells a lie in 2006, or it does not—and in neither case should its existence preclude the truth of the relevant can-claim (C). If the monument’s existence does entail that van Inwagen tells a lie, then the fact of its existence appears, at least by Ockhamist lights, to be a soft fact. And if it is a soft fact, then there is little reason for the Ockhamist to doubt the truth of (C). If, on the other hand, the existence of the monument does not entail that van Inwagen tells a lie, then the backtracker (D) is false, and hence there is no reason for the Ockhamist to doubt the truth of (C). Thus, whether or not this putative FDPO entails the occurrence of van Inwagen’s lie at the relevant time, there is little pressure on the Ockhamist to doubt the truth of the claim that van Inwagen can refrain from telling the lie at that time. We conclude that van Inwagen’s argument, like Widerker’s, fails as an implementation of the indirect anti-Ockhamist strategy.

§7 Widerker’s rejection of entailment criteria

Since the notion of entailment plays a significant role in our critique of the indirect strategy, we should acknowledge and briefly discuss Widerker’s rejection of various entailment criteria for soft facthood. Widerker considers and dismisses three attempts to analyze the concept of a soft fact in terms of entailment: William Rowe’s, Marilyn Adams’s, and Alvin Plantinga’s.²⁹ In each case, the basis of his dismissal is the following claim. He says that entailment is not a reliable indicator of soft facthood, since there are clear examples of past facts that are hard (i.e., past facts that are now “over-and-done-with”) but nonetheless entail certain things about the future. So, for example, Widerker construes Adams as claiming that a hard fact about a time t “must not entail the obtaining of a state of affairs at a time later than t .”³⁰ But Widerker asks us to consider the fact of God promising Smith at t_1 that Jack will sign a contract at t_3 :

(6) God promised Smith at t_1 that Jack will sign the contract at t_3 .

²⁹Widerker (1990, pp. 465–470).

³⁰Ibid., p. 466.

This fact entails something about t_3 , namely that Jack will sign a contract at t_3 .³¹ But, continues Widerker, “It seems intuitive to assume that the fact that God promised Smith at t_1 that Jack will sign the contract is fully accomplished and over-and-done-with [i.e., hard] at t_2 .”³² So it appears incorrect to say that a hard fact must not entail anything about the future.

We hope that it is clear by now that this argument, based on the divine promise example, is insufficient (at least by itself) to show that entailment cannot be part of an account of soft facthood. The same reasons that support the softness of a fact about God’s past beliefs will also support the softness of a fact about God’s past promises (or warnings). In fact, this point holds in the other direction as well: the claim that a fact about one of God’s past beliefs is hard is no less intuitive than the claim that a fact about one of God’s past promises is hard. We agree that it is intuitive to claim that past facts about God’s promises are hard facts. But it is also intuitive to claim that past facts about God’s beliefs are hard facts. (This intuitiveness is, we take it, precisely the reason that contemporary formulations of the standard argument for theological incompatibilism typically rely on a fact about God’s past beliefs rather than a fact about his past knowledge.) The Ockhamist innovation was that God’s beliefs might not be hard after all. If this (putative) insight applies to God’s beliefs, but not God’s promises (warnings, etc.), then we need a reason why.

Widerker does attempt to provide a reason why facts about God’s promises are different than facts about God’s beliefs. He points out that a divine promise “can be remembered by Smith, can be empirically detected by him . . . and may have traces . . . just like regular hard facts about the past.”³³ But this reason is inconclusive at best. To see why, first note that whatever the similarities between facts about God’s promises and regular (i.e., obvious and uncontroversial) hard facts about the past, there remains a crucial difference: facts about God’s promises—like facts about God’s beliefs—are always going to have two parts. And the part that pertains to the future (i.e., the part that pertains to the act or state of affairs promised by God) will make its enclosing fact a soft fact. In other words, the future-directed part of facts about God’s promises appears to

³¹We should note that (6) only entails that Jack will sign a contract at t_3 on the traditional theistic assumption that God is essentially immutable. For if God is not essentially immutable, then he could promise one thing at t_1 and later change his mind and promise something else at t_2 . (For an illuminating discussion of some of the issues surrounding the question of whether God’s promising that p entails that p , see Kvanvig (2011)) But notice also that rejecting this assumption is of no help to Widerker, because the restricted entailment criterion that we have been relying upon merely says that entailment of a particular sort of fact is sufficient for soft facthood; nothing follows from a failure of entailment. Thus we can also construe the point of this section as a dilemma: God’s promise that p either entails that p or it doesn’t entail that p . If it entails that p , then it is on a par (as regards its hardness or softness) with God’s belief that p . If it doesn’t entail that p , then the restricted entailment criterion doesn’t apply. Either way, divine promises do not provide any reason to reject the criterion.

³²Ibid., pp. 467–468.

³³Ibid., p. 469.

provide a “handle” by which those facts can be altered.³⁴

Consider again Plantinga’s view on facts about God’s beliefs. On his view, the reason why

(4) God believed that Jones would mow his lawn tomorrow

is a soft fact (relative to today) is that (4) is logically equivalent to

(7) God believed that Jones would mow his lawn tomorrow and Jones will mow his lawn tomorrow,

which is soft because no conjunctive fact that contains “Jones will mow his lawn” as a conjunct can be hard. And since hardness is closed under logical equivalence, (4) cannot be hard either.³⁵

Now return to Widerker’s divine promise example. Notice that the relevant fact, (6), is logically equivalent to the following conjunctive proposition:

(8) God promised Smith at t_1 that Jack will sign the contract at t_3 and Jack will sign the contract at t_3 .

On Plantinga’s view, (8) is soft because no conjunctive past fact (relative to t_3) that contains “Jack will sign the contract at t_3 ” as a conjunct can be a hard fact. And (6) will turn out soft as well, because it is logically equivalent to (8). So Plantinga appears to be committed to the softness of (6).

Widerker takes this to show that Plantinga is wrong, because (6) “surely expresses a hard fact about the past.”³⁶ How can we be so sure that (6) expresses a hard fact about the past? Perhaps it is because God’s promises are going to leave empirical traces that can be detected (facts about which would, were it not for God’s involvement, obviously be hard). But note that the empirical traces issue only from a *part* of a fact such as (6). That part of such a fact leaves empirical traces does not imply that the fact as a whole is hard, especially given that it has *another* part that is intuitively about the future.³⁷

For those who remain uncomfortable with our use of the restricted entailment criterion, notice that we could run our argument with a different analysis of the distinction between hard and soft facts. On all of the analyses that we have encountered, facts such as (W1) and (6)—which involve a divine warning and a divine promise, respectively—will come out soft. But pointing out that (W1) and (6) intuitively seem hard is not enough to defeat these analyses—at least not if it is granted that facts about God’s beliefs might be soft. For facts about God’s beliefs intuitively seem hard at first, but upon analysis turn out to be at least arguably soft. Why should God’s warnings or promises be any different?

³⁴Thanks to John Fischer for this helpful metaphor.

³⁵Plantinga (1989, pp. 193–194).

³⁶Ibid.

³⁷This point might lead someone to reject the claim that some of God’s past beliefs can be soft facts about the past. That result is acceptable to us, as our point here is merely that beliefs and promises (decrees, warnings, etc.) stand or fall together as regards their softness.

§8 Conclusion

We will close by summarizing the dialectic and our contribution to it. First, the Ockhamist claims that facts about God's past beliefs are soft. One reason, although certainly not the only reason, to suppose that such facts are soft is that they satisfy Plantinga's "restricted entailment" criterion for soft facthood. And, given that they satisfy this entailment criterion, their being facts about the past depends on something basic happening at a future time. That is, they are temporally relational. Perhaps the fact about God's belief in the past "includes" the subsequent fact, or perhaps it depends on it in some weaker way—but in either case, says the Ockhamist, the past fact about God's belief cannot uncontroversially be considered "over and done with" in the past.³⁸

But this question—of whether God's beliefs are temporally relational—is too contentious for some of the opponents of Ockhamism. These opponents, Widerker and van Inwagen among them, would rather sidestep this question. So, for example, Widerker and van Inwagen claim (or at least would claim) that their proposed freedom-denying prophetic objects (or events), being physical objects or events, with empirical traces, succeed in avoiding the controversy. But the dilemma we present above shows that they cannot in this way sidestep the controversy. More specifically, we have argued that when the facts in question are construed so as to involve God, their obtaining (i.e., the existence of the object, or the occurrence of the warning) entails that something basic occurs later. Thus, although it might at first appear that the existence or occurrence of the object or event is temporally non-relational, arguably it is not. And finally, if the facts in question are construed so as not to involve God, then there is none of the counterfactual dependence (between belief and action) that is supposed to force the Ockhamist to concede the falsity of the relevant can-claim.

Notice that we are not arguing for the acceptance of Ockhamism. There may still be a successful anti-Ockhamist strategy. Our main point is simply that the only way for the *indirect* anti-Ockhamist strategy to get any traction is if the fact under consideration (the fact about a putatively freedom-denying prophetic object or event) is truly and uncontroversially a hard fact about the past. Furthermore, and crucially, the fact has to be such that refraining from the putatively free action will require altering that uncontroversially hard fact. But, at least in the specific cases we have examined above, the posited fact is not what it needs to be. If it entails the occurrence of the action in question, then it is arguably not a hard fact after all. On the other hand, if it does not entail anything about the action, then the relevant backtracker will be false and so nothing about the posited fact rules out the possibility of some agent freely rendering *it* false as well. In neither case should the Ockhamist conclude that our freedom to do otherwise is in danger.³⁹

³⁸Thanks to John Fischer for helping us summarize this stage of the dialectic clearly and concisely.

³⁹Many thanks to John Fischer for his thoughtful and penetrating comments on earlier drafts. Thanks also to Kenny Boyce, Brandon Carey, Trent Dougherty, Chris Franklin, Carl Ginet,

REFERENCES

- Adams, Robert. 1989. "Reply to Kvanvig." *Philosophy and Phenomenological Research* 50.2: 299–301.
- Fischer, John Martin. 1983. "Freedom and Foreknowledge." *Philosophical Review* 92: 67–79.
- Fischer, John Martin. 1997. *The Metaphysics of Free Will*. Malden, MA: Blackwell.
- Freddoso, Alfred J. 1989. "Accidental Necessity and Logical Determinism." pp. 136–158. Stanford, CA: Stanford University Press.
- Hasker, William. 1989. "Foreknowledge and Necessity." In *God, Foreknowledge, and Freedom*, edited by John Martin Fischer. Stanford, CA: Stanford University Press.
- Hoffman, Joshua and Gary Rosenkrantz. 1989. "Hard and Soft Facts." In *God, Foreknowledge, and Freedom*, edited by John Martin Fischer. Stanford, CA: Stanford University Press.
- Kvanvig, Jonathan L. 2011. 140–178. pp. 140–178.
- Pike, Nelson. 1965. "Divine Omniscience and Voluntary Action." *The Philosophical Review* 74.1: 27–46.
- Plantinga, Alvin. 1989. "On Ockham's Way Out." In *God, Foreknowledge, and Freedom*, edited by John Martin Fischer. Stanford, CA: Stanford University Press.
- van Inwagen, Peter. 1983. *An Essay on Free Will*. Oxford: Clarendon Press.
- van Inwagen, Peter. 2007. "What Does an Omniscient Being Know about the Future?" pp. 216–230. Oxford: Oxford University Press.
- Widerker, David. 1990. "Troubles with Ockhamism." *Journal of Philosophy* 87: 462–480.
- Widerker, David. 1991. "A Problem for the Eternity Solution." *International Journal for Philosophy of Religion* 29: 87–95.

David Hunt, Jon Kvanvig, Ben Mitchell-Yellin, Wes Morriston, Philip Swenson, Patrick Todd, Neal Tognazzini, David Widerker, and anonymous referees from Oxford University Press for insightful comments on earlier versions. None of God's past beliefs were falsified in the writing of this paper.

SOCIAL EVIL*

Ted Poston

Abstract: Social evil is any pain or suffering brought about by game-theoretic interactions of many individuals. This paper introduces and discusses the problem of social evil. I begin by focusing on social evil brought about by game-theoretic interactions of rational moral individuals. The problem social evil poses for theism is distinct from problems posed by natural and moral evils. Social evil is not a natural evil because it is brought about by the choices of individuals. But social evil is not a form of moral evil because each individual actor does not misuse his free will. Traditional defenses for natural and moral evil fall short in addressing the problem of social evil. The final section of this paper discusses social evil and virtue. I begin by arguing that social evil can arise even when individual virtue is lacking. Next, I explore the possibility of an Edwardsonian defense of social evil that stresses the high demands of true virtue. In this context, I argue that social evil may arise even when all the participants are truly virtuous. The conclusion of this paper is that social evil is problematic and provides new ground for exploring the conceptual resources of theism.

Discussion on the problem of evil assumes that there are two classes of evils: natural evil and moral evil.¹ Richard Swinburne divides “the world’s evils in the traditional way into moral evils and natural evils.”² Swinburne characterizes moral evils as “those brought about by human intentional choice, or knowingly allowed to occur by humans, together with the evils of their intentional bad actions or negligence.”³ Natural evils, according to Swinburne, are “all other evils, such as bad desires that we cannot help, disease, and accidents.”⁴ Alvin Plantinga provides a similar division of evils. He writes, “In addition to natural’ evils such as earthquakes, tidal waves, and virulent diseases there are evils that

Thanks to Lara Buchak, Trent Dougherty, Allan Hillman, Jonathan Jacobs, Clayton Littlejohn, Kevin Meeker, Bradley Monton, Alvin Plantinga, Alexander Pruss, Meghan Sullivan, Richard Swinburne, Peter Vallentyne, Paul Weirich, and Heath White for excellent comments on earlier drafts. I am especially grateful to William Wainwright for detailed and penetrating comments on an earlier draft.

¹The distinction between two classes of evil goes back, at least, to Augustine. On natural evil see Augustine’s *De Ordine* and on moral evil see *Confessions and On Free Will*.

²Swinburne (2004, p. 236); cf. van Inwagen (2006, p. 9).

³Ibid., 236.

⁴Ibid.

result from human stupidity, arrogance, and cruelty.”⁵ William Rowe’s famous Bambi and Sue cases represent, respectively, natural and moral evil. The Bambi case involves the prolonged suffering and death of a fawn; the Sue case relates the horrible beating, rape, and murder of a five-year old girl.⁶ Both these forms of evil exercise the conceptual resources of theism to explain how an omnipotent, omniscient, and morally perfect being might allow these evils to occur. But there is another form of evil that has not received the attention of philosophers working on the problem of evil. This form of evil I call *social evil*.

At the outset I want to be clear about the nature of evil and its various types. For the purposes of this paper, I assume that evil is any instance of pain or suffering and that types of evils are individuated by the processes that bring them about. Moral evil is an instance of pain or suffering brought about by the direct agency of a person, and natural evil is an instance of pain and suffering brought via the operation of laws of nature on matter. The nature of direct agency is difficult to analyze. I think of direct agency via moral responsibility. A person exercises direct agency in performing an action when they have full responsibility for that action. Even though action may require background conditions such as the operation of laws of nature, a person can still be held fully accountable for her act. Thinking about natural evil in terms of responsibility, the “blame” for natural evil lies with impersonal forces. If John is killed by an avalanche, the reason for his death lies in inanimate forces that produced the avalanche. Social evil is an instance of pain or suffering that results from the game-theoretic interactions of many individuals. When a social evil occurs, responsibility for the outcome lies with no particular person and no impersonal force of nature; rather it lies with a group of people, each of whom may be morally in the clear. Until the last section of this paper, I focus on the most problematic form of social evil, social evil that occurs because of the game-theoretic interactions of rational moral individuals. It is widely acknowledged in the literature on game theory that rational well-intentioned agents can bring about horrible social outcomes. Russell Hardin (1995) provides an arresting example of this by offering a game-theoretic analysis of violent group conflict. Once we see the game-theoretic machinery in play, it is hard to resist the thought that much evil in our world is the unintended result of collective agency among individually rational participants. My goal in this paper is to introduce and explain social evil and the problems it poses for traditional defenses. Furthermore, I discuss the prospects for a theistic diagnosis of social evil. I intend my discussion to be a starting point for deeper reflection on the nature of social evil and the theoretical lessons it has for theism. The concept of social evil has wider applications to ethical theory. Reflection on social evil shows that, in at least some cases, an individual’s actions are not the result of a bad will but rather the result of an individual being caught in a tragic game-

⁵Plantinga (1977, p. 8).

⁶The original “Bambi” case comes from Rowe (1979) and the “Sue” case is in Rowe (1988). William Alston (1991) dubbed these cases “Bambi” and “Sue” and, following Alston, the terminology has stuck.

theoretic scenario. An understanding of social evil suggests that there are far fewer moral monsters that we might otherwise think.

I. What is social evil?

I begin with a clear case of pain and suffering that results from the game-theoretic interactions of rational, well-intentioned individuals. Suppose you are a resident of Los Angeles and the greater Los Angeles area is facing a serious water shortage. The reservoirs in northern California are running dry; the Los Angeles and San Gabriel rivers are bone dry; even Oregon's plentiful lakes and rivers are ominously diminished. Without a significant decrease in overall water consumption, the Los Angeles area will run out of an adequate water supply. City planners foresee the possibility of severely restricting residential water use. However, if most residents significantly decrease their water consumption—by not watering lawns, washing cars, or letting the tap run unnecessarily—the Los Angeles area will manage until the winter rains come. Obviously, it is in the best interest of all that *most* everyone follows this advice. But this represents a considerable cost to *each* person. If, for example, you decrease your water usage, your carefully cultivated garden and fruit trees will wilt and die. This is a hefty burden to pay. However, if no one decreases his water usage, each will pay an even greater cost. Yet you realize that if most everyone decreases his water consumption, then you may continue your normal usage without any ill consequence. Moreover, because the benefit of decreased water usage requires a very large number of participants—well over a million homeowners—your own individual contribution does not affect whether or not the benefit is realized.

The Los Angeles water shortage case is a simultaneous move game. A simultaneous move game is one in which you (and everyone else) act in ignorance of what the other people do. In this case let us assume that you are rational and blameless. You do not suffer from a failure to realize that you are in this kind of circumstance. Moreover, you do not suffer from a moral fault; you do not want to harm anyone and you do not want your action to bring about a worse circumstance. Furthermore, while you could reason in a self-interested manner, you do not. You reason from states of value in the world. Your garden is an item of beauty and it is worth preserving. In this situation you realize that your decision whether to continue to water affects only whether an item of value is preserved, i.e., your beautiful garden. Taking into consideration all the relevant factors, the best option for you is not to conserve water. But many other Los Angeles residents face similar situations in which the best option for each is not to conserve water. If most everyone is rational and blameless, each individual will play his best strategy and the collective result will be unintended disaster. The disaster that results is a social evil.

This is a standard form of a multiplayer prisoner's dilemma, also known as "the tragedy of the commons."⁷ Cases of this sort are widespread. Achieving

⁷See G. Hardin (1968), Kuhn (2009).

adequate healthcare, decent education, effective inoculations, safe freeways, fertile fishing waters, and pristine national parks all require the cooperation of a sufficiently large group of individuals. In these cases the goods achieved and the evils avoided require solving a multiplayer prisoner's dilemma. Because each individual's dominant strategy is to depart from the cooperative act (i.e., to defect), governments aim to avoid the destructive logic of these games by, among other things, imposing significant penalties for defecting. While this is a pressing practical problem, there is a theoretical problem for theism that has not been addressed in the literature to date.

To guide our discussion I present a standard two-person prisoner's dilemma.⁸ We can consider a two-person dilemma as the smallest case of a social evil, a case in which an unintended and worse outcome results for each player enacting his best strategy. Consider the following game.

Two-Person Prisoner's Dilemma

		Player II	
		Defect	Cooperate
Player I	Defect	2,2	4,1
	Cooperate	1,4	3,3

In this game Player I has two options: either defect (maintain his water usage) or cooperate (decrease his water usage). Player I achieves his *best* result by defecting given Player II's cooperation, and he achieves his *worst* result by cooperating given Player II's defection. The numbers 4-1 represent an *ordinal* ranking of the players' preferences. An outcome 4 is preferred to 3 which is preferred to 2 which is preferred to 1. Ordinal rankings reflect a series of ordered preferences, and they do not reflect the *strength* of preference. With an ordinal ranking you cannot infer that Player I's highest ranked option, 4, is twice as desirable as his second ranked option, 2. Player II's rankings are symmetrical with Player I's. Thus, for Player II his highest ranked option is where he maintains his current water usage and Player I decreases water usage.

What should each Player do? Given Player I's situation, he can always improve his outcome if he defects. To see this suppose Player I cooperates. Then if Player II defects, Player I would do better by *defecting*. He would move from an outcome of 1 to an outcome of 2. If Player II cooperates, then, just as in the other situation, Player I would do better by *defecting*. He would move from an outcome of 3 to an outcome of 4. In this situation Player I's option of defecting *strictly dominates* cooperating because regardless of what Player II does, Player

⁸I give a multiplayer prisoner's dilemma below in section IV. At this point in our discussion, we will gain no appreciable advantage by using the more complex multiplayer PD.

I can do better by defecting. Player II, of course, should engage in the same reasoning and so Player II should defect. Thus, these two rational well-intentioned players achieve an outcome neither wants, viz., mutual defection. Both players would prefer mutual cooperation to mutual defection but the logic of the game lands rational, well-intentioned players in mutual defection.

The water shortage case illustrates the painful logic of group action. Each participant realizes the structure of the game they are in and each participant is blameless. But *given* the structure of the game, it is rational for each player to defect. Note that because of the nature of group action, each individual player does not bring about a bad state of affairs. In large multi-player prisoner dilemmas the benefit (or detriment) is achieved regardless of what any one individual does. In more realistic games—like the water shortage case—the benefit (or detriment) is insensitive to at least one tenth of a percent of the total number of players. For example, if a multiplayer prisoner's dilemma is realized with 10^6 players then the social outcome will obtain regardless of what 10^3 players do. If 1,000 players don't make a difference to this game your decision surely does not make a difference. Consequently, each individual player's choice does not affect the outcome of the game. Moreover, in large multiplayer games each player may know that defection will bring about a Pareto improvement. A Pareto improvement to an allocation of resources is a redistribution of those resources that makes someone better off without making anyone worse off. In a large multiplayer prisoner's dilemma, any change in any one individual's strategy doesn't affect anyone else, so a player can know that defection will be a Pareto improvement. We might say that the problem of social evil is that the road to hell is paved with Pareto improvements.

I assume that this discussion is sufficient to motivate the thought that social evils are distinctive because they occur as the result of game-theoretic interactions among many individuals and *not* on account of some individual's choice (as in the case of moral evil) or some natural process (as in the case of natural evil). Let us now examine challenges to the idea that *social evil* is a distinctive kind of evil.

II. Three objections to the distinctness of social evil

Is social evil really distinct from natural and moral evil? In one sense it is obvious that social evil is distinct from natural and moral evils. Natural evils are instances of pain and suffering brought about by natural processes like earthquakes and tidal waves. Moral evils are instances of pain and suffering brought about individual human choice. If Sam loathes Bill and breaks his nose, then this is moral evil. Yet in another sense, one may very well wonder whether social evil is distinct from moral and natural evil. Might not social evil really be the problem of limited resources? Might not social evil be a subtle form of natural evil, viz., human stupidity? Finally, might not social evil be a subtle form of moral evil? In this section I respond to these questions.

A. *Social evil and natural evil*

One objection to the distinctness of social evil is that social evil is really a form of natural evil and so an adequate defense for natural evil will carry over to social evil. The water shortage example involves a severe drought; droughts, like earthquakes, are natural evils. Even so, the extent of the pain and suffering of the Los Angeles drought depends upon the collective response of a great number of people. If most everyone cooperates, the drought will not drastically affect the Los Angeles area. The reason a drought can be so devastating lies not merely in the severity of the natural conditions but also in the collective response of many individuals. A group is accountable for the severity of the drought even though no particular person is accountable. So even though there is a natural evil in the background, the problem in the foreground is social evil.

The present objection is that any form of social evil is a form of natural evil on account of a scarce resource. In the standard cases of multiplayer prisoner's dilemma there is a limited resource that needs to be distributed over a large population. But the present objection infers an evil from a scarcity of resources. This move is dubious. What reason do we have for thinking that some pain or suffering will result from the mere fact that a resource is limited? Even in normal circumstances, water is a limited (i.e., a scarce) resource. Everyone can't run the tap all day long. The fact that a resource is limited doesn't imply that any pain or suffering results. It may only imply that a limited resource has no Pareto-efficient allocation. A Pareto-efficient allocation is one that admits of no Pareto improvement. As we saw in the water shortage case, there is always a Pareto improvement to the allocation of water resources; there is always a reallocation of water related benefits and burdens in which some individual is better off and no one else is made worse off. Any individual can continue to use water thereby increasing overall value in the world in some respects without making the world any less valuable in other respects. In general, the structure of the tragedy of the commons always permits a Pareto improvement. If the commons is a grassy green then one could always permit one more sheep to graze without affecting the overall outcome. To summarize: a situation without any Pareto-efficient allocation of resources need not be one in which there is suffering and so such a situation need not be an instance of natural evil.

Another response to the objection that social evil is a natural evil on account of limited resources is that social evils can arise from resource *abundance*.⁹ Many harms to our environment are caused by society overusing a certain resource. It is morally permissible for any one individual to drive a car to work, but the result of millions of individuals driving is an overabundance of carbon which results in global warming, thereby causing some pain or suffering. Similarly, any one individual is morally permitted to fertilize his yard. Yet the result of many individuals doing the same thing is an overabundance of fertilizer in our lakes and rivers, which ruins our recreational fishing grounds.

⁹Thanks to an anonymous referee for this suggestion.

B. Social evil and rationality

A second objection to the distinctness of social evil is that it is a subtle form of natural evil, specifically human stupidity. If an individual defects while fully realizing that everyone else faces the exact same reasoning, then the person is stupid. But why should we think that the individual defector is stupid? One reason is that the individual does not realize that his defection brings about a worse state of affairs, viz., one confederate fewer. But this reason rests on a serious misunderstanding of the logic of multiplayer prisoner's dilemmas. The individual defector does not *bring about* a worse state of affairs by bringing it about that there is one fewer confederate. This is because the social benefit will be achieved (or not) regardless of what an individual person does. That is, for any n , $n \pm 1$ confederates do not change the outcome of the game. If one million confederates realize the outcome then one additional or one fewer confederate will not change the outcome. To suppose that an individual's choice affects the outcome of a multiple player game is a gross misunderstanding of the logic of such games.

The principle that *for any n , $n \pm 1$ confederates do not change the outcome of the game* is true of cases in which the threshold of confederates required to achieve the good and avoid the evil is vague. In cases of vagueness, one cannot apply the principle iteratively without at some point losing knowledge about whether the threshold is met. In the Los Angeles water shortage case, let us suppose that three million confederates will achieve the social good but one hundred thousand will not. One can gradually diminish the numbers from three million participants to one hundred thousand, but at some point in this series it becomes vague whether that number of participants will achieve the good. Still, it remains true that one knows that small changes, i.e., plus or minus 1, will not affect the outcome to be achieved.

Readers familiar with the vagueness literature will recognize that the epistemicist claims that this margin of error principle is false.¹⁰ According to epistemicism, vagueness arises because of ignorance. We do not know the position of the cutoff, the position at which $n + 1$ number of grains makes a heap while n does not make a heap. Consequently, the epistemicist claims, we mistake our ignorance about cutoffs for the non-existence of cutoffs (i.e., for the margin of error principle being *true*). There is much to be said in favor of epistemicism: it is elegant, upholds classical logic, and handles higher-order vagueness. But it is false. Epistemicism is primarily developed as a semantic theory designed to get classical truth-conditions for vague terms. But in our case we are interested in causal influences. Given the nature of a multiplayer prisoner's dilemma any single player does not affect the overall outcome. Relatedly, if epistemicism is true, then game theory must be revised since the claim that there is no Pareto-efficient allocation of a resource rests on a margin of error principle.

¹⁰See Williamson (1994).

One final attempt to argue that defection is irrational is worthy of mention. Douglas Hofstadter provides a sophisticated defense of the line that defection is not rational, which explicitly acknowledges that an individual has no causal influence on the social outcome.¹¹ Hofstadter's analysis depends on the claim that each player in the game is superrational. A superrational player realizes that reason alone will recommend a strategy in a collective action game and also realizes that each other player, in so far as they are superrational, will receive the same recommendation from reason. Thus, a superrational player will renormalize her opinion based on the common knowledge that each player is superrational. Consequently, Hofstadter reasons that each player faces the choice between universal cooperation and universal defection, since reason itself will not offer one recommendation to one player and a different recommendation to a different player.¹² Hofstadter claims that for the superrational it is evident that universal cooperation is the best option.¹³

Hofstadter's appeal to superrationality is intriguing but it does not substantiate the claim that defection is always irrational. First, ordinary agents who have not renormalized their opinions in line with superrationality are not to be faulted for following dominance reasoning. An individual in a multiplayer prisoners' dilemma often knows that defection will bring about a better world. In the water shortage case, a resident knows that regardless of what other people do, defection results in a more valuable world. It's hard to see what is irrational about defection in this case. Cooperation, in effect, knowingly leaves a real opportunity to make the world a better place unrealized. Second, Hofstadter's solution turns the prisoner's dilemma into a Newcomb's problem. That is, on Hofstadter's analysis in a prisoner's dilemma case the superrational face a conflict between dominance reasoning and the principle to maximize expected utility. This latter principle is relevant because, according to Hofstadter, reason cannot justify an asymmetry between any two superrational players. Thus, one's decision to defect—for example—should reflect the fact that if you conclude to defect, everyone concludes to defect since reason cannot give different recommendations to different players. But, as David Lewis has argued, it's unclear that reason recommends following the principle to maximize expected utility when one is faced with straightforward dominance reasoning.¹⁴ After all, one knows that regardless of how the world turns out, one has an available act that results in a better world. In sum, Hofstadter's appeal to superrationality does not show that defection is always irrational.

¹¹See Hofstadter (1985) *Metamagical Themas*, chapters 30-32. Thanks to Brad Monton for pointing out the relevance of Hofstadter's superb essays.

¹²For ease in exposition, I ignore the possibility of probabilistic strategies.

¹³See Hofstadter (1985, p. 746).

¹⁴David Lewis, independently of Hofstadter's ideas, claims that the prisoners' dilemma is at root a Newcomb's problem. Lewis then argues that defection is rational on the basis of an appeal to dominance reasoning. See Lewis (1979). Also see Binmore (1994), 3.4-3.5 for an extended rebuttal to a line similar to Hofstadter's.

C. Social evil and moral evil

A third objection to the special nature of social evil is that it is a form of moral evil. Specifically, an individual defector is morally to blame for defecting. While there *might* be something to this claim (though see sections III & IV below), it will require serious argumentation that goes beyond standard characterizations of moral evil. On the traditional view, moral evil is a form of pain and suffering that results from the agency of another person. Typically, this involves the *misuse* of free will.¹⁵ Clearly, in a multiplayer prisoner's dilemma an individual defector does not bring about any (relevant) pain or suffering. The individual defector is not an outcome-cause of the effect. The effect will occur regardless of what the individual does. Just as one raindrop does not make France fertile, so one person's decision in a large multiplayer prisoner's dilemma does not affect the overall outcome.

One response to this line of argument is that it mistakes the logic of rational choice with the logic of moral responsibility. A prisoner's dilemma represents the interaction of preferences between various agents. If one agent has a dominant strategy, one cannot infer from dominance alone that an agent is morally in the clear to enact that strategy. In particular, to be morally just an agent's action must be universalizable; it must be possible for everyone to act on the agent's maxim. But clearly, so the objection goes, an individual's strategy to defect is not universalizable in a multiplayer prisoner's dilemma. If a defector acts on the maxim *I will defect to bring about a better world*, then this is not universalizable because if everyone acted on that maxim a much worse world would result. Thus, the objection is that defection is immoral because it doesn't pass the universalization requirement.

This response may appeal to those with Kantian moral sensibilities. The problem with this reply, though, is that universalizability is not a necessary condition for moral permissibility. Suppose you are considering what kind of life you should live. Should you be a surgeon by which you may save many lives or should you develop your talent as an artist through which you may enrich the lives of many? Each decision is made against the backdrop of a diverse population with different aims and goals. A life as a surgeon assumes that many people are not doctors, and a life as an artist assumes that not everyone is an artist. One cannot simply universalize the maxim *I will become a surgeon to save the lives of many* because if everyone became a doctor that specialization would collapse.

This objection to universalization shows that if the requirement is to be plausible at all, it must allow that diversity is represented within the universalization test. To test the moral appropriateness of one's choice to become a doctor one must universalize the maxim assuming that other people choose different lives to lead. But once we allow for diversity in the universalization test, the original objection that defecting is blameworthy falls apart. For one can represent in the universalization test that one's action will not influence what the group will do.

¹⁵Though see Robert Adams (1985) "Involuntary Sins."

Even more strongly, one can represent that one knows this to be the case. Once this is represented then the defector's maxim passes the universalization test. There is nothing morally inappropriate about defection when one knows that one's action will bring about a better world without making anyone worse off.

Another possible attempt to explain the moral wrongness of defection is rule utilitarianism. According to this view, an act is morally good if and only if it is in accord with the optimific rule, the rule that has the best consequences. Clearly, if everyone cooperated then we would achieve the full benefits of cooperation. So, the objection goes, complete and full cooperation is the optimific rule, and in this connection defection is unjust.

The problem with this reply is that there is no unique optimific rule in a multiplayer prisoner's dilemma. For any number of confederates, one less confederate will achieve the same social utility. Moreover, an individual's defection will bring about more value in the world. An individual can know all this and act to bring about a better world without adversely affecting anyone. Rule utilitarianism does not imply that an individual defector has done anything amiss.

Perhaps, though, if we examine Kant's second formulation of the categorical imperative—to treat people as an ends, never as a means—we can discover why it is morally wrong to defect in a multiplayer prisoner's dilemma. Perhaps, in defecting one treats people as a means for one's action rather than treating people as ends. The suggestion is that the individual in the Los Angeles water case reasons to the conclusion to continue to water by not properly taking into account the autonomy of other individuals. But this suggestion fails. Granted, the language of defection' suggests that one is defecting from a group consensus in order to use the group to further one's own aims, but the reasoning of the individual in the water case doesn't proceed like this. The individual in the water case recognizes that regardless of what people do, he can bring about an improvement without harming anyone else. It's entirely consistent with this reasoning that the individual treats each other person as an end. After all, he may very well be concerned to uphold each other individual's dignity by not interfering with their ends. And, of course, there's nothing he can do in the situation to further (or subvert) their ends.

Some people persist in thinking that there is an attenuated sense of 'bring about' in which the individual brings about a worse state of affairs. But, to repeat a point made above, this is a failure to realize the logic of a multiplayer prisoner's dilemma. An individual doesn't affect the social outcome. Further, an individual's change of strategy doesn't affect the probability that the social outcome is achieved. Removing one grain of sand from the beach doesn't affect the probability that the beach has enough sand to make a sand castle. A related but confused objection is that an individual's act of defecting represents a cost itself, so the individual is not faultless for defecting. This objection is confused because it changes the nature of the game. A prisoner's dilemma is a mathematical object that represents moves and preferences. If one thinks that defection itself is a cost, then that should be reflected in the system of preferences. Either

the new game will be a prisoner's dilemma or not. If it is, then the problem of social evil will arise; if it is not, then the problem may not arise.¹⁶

III. Standard Defenses

In this section, I examine standard defenses to determine how they might handle social evil. I argue that social evil is difficult to incorporate within standard theistic responses to evil. I consider three defenses: the value of natural laws, a soul-making defense, and the free will defense. At the outset I should make it clear that I am not merely interested in the logical compatibility of social evil and theism. Alvin Plantinga has conclusively shown that there is no logical problem of evil, and, as I explain below, Plantinga's story can be extended to show that there is no logical incompatibility between theism and the existence of social evil. Even so, granting the compossibility of social evil and theism, the distinctive problem social evil poses for theism remains. What the theist needs to offer is a reason for thinking that God would permit the destructive logic of multiplayer prisoner's dilemmas when we already have enough natural and moral evils to secure the kinds of goods discussed in the standard defenses. The theist need not offer anything amounting to a theodicy, but she should be able to sketch a not completely implausible explanation for why God would allow multiplayer prisoner's dilemmas. In this connection, we can set aside a skeptical theist position.¹⁷ The skeptical theist eschews the task of offering explanations-why for evil and instead focuses on our epistemic limitations to discern a God-justifying good, if one exists. If skeptical theism is plausible, then it has the resources to handle most any evil we are acquainted with. Skeptical theism has been developed with an eye to moral and natural evils, but it can easily be extended to social evil. Even so, the concept of social evil forms a fruitful area of investigation since the extant discussion on the problem of evil focuses on values that are applicable only to natural and moral evils. The problem of social evil should lead the theist to search out new kinds of values that might justify a perfect being in permitting this destructive logic.

A. *The value of natural laws*

Richard Swinburne argues that the problem of natural evil is lessened by the value of natural laws.¹⁸ A universe with recognizable regularities allows persons to successfully predict the consequences of their actions. If I want to help you by offering you nourishment, it is valuable to know that if I offer you bread it won't kill you. Similarly, if I want to study the behavior of atoms in very cold environments, it will help to know that there are reliable ways to bring about freezing temperatures. Because a world with laws is beneficial in this respect, it may be a consequence that some natural evils occur. For instance, a world in

¹⁶Besides the prisoners' dilemma, social evil may arise in assurance games or games of chicken. In general, collective action games provide occasions for social evil.

¹⁷For an explanation and defense of skeptical theism see Bergmann (2001).

¹⁸See Swinburne (2004, 245ff).

which substances have stable properties may imply that in some cases the stable properties of substances harm individuals (e.g., a tree falls and breaks Joe's leg). Or, a world in which environments have stable properties may imply that at boundary zones violent storms occur (e.g., a cold front meets a warm front).

Regardless of what one makes of Swinburne's claim, it is clear that this will not help with social evils. Social evils arise because of the collective agency of rational well-intentioned individuals. This requires that individuals can effectively reason about their options, which in turn requires observable regularities in the world, but that's true of moral evil as well. The crucial difference between social evil and natural evil is that social evil occurs because of human agency. Thus, the appeal to natural laws does not handle social evil.

B. Soul-Making

A different defense appeals to the value of certain kinds of character traits—patience, fortitude, courage, and compassion—and then argues that these valuable traits require evils. Compassion requires suffering; courage requires injustice. The soul-making defense argues that a world that contains the great goods of character must also contain the great evils of character as well. John Hick provides a nice summary of this defense:

The value-judgement that is implicitly being invoked here is that one who has attained to goodness by meeting and eventually mastering temptations, and thus by rightly making responsible choices in concrete situations, is good in a richer and more valuable sense than would be one created *ab initio* in a state either of innocence or of virtue. In the former case, which is that of the actual moral achievements of mankind, the individual's goodness has within it the strength of temptations overcome, a stability based upon an accumulation of right choices, and a positive and responsible character that comes from the investment of costly personal effort.¹⁹

Does the soul-making defense offer promise for handling social evil? An initial hurdle for the soul-making defense is that social evils are not cases in which an individual *brings about* a worse outcome. There is no causal connection between an individual's choice and the outcome that is realized by the group. This verdict is especially clear if we apply the distinction between outcome causation and aspect causation. An outcome cause is a difference maker to whether or not the effect occurs. An aspect cause is a difference maker to the effect *occurring as it does*. In a large multiplayer prisoner's dilemma each individual's action is not a difference-maker to the outcome being achieved or not. It is only in the most attenuated sense that an individual's action is an aspect cause of the effect, i.e., by contributing to the group, the individual causes the effect to occur as it does with $n + 1$ confederates rather than n . So, there is very little room for an individual in such a case to achieve valuable character traits. Perhaps, though, a defender of this move will stress the value of solidarity or the value of self-inflicted loss even when those choices do not have any larger social consequences. But it's doubtful whether this move could be sustained because there is no causal

¹⁹Hick (1977, pp. 255–266).

connection between what an agent does and what social outcome is achieved. If an individual knows that her choices will have no negative effects while also increasing value in the world, it's hard to see how the individual could be blamed for doing what she does.

There is another problem with the soul-making defense. The soul-making defense is deeply individualistic. It focuses on the value of an individual's own character traits. Thus the soul-making defense doesn't have the conceptual resources to explore the value of traits of societies, traits that do not reduce to traits of individuals. Exploring the value of traits of societies seems to be a productive area for reflection on social evil. Perhaps further investigation on the value of certain types of society will uncover reasons God may have to permit social evils. But unfortunately our current discussion of evil is largely informed by individualistic assumptions to the extent that we lack the language and perspective to discuss societal goods and evils.

C. *The Free Will Defense*

The free will defense focuses on the immense value of free will and its irresponsible use in generating horrendous evils. A free will defense is a story about why God may permit moral evil. A crucial part of the free will defense is that persons are responsible for actions that they freely bring about. But as we've seen above, social evil lacks the feature that an individual brings about the evil. Social evils result from the collective agency of individuals, not from any particular individual's choice. Another way to see this is that social evils do not result from any specific individual *misusing* her free will. If an individual in a multiplayer prisoner's dilemma enacts her best option, she is not thereby stupid or wicked. As explained above, individuals in prisoner's dilemmas need not be either irrational or mean. In the water shortage case, I stipulated that each individual was rational and blameless. They each know the situation, and they don't want to harm anyone by their actions. But given the situation each person's best play is defection. In this case, defection is not morally unjust. Thus, the value of an individual's free will doesn't account for social evils because there's no sense to be made of a culpable misuse of free will in this case.

Alvin Plantinga, pursuing a line by Augustine, considers the possibility that natural evil is the result of the free action of non-human spirits.²⁰ It's possible that, as Plantinga says, "Satan rebelled against God and has since been wreaking wherever havoc he can. The result is natural evil."²¹ Given this possibility, the free will defense shows that there is no logical inconsistency between theism and natural evil. In the past, some free agent performed a culpable act and one of the enduring results of that act is natural evil. A similar move may be considered in connection with social evil. It is possible that the occurrence of the destructive logic of a multiplayer prisoner's dilemma is due to the past misuse of free will. Following Plantinga's Augustinian lead, it's possible that "Satan rebelled against

²⁰Plantinga (1977, p. 192).

²¹Ibid.

God and has since been wreaking wherever havoc he can. The result is social evil.”

How plausible is this response? It bears acknowledging that if it is possible that natural evil occurs because of Satan’s activity then it is possible that the occurrence of the destructive logic of multiplayer prisoner’s dilemmas is also the result of Lucifer’s actions. After all, one might contend, wouldn’t the prince of darkness delight in bringing about games in which rational moral individuals collectively produce disaster? Yet I confess that I find this response completely implausible. Given what we know about natural processes—ocean temperatures, weather patterns, plate-tectonics—we know how many natural evils occur. Similarly, given what we know about how society organizes itself—the division of labor, the need for protection and access to resources—we understand how multiplayer prisoner’s dilemmas occur. There is no need to introduce a non-human agent to explain the occurrence of natural and social evils. Consequently, the Augustinian / Plantingian line does not provide a plausible solution to this problem.

A different possibility for the free will defense is that prior to human sin God prevented the occurrence of multiplayer prisoner’s dilemmas.²² Perhaps prior to the misuse of free will there were no scarce resources, or individual’s incentives and options never realized the destructive logic of certain games. But upon the misuse of free will, God removed these protections. So even though social evil presently arises from no wrongdoing, the occurrence of multiplayer prisoner’s dilemma is the tragic result of *previous* sin. This strikes me as the best option for a free-will defense. But it stands in need of further elaboration and defense. One issue is that even if it’s true that God prevented the occurrence of multiplayer prisoner’s dilemmas prior to human sin, we need some additional account of what might justify God in permitting this destructive logic to occur now. As I’ve argued, horrible social outcomes can result from the collective choices of rational well-intentioned individuals, individuals who use their free will responsibly. It seems a harsh consequence for *previous* human sin that the proper use of free will would now result in evil. So the possibility that multiplayer prisoner’s dilemmas arise because of past human sin does not exonerate the free will defense.

IV. Social Evil and Virtue

To this point I’ve argued for the claims that social evil is distinct from natural evil and moral evil and also that standard theistic defenses do not handle social evil. In this section, I have two goals. First, I weaken the assumptions I made regarding social evil. I have argued above that social evil arises by the collective action of rational, well-intentioned individuals in multiplayer prisoner’s dilemmas. This is but one form of social evil. Below, I substantiate the claim that social evil can arise from the collective action of rational *blameworthy* agents in multiplayer prisoner’s dilemmas. This can occur when game-theoretic scenarios make it very

²²Peter van Inwagen (1988) suggests this possibility for natural evils. Thanks to Alex Pruss for extending this suggestion to social evils.

difficult to avoid doing the wrong thing. In these kinds of cases an individual can be strongly tempted to perform an act that, apart from the wider game-theoretic scenario, would be a peccadillo. But because of the destructive game these individuals find themselves in, small sins add up to horrendous evils.

The second goal of this section is to explore the possibility of an Edwardsian response to social evil. Jonathan Edwards claims that true virtue consists in love for being in general.²³ A truly virtuous person does not love merely a limited system—e.g., himself, his family, or his town—but loves every being and seeks the good of all. A truly virtuous person will perform the act that is best for all even if that act requires shunning his own private good. Since God has a reason to bring about truly virtuous people, one might reasonably think that the value of true virtue can provide God with a reason for permitting the game-theoretic machinery that produces social evil. The ensuing discussion will show that the Edwardsian defense solves a two-person prisoner's dilemma but it does not solve a multiplayer prisoner's dilemma because social evil may arise for the truly virtuous.

A. *Immoral man and very immoral society*²⁴

Our initial examination of social evil investigated instances of pain and suffering brought about by the collective action of rational well-intentioned agents. The key feature that distinguishes social evil from natural and moral evil is that the pain and suffering that occurs does not arise from the direct choice of any individual nor from the result of natural processes. It is the cumulative effect of very many choices within the multiplayer prisoner's dilemma that produces pain and suffering. Where each individual enacts their best strategy, the structure of the game the actors are in produces horrendous outcomes.

This description of social evil does not assume that the individuals are blameless. Social evil requires only that game-theoretic machinery produce an amount of pain and suffering that is disproportionate to the individual choices in the game. If a game is realized in which each individual is more self-interested than not and each chooses his best strategy, a horrendous outcome can occur that is not mirrored in the actions of each individual. In these cases the engine of society produces a magnitude of pain and suffering disproportionate to the fuel of individual animosity.

Examples of this kind of social evil are all too common. Russell Hardin in his book *One for All* provides a game-theoretic account of several well-known conflicts: Yugoslavia, Northern Ireland, Somalia, Rwanda, and the nationalist movement in Quebec.²⁵ Each conflict shares a basic structure. The roots of the conflict lay in norms of group identifications, norms that govern group inclusion and exclusion. These norms are important because there are significant benefits

²³Edwards (1960).

²⁴I'm told by a reliable source that Reinhold Niebuhr later thought that this would be a more apt title for his famous book *Moral Man and Immoral Society*.

²⁵See Hardin (1995, pp. 155–177).

realized by being a member of the group. In 17th century England, for example, identification with the aristocracy carried the promise of better jobs, education, and social mobility. But identification requires norms of exclusion, norms that distinguish one group from the other. In 17th century England, one powerful norm of exclusion was the duel; an aristocrat was required on pain of loss of honor to risk injury and death at the smallest of offenses (e.g., disagreement over the merits of John Donne). Once the norms of identification are in place, each individual's incentive structure favors identifying with a group. This leads to competition between groups for access to benefits. Often this competition leads to conflict and even violence. But once violence is realized, this provides the tipping point of group conflict in which escalating reprisals are rewarded. In many cases, this leads to open war. There is, thus, this basic structure: norms of group identification and consequent benefits from identification; competition; conflict; tipping point; system in which reprisals are rewarded; and escalating violence.

The Croatian War of Independence from 1991–1995 provides a case in point of the dreadful game-theoretic situations that produce horrendous suffering.²⁶ Croatia was one of six republics in Yugoslavia, and in 1990, faced with the prospects of Serbian dominance in Yugoslav politics, Croatian leaders decided to move for independence. Croatian Serbs, in addition to being a substantial minority in Croatia, constituted a sizable portion of the Croatian military and police force. When the Croatians moved for independence, they faced the prospects of an internal military revolt from the Croatian Serbs. Regrettably, these Croatian Serbs found themselves in a game in which they could do nothing to ensure loyalty to the new Croatian government. Any promise of loyalty to the new Croatian state would be perceived as subterfuge. In this situation, Croatian leaders preempted the possibility of a fifth column by dismissing Croatian Serbs from military and police positions. This move, though favored by the game the Croatians found themselves in, created the tipping point for internal Serbian dissent and led to revolt. The ensuing series of moves and countermoves resulted in the devastating Croatian civil war.²⁷

This story highlights the awful logic of some games. Further, this story doesn't assume that each actor is morally in the clear. Arguably, events around the tipping point are replete with bad intentions. But the bad intentions and wrongful actions do not themselves explain the descent into open war and the carnage that followed. In this case we can clearly see the effects of game-theoretic scenarios in producing horrendous pain and suffering.

In the discussion of the first several sections, I focused on pain and suffering that resulted from the collective action of rational moral agents. Now we have seen that social evil can result when the game-theoretic scenario magnifies the consequences of individual blameworthy action. Sometimes pain and suffering

²⁶This paragraph is a condensed summary of Hardin's analysis (see pp. 156–163).

²⁷Hardin also notes that the Croatian move for independence made group cooperation in Bosnia "virtually impossible" (p. 159).

result from collective action even though no one individual has done anything amiss. This is the first kind of social evil. But in other cases, social evil results from collective action when the individuals are to blame for their acts and yet the social situation they are in amplifies the effects of those wrong acts. This is the second form of social evil.

B. An Edwardsian Defense of Social Evil

Jonathan Edwards, in his book *The Nature of True Virtue*, argues that true virtue is love for being in general. The truly virtuous person does not seek to benefit only a limited group of participants but rather seeks the good of every being. Edwards' discussion on true virtue includes a penetrating exposition on the consequences of self-love. He offers what is, in effect, a game-theoretic explanation of apparently moral behavior. He observes that a person's self-interest will motivate them to appear to be altruistic. Consider giving to the needy. Edwards argues that this act may often arise from a limited benevolence. One can be concerned only with the needy in one's own town, or one can give to the needy to identify with the "moral" crowd and then reap the benefits of being a member of that crowd. In this case, adopting Hardin's terminology, sacrificial giving is a norm of exclusion. Edwards underscores that any such act does not arise from true virtue. True virtue consists in love for being in general.

Given Edwards' emphasis on the high demands of true virtue together with the claim that a perfect being desires to bring about truly virtuous individuals, there may be the makings here for an Edwardsian defense of social evil. A case can be made for an Edwardsian defense, but, if my analysis is correct, an Edwardsian defense answers only social evil arising from a two-person prisoner's dilemma; it does not extend to a general solution to the multiplayer prisoner's dilemma. Even though the Edwardsian defense is unsuccessful, it holds important lessons for reflection on social evil.

I begin by examining an Edwardsian solution of a two-person prisoner's dilemma. I will assume that the truly virtuous person will favor the option with highest social utility. In a two-person prisoner's dilemma, I will take social utility to be the sum of the individual preferences. This way of generating a social utility makes the most sense under the assumption that the preferences represent *strength of preference* rather than order of preference. Since nothing hangs on this change, in the following game the reader may assume we are summing across preference strengths. In the matrix below I put the social utility in parentheses. Consider the following game.

Two-Person PD with social utility

By inspection we see that the option with the highest social utility is the one in which everyone cooperates. This option achieves 6 units of social utility. What does this revised game look like for the truly virtuous person? Let us assume that Player I is truly virtuous. Player I faces the following choice: defect or cooperate. If Player I defects, then he realizes either the social utility of 4 (in

		Player II	
		Defect	Cooperate
Player I	Defect	2,2 (4)	4,1 (5)
	Cooperate	1,4 (5)	3,3 (6)

which Player II defects) or the social utility of 5 (in which Player II cooperates). In either case, Player I would realize more social utility by cooperating. For the truly virtuous person, cooperating is the dominant action. In this way, the truly virtuous person solves the two-person prisoner’s dilemma.

Does the Edwardsian position solve the general multiplayer prisoner’s dilemma? No. To see this I will first describe the general form of a multiplayer prisoner’s dilemma and then describe an instance of this for the truly virtuous. This shows that multiplayer prisoner’s dilemmas can arise for the truly virtuous. Consequently, true virtue doesn’t lead to the dissolution of multiplayer prisoner’s dilemmas.

The multiplayer prisoner’s dilemma I describe below assumes that there is a vague threshold for cooperation required to achieve the social good and avoid the social evil. As I argued above in the water shortage cases, it is eminently reasonable that the threshold for providing adequate water supply is vague. Any change of one person’s strategy will not make a difference to the outcome achieved. Standard representations of multiplayer prisoner’s dilemmas assume that the threshold for cooperation is exact. To provide a suitable matrix for the game I focus on three states: the state that doesn’t meet the threshold, the state that is at or exceeds the threshold, and the penumbral state. In the penumbral state, it is unclear whether the social good will be achieved or not. I assume that in the penumbral state there is enough cooperation to achieve at least some (but not full) benefit. When the threshold of cooperation is met then the full benefit is achieved.

In a multiplayer prisoner’s dilemma each player faces the same matrix. The table below represents the matrix for a random player i.

Multiplayer Prisoner’s Dilemma with a vague threshold

What should player i do? As before with the original two-person prisoner’s dilemma defection is the dominant strategy. If player i cooperates, then, whatever the other players do, i will do better by defecting. Inspect the above table: if player i cooperates then if the threshold isn’t met, i does better by defecting; if the penumbral state is realized, i does better by defecting; and if the threshold is exceeded, i does better by defecting. The reason this situation arises is that each individual can bring about a Pareto improvement of resources by defecting. Since an individual doesn’t affect the social outcome but does affect some item of value in the world, each individual faces the prospect of a wasted sacrifice and a loss of some item of value by cooperating. Perhaps, though, if all players

		<i>Below the Threshold of Cooperation</i>	<i>The penumbral state</i>	<i>At or above the threshold of Cooperation</i>
Player i	Defect	No cost & no benefit Ordinal rank: 2	No cost & some benefit Ordinal rank: 4	No cost & full benefit Ordinal rank: 6
	Cooperate	Cost & no benefit Ordinal rank: 1	Cost & some benefit Ordinal rank: 3	Cost & full benefit Ordinal rank: 5

were truly virtuous we could avoid this disastrous game. In the following, I argue against this by providing a multiplayer prisoner's dilemma for the saints.

Suppose ten saints each manage their own orphanage. Christmas time is approaching and it is time to solicit gifts. Each orphanage draws upon a common pool of resources that in normal times is not adequate to meet every need but it is just enough for most. Each orphanage sends out letters to the community asking for donations. Now each saint faces the following decision. Should I *cooperate* by sending out only letters and relying on the donations generated by those letters, or should I *defect* by sending out letters and then making personal phone calls? The strategy to defect should not be thought of as a departure from true virtue. In the case at hand, defection amounts to making that extra call to ensure that Johnny gets the red fire truck he wants. It's implausible to think that the saint that defects from the group is acting out of anything but love. But if the saints each defect what they collectively bring about is a race to request charity. In addition to raising the required effort each year on the part of the saints, it is likely to lead to less overall charity. Besieged by letters, emails, phone calls, and visits, people grow weary of the increasing intrusions. Thus, the collective effect of departure from the standard is to realize an overall worse situation. However, whatever each individual saint does will not affect the overall overcome. If Theresa, at the last minute, makes that extra phone call she will bring about a better situation—Johnny gets the fire truck and all the giving is as it would otherwise be.

The orphanage case illustrates the dreadful logic of a multiplayer prisoner's dilemma. Even if each player is truly virtuous, the painful consequences of enacting each player's best option will not be avoided. Ultimately, the reason the Edwardsian defense fails is that each truly virtuous person can know that her act will bring about a better world but the collective result of each person making a better world is unintended disaster. Even with the truly virtuous, the road to hell is paved with Pareto improvements.

Conclusion

Social evil, pain or suffering resulting from the collective agency of rational individuals in multiplayer prisoner's dilemmas, is a pervasive feature of our world. Yet traditional defenses do not address it. Traditional defenses have focused exclusively on pain or suffering that results from either natural processes or from moral agency. Theists should view this problem as an opportunity to further mine the conceptual resources of theism. Additionally, social evil provides strong motivation for everyone to be concerned about the structures of society. We should aim for societal structures that minimize this dreadful logic.

REFERENCES

- Adams, Robert. 1985. "Involuntary Sins." *Philosophical Review* 94: 3–31.
- Alston, William. 1991. "The Inductive Argument from Evil and the Human Cognitive Condition." *Philosophical Perspectives* 5: 29–67.
- Bergmann, Michael. 2001. "Skeptical Theism and Rowe's New Evidential Argument from Evil." *Noûs* 35: 278–296.
- Binmore, Kenneth. 1994. *Playing Fair: Game Theory and the Social Contract*. Cambridge: MIT Press.
- Edwards, Jonathan. 1960. *The Nature of True Virtue*. Ann Arbor, MI: University of Michigan Press.
- Hardin, Garrett. 1968. "The Tragedy of the Commons." *Science* 162: 1243–1248.
- Hardin, Russell. 1995. *One for All: The Logic of Group Conflict*. Princeton, NJ: Princeton University Press.
- Hick, John. 1977. *Evil and the God of Love*. New York: Harper Collins.
- Hofstadter, Douglass. 1985. *Metamagical Themas*. New York: Basic Books.
- Kuhn, Stephen. 2009. "Prisoner's Dilemma." *Stanford Encyclopedia of Philosophy* <http://plato.stanford.edu/archives/spr2009/entries/prisoner-dilemma/>.
- Lewis, David. 1979. "Prisoner's Dilemma is a Newcomb's Problem." *Philosophy & Public Affairs* 8:3: 235–240.
- Niebuhr, Reinhold. 1932. *Moral Man and Immoral Society*. New York: Charles Scribner's Sons.
- Plantinga, Alvin. 1977. *God, Freedom, and Evil*. Grand Rapids, MI: William B. Eerdmans Publishing Co.
- Rowe, William. 1979. "The Problem of Evil and Some Varieties of Atheism." *American Philosophical Quarterly* 16: 335–341.
- Rowe, William. 1988. "Evil and Theodicy." *Philosophical Topics* 16.2: 119–132.
- Swinburne, Richard. 2004. *The Existence of God*. New York: Oxford University Press. 2nd edn.
- van Inwagen, Peter. 1988. "The Magnitude, Duration, and Distribution of Evil: A Theodicy." *Philosophical Topics* 16.2: 161–187.
- van Inwagen, Peter. 2006. *The Problem of Evil*. Oxford: Clarendon Press.
- Williamson, Timothy. 1994. *Vagueness*. New York: Routledge.

THE ARBITRARINESS OF THE PRIMAL SIN

Timpe

Abstract This paper focuses on the first sin, often referred to as the primal sin. In particular, it contrasts a voluntarist account of the primal sin with an intellectualist account. One frequent objection to voluntarist accounts of free will is that they leave an agent's choice fundamentally inexplicable or arbitrary. I argue that with respect to explain the primal sin, intellectualism fails no better as both accounts involve an unresolved arbitrariness. This suggests that, at least for theists, voluntarist accounts of free will are not as problematic as many assume.

Who made me? Is not my God not only good but the supreme Good? Why then have I the power to will evil and to reject good? Is it to provide a reason why it is possible for me to undergo punishment? . . . If the devil was responsible, where did the devil come from? And if even he began as a good angel and become devil by a perversion of the will, how does the evil will by which he became devil originate in him, when an Angel is wholly made by a Creator who is pure goodness? Saint Augustine¹

§1 Introduction

It is common, especially among medieval treatments of free will, to distinguish voluntarist accounts from intellectualist ones. Colleen McCluskey differentiates the two approaches as follows:

In light of a common theory of human psychology, the medieval debate centered upon whether human beings act freely primarily in virtue of their wills or in virtue of their intellects. Those who argue that freedom is primarily a function of the intellect are known as intellectualists while those who argue that freedom is primarily a function of the will are known as voluntarists, from the Latin word for will, *voluntas*.²

While both of these approaches have their share of proponents, Robert Brown claims that the voluntarist approach is the more common in the history of Christian theology: "The early Christian tradition, after several centuries of vigorous debate, chose to locate the source of human evil exclusively in the exercise of will rather than the deficiencies of intellect."³ This distinction is not restricted just to medieval writers, as one can find defenders of both approaches among contemporary philosophers. For example, Eleonore Stump's Thomistic inspired view

¹Augustine (1993, p. 11f, (7.3.5))

²McClusky (2007). See also Brian Leftow (1998).

³Brown (1978).

is perhaps not surprisingly an intellectualist account⁴, while many agent-causal accounts, such as Timothy O'Connor's, have voluntarist leanings.⁵

A common objection to voluntarist accounts is that they leave the exercise of an agent's free will as inexplicable or incomprehensible. In what follows, I evaluate voluntarism and intellectualism regarding their explanatory strength with respect to the primal sin. (My reasons for focusing on this particular sin will be made clear in the next section.) It turns out within such a theological context, the explanatory benefit of intellectualism is significantly less than critics of voluntarism often suggest. Considerations of the primal sin show that both voluntarist and intellectual accounts involve an unresolved arbitrariness at the heart of their accounts of free agency. This suggests that, at least for theists, voluntarist accounts are not as problematic as many believe them to be.

§2 Primal Sin

It need not be said that Augustine was perplexed by evil throughout his life. One author has recently written that "Augustine's account of the problem of evil came in the end to embrace almost every other area of his writing."⁶ It was the attempt to reconcile evil with the existence of God, for example, that constituted one of the main reasons for his nine-year affiliation with Manichaeism.⁷ And subsequent to his conversion to Christianity, Augustine's reflection on the nature and origin of evil produced the early work *On Free Choice of the Will*, as well as some of the most alluring passages of the latter *Confessions* and *The City of God*. Augustine played a key role in developing and popularizing what is now commonly referred to as the free will defense.⁸ According to Augustine:

We have determined that the choice to follow and embrace one or the other [of good or evil] lies with the will, and that only the will can depose the mind from its stronghold of power and deprive it of right order. And it has become clear that we not blame anything when someone uses it wrongly; we should blame the one who uses it wrongly. . . . We are now in a position to ask whether evil doing is anything other than neglecting eternal things, which the mind perceives and enjoys by means of itself and which it cannot lose if it loves them; and instead pursuing temporal things which are perceived by means of

⁴See, for instance, Stump (2003, 1990).

⁵See, for instance, O'Connor (2000).

⁶G. R. Evans (1999, p. 340).

⁷The Manicheans, as well as gnostics in general, thought that reflection on primal sin gave credence to their own dualism by posing the following dilemma: "Was Adam created perfect or imperfect? If perfect, then how could he fall? If imperfect, why did a perfect God create an imperfect being? Would not God [then] be ultimately responsible for the fall?" (Charles Seymour (1999, p. 256)). The present paper attempts to provide an answer to the second of these questions.

⁸The locus classicus for the free will defense to the logical problem of evil is Plantinga (1977). For a more elaborate, and technical, discussion of the same issues, see also Plantinga (1974).

More precisely, Augustine was likely giving a free will theodicy. However, insofar as giving the actual reason God has for allowing evil entails giving a possible reason for the same, it is not inappropriate to describe Augustine's project as a defense.

the body, the least valuable part of a human being, and which can never be certain as if they were great and marvelous things. It seems to me that all evil deeds that is, all sins fall into this one category.⁹

However, as the quotation at the start of the chapter indicates, even after developing the free will defense, Augustine was still perplexed by the origin of evil given the role of an essentially good God who created the world in general, and free creatures more specifically. Given his nature, God wouldn't create anything that was intrinsically evil. But if all the agents that God creates are created good, how is it that at least some of them do evil? The answer, of course, is that an agent wills to do evil, but this answer immediately raises another question: where does the evil will originate, particularly if the reasons-constraint on free choice is true? A creature's free will in the status integritatis might account for the possibility of evil, but that isn't sufficient to explain why that possibility was actualized by an otherwise morally untarnished creature. According to William Babcock, this is a question that "Augustine never fully escaped nor finally solved."¹⁰ In the present chapter, I seek to address how an agent that is created as all good could nonetheless will to do evil, thus dealing with the transition from the status integritatis to the status corruptionis.

According to a common strand prevalent throughout much of the Christian tradition, the devil was an angel and the first creature to sin.¹¹ This sin has traditionally been seen as pride, though others have held it to be primarily a sin of envy.¹² The Catechism of the Catholic Church, for example, puts it as follows: Behind the disobedient choice of our first [human] parents lurks a seductive voice, opposed to God. . . . Scripture and the Church's Tradition see in this being a fallen angel, called "Satan" or the "devil". The Church teaches that Satan was at first a good angel, made by God: "The devil and the other demons were indeed created naturally good by God, but they became evil by their own doing." Scripture speaks of a sin of these angels.

This "fall" consists in the free choice of these created spirits, who radically and irrevocably rejected God and his reign.¹³

⁹Augustine (1993, p. 27).

¹⁰Babcock (1988, p. 30).

¹¹In discussion Augustine's account of primal sin, which we'll return to below, Scott MacDonald describes it as "an idea without which he [Augustine] thinks no account of sin can be complete: the idea that imitation of God in the form of prideful self-assertion is at the bottom of all sin" (MacDonald 2003, p. 408). For a very interesting discussion of the origin of the fall in sixteenth- and seventeenth-century theology, see Steadman (1960).

¹²For discussions of the primal sin as pride, see the entry on the devil in Fitzgerald (1999, p. 268) and Aquinas (1948, 63.7) and *De Doctrina*, in Milton (1931, XV, 181), and Lewis (1940, pp. 69ff). For a discussion of the minority position which roots the primal sin in envy, see King (2012).

¹³2nd edition, paragraphs 391 and 392; footnotes omitted. For an account of sin originating in the free will of the devil from a very different theological perspective, see Gregory Boyd (2001): "The New Testament repeatedly identifies the originator and head of the rebellion against God as Satan. Indeed, although it does not locate the entire responsibility for all evil on Satan, it does trace all evil back to him. . . . Second, and closely related to this, because Scripture depicts Satan as being far more powerful than any of the demonic or human agents

Following Scott MacDonald, I will refer to this first purported case of sin as the primal sin':

The fall of the angels constitutes the paradigm case [of evil-originating free choice] since, unlike Adam's and Eve's sin in the garden, the first angelic sin is entirely unprecedented. We can think of that first evil free choice as constituting primal sin. The first sin deserves to be called primal, however, not just because it is temporally first but also because it is something radically new in creation: the first evil appears against a backdrop of utter goodness. All things created by God, including the rational creatures whose free choices are the original evils, are wholly good and without flaw. . . . There can be no context of defect or corruption into which the first sin fits. Good creatures with good wills voluntarily introduce evil into a world where there was none before. Primal sin is not only unprecedented but also seemingly unprepared for and unprompted.¹⁴

Focus on the primal sin thus makes the discussion of the relationship between free will and sin a bit cleaner than it would be otherwise, particularly given the dynamic relationship between an agent's moral character and her choices discussed at the end of the previous chapter. For every sin other than the primal sin, that choice to sin will have in its causal antecedents the causal consequence of an earlier sin.¹⁵ To take an example, consider a particular sinful act by gluttonous Gene; say, that on a given occasion, he eats an entire carton of Moose Tracks ice-cream and then, despite knowing better, gets into the freezer for yet even more.¹⁶ Given that he is aware that his motivational reasons do not align with the normative reasons, part of the explanation for this choice will be the thoroughly gluttonous disposition Gene has fostered over the previous years. But an investigation of primal sin will avoid these complications arising from previous sinful choices and a bad moral character, as well as the further issues of the bondage to sin and original sin. Katherin Rogers makes a similar point in a recent study of Anselm's account of freedom:

I have been discussing Anselm's theories regarding human free will. In fact, much of his central argument occurs in *De casu diabolic*. The contemporary reader may find it odd

that are under him, he represents the ultimate challenge for our theodicy. The challenge of explaining how God could create beings who can resist his will and genuinely war against him is epitomized in Satan. If we can account for his existence, we shall have thereby accounted for the existence of all lesser evil against" (17).

¹⁴MacDonald (1998, 110f).

¹⁵This will be particularly true with respect to human freedom if one takes seriously the doctrine of original sin. For a very useful philosophical discussion of this doctrine, see Rea (2007) and Wyma (2004). For more on the distinction between actual and original sin, see Quinn (1998).

¹⁶Jonny Brown has suggested that the use of a fairly pedestrian sin' such as that involved in the example of gluttonous Gene, to which we'll return later, may inadvertently trivialize the seriousness of the subject-matter. There is a growing body of literature which does suggest that the level of moral seriousness in an example can greatly impact one's intuitions regarding freedom and moral responsibility, as can the difference between an actual example and a hypothetical example; for an overview of these data, see Sommers (2010), particularly section 3. While I am sensitive to this data, I do not think the use of the present example is infelicitous. For one, the use of a more serious example may illicit more compatibilist intuitions. Insofar as the present paper is working on the assumption of incompatibilism, I want to minimize that tendency.

that Anselm would focus on the will of Satan. Whatever one's view of the Heavenly Host and their fallen brethren, it should be appreciated that Anselm chooses to discuss the fall of the devil from the best analytic motives. He is interested only in morally significant choice, and he is deeply concerned to get to the bare metaphysics of free will. He prefers to set aside instances of choice where the core act is difficult to discern, being encrusted with layers of competing desires born of years of lived history. He wants to examine a pure instance of choice, and he wants to put the central and most difficult puzzle of created freedom in the starkest terms: how could a being made perfectly good, with no one and nothing already evil in the world to tempt him, possibly choose against the will of God?¹⁷

Rogers' comments also suggest a reason for not focusing on the fall of humans from the status *integritatis*, as according to traditional theology their fall was influenced by agents who had already fallen. My focus here on the fall of the devil is then, in one sense, merely a placeholder for whatever the first temporal sin was; I am not doing what Robert Brown has referred to as "theological history."¹⁸ As such, I do not intend the following discussion to have purchase only for those who believe in the literal existence of the devil, and his role in the fall of the human race.¹⁹

There is also one additional reason for focusing on the primal sin. If a satisfactory account of the primal sin can be given, then that account plus the causal contribution to other sins that it makes possible should also go a significant distance towards providing an understanding of how other subsequent sins are possible in the status *corruptionis*. Thus, to quote MacDonald, the primal sin whatever they may be "are important not merely because of their temporal priority and causal significance but also because they are in a certain way paradigms."²⁰ But, as we'll see below, the issue of primal sin also appears, at least initially, to be a philosophical conundrum "a problem all of whose possible solutions are unsatisfactory."²¹ And if the primal sin can be shown to only appear to be a conundrum, then the same would also be true of sins in general.

¹⁷Rogers (2008, 7).

¹⁸Brown (1978, 319).

¹⁹Whatever the other specifics of one's theology, it seems that Christian orthodoxy commits one to believe in a temporally first sin given the Christian doctrine of God's creation. Furthermore, there will also be a temporally first sin of a human, regardless of whether one interprets the second chapter of Genesis literally, as the following passage from Swinburne makes clear:

At some state in the history of the world, there appeared the first creature with hominoid body who had some understanding of the difference between the morally obligatory, the morally permissible (i.e., right), and the morally wrong; and an ability freely to choose the morally right. So much is obvious; since on modern evolutionary views, as well as on all views held in Christian tradition, once upon a time there were no such creatures and now there are some, there must have been a first one. It seems reasonable to consider such a creature the first man; and we may follow biblical tradition and call him Adam'. (The Hebrew word means man'.) (Swinburne (1989, 141)).

Shane Glackin has suggested that an evolutionary account, there will be vagueness with respect to not only human nature, but also responsibility and sin. I'm inclined, however, to think that that vagueness will be epistemic in nature, and not metaphysical. Even if I'm wrong, the implications of vagueness regarding human nature and moral responsibility must be left for another time.

²⁰MacDonald (2003, 397).

²¹O'Connor (2005, 207).

In the present paper, I'm interested in comparing this paradigm sin from voluntarism and intellectualist approaches. I'll approach the issue of primal sin by looking at the two most developed extant accounts of it in the contemporary literature. Both accounts are libertarian accounts insofar as they suppose that the truth of theological determinism would render the devil unfree, and thus not responsible, in his fall. Furthermore, both accounts are inspired by medieval theologians, though they aim to provide satisfactory philosophical accounts of the primal sin and not be mere historical exegesis.

§3 Voluntarism: Katherin Rogers on Anselm

I'll approach the issue of primal sin by looking at the two most developed extant accounts of it in the contemporary literature. Both accounts are libertarian accounts insofar as they suppose that the truth of theological determinism would render the devil unfree, and thus not responsible, in his fall. Furthermore, both accounts are inspired by medieval theologians, though they aim to provide satisfactory philosophical accounts of the primal sin and not be mere historical exegesis.

Having already mentioned Rogers' discussion of the primal sin in her study of Anselm's account of freedom, I'll begin there.²² Rogers' treatment of the issue is embedded within the larger context of her discussion of Anselm's departure from Augustine on the nature of free will, and presupposes a number of interpretive issues.

[Anselm's] basic metaphysics and epistemology are solidly Augustinian. And yet Anselm did not agree with everything the Master had said. He does not say it explicitly, but it is clear that his work on free will is motivated by a fundamental problem which he finds in the work of Augustine. It is in his book *De casu diabolico* that he comes to grips with the basic mechanics of created freedom, and the question which drives the work is this: how could Satan, created perfectly good, choose to sin? Anselm's student interlocutor spells out the standard answer he has heard. Satan sinned because he chose not to persevere in the good will which God had given to him at creation, and the reason he did not persevere was that God had not given him the necessary perseverance. Though Anselm does not cite the source, this is Augustine's position. But in Anselm's view this answer is, at best, radically incomplete. How, if Satan could not help but fall without the God-given perseverance, can we avoid the conclusion that God is responsible for the sin?²³

As this quotation indicates, Rogers argues that Augustine's view is best understood as a form of compatibilism. More specifically, she thinks that "Augustine is a compatibilist from his earliest work on freedom through his anti-Pelagian writings, and the freedom possessed by the un-fallen and fallen will is a compatibilist freedom."²⁴ While this is a controversial claim, it need not concern us here insofar as Augustine's scholarship is not my primary goal.²⁵ What is less controversial is that Anselm was a

²²Rogers seems to endorse Anselm's view insofar as she describes him as "offer[ing] viable solutions to some of the puzzles which have plagued Christian philosophers since the days of Augustine and which are still hotly debated today" (Rogers 2008, 1). In what follows, I shall interpret Rogers to be endorsing the Anselmian view that she elaborates.

²³Rogers (2008, 31). See also 93ff.

²⁴Rogers (2004, 415).

²⁵For a discussion of whether or not Augustine is best interpreted as a compatibilist or incompatibilist, see chapter 1, note 39.

libertarian regarding free will.²⁶ And it is precisely Anselm's rejection of Augustine's compatibilism (as Rogers interprets him) and his embracing of theological determinism which underscores Anselm's problem with Augustine's account of the primal sin. She continues:

As Anselm recognized, Augustine's analysis of the choice of the unfallen will, the will in its ideal condition, raises exactly the same problem as his view of saving grace. According to Augustine, the created will chooses on the basis of what it most desires.

There is nothing in Augustine's work to suggest that on this most fundamental point about the working of the will he distinguished between the pre-lapsarian and the post-lapsarian condition. But everything about the creature, including its knowledge, will, and desires, and everything about his situation, including whatever can be a possible object of desire, are from God. Thus, as I shall argue, on Augustine's understanding, God is not only the architect of the original situation in which the created agent finds himself, He also controls the outcome. But then God is responsible for created choices, even in the beginning when the will is in its original, pristine condition. The upshot seems to be, as Anselm acknowledges in *De casu diaboli*, that the very first choice for evil can be traced to God.²⁷

Rogers thinks that the choice for sin, on the assumption of both theological determinism and compatibilism, is either "unintelligible"²⁸ or leads to the conclusion that "the responsibility for the original evil lies with God,"²⁹ an evaluation with which I am inclined to agree. Instead, Anselm's account of the primal sin is thoroughly incompatibilist in nature:

God does not cause sin, nor does He bear the ultimate responsibility for it as something He could and should have prevented. The source of sin is the created agent. Given Anselm's analysis of what free choice means one can only be unjust under one's own steam. As he writes in *De casu diaboli*, I think you realize that God cannot cause [one to be] unjust in any way at all, unless it is by not causing [one to be] just, when He could do so. Before having received justice, no one is just or unjust, and after having received justice no one becomes unjust except through abandoning justice on their own (sponte). . . . But if God is not the cause of sin, then the rational creature must be a primary agent. Choice must in some way originate in the creature.³⁰

Rogers' reconstruction of Anselm's metaphysics of free will is complex, and many of the details are unnecessary for present purposes. But it will be helpful to have in mind the two inclinations (or affectiones) that Anselm thinks are necessary if a will is to be free: the desire for benefit and the desire for justice.³¹ These are not two exclusive objects of desire; rather they differ in terms of being different orders of desires, in a way reflective of Harry Frankfurt's more recent discussion of first- and second-order desires. The desire for benefit is the desire

²⁶See, for instance, Visser and Williams (2008).

²⁷Rogers (2008, 32).

²⁸Rogers (2008, 47).

²⁹Rogers (2008, 51).

³⁰Rogers (2008, 92f). Rogers later says that "the first sins of Adam and Eve were in essence the same as the sin of the devil" (129).

³¹Justice' here is taken to refer primarily to the cardinal virtue, understood as rectitude of will preserved for its own sake.'

for things the possession of which one thinks will lead to her happiness; the desire for benefit is thus a first order desire. The desire for justice, in contrast, is “a desire for rightness of will preserved for its own sake’. It is therefore a second order desire that one’s first order desires should be properly ordered, should be as they ought to be.”³² Both kinds of desires, however, provide the agent with motivational reasons for acting. Sin in general occurs when an individual pursues her desire for benefit in a way that is not properly ordered, that is, in a way that contradicts the normative reason of justice:

In order to allow created freedom, God bestows upon the created agent the two affections. Thus morally significant choice consists in a struggle with the agent, due to the conflict between the desire for the inappropriate benefit, and the desire for justice which would lead him to endorse only the appropriate desires. Preceding the final decision they are, as it were, two streams of desire competing for ascendancy. Or, to put it another way, the agent is trying to pursue two desires, where ultimate success regarding one entails the abandonment of the other. Sin occurs when the agent succeeds’ in following the desire for the inappropriate benefit.³³

Anselm, like the majority of the medievals, does not think that anyone wills injustice (or any other evil) for its own sake; rather, they will it under the description of something beneficial: “Anselm explains that the injustice of the bad angel consists not in willing injustice per se, but in willing benefits which he should not have willed. . . . There is absolutely nothing intrinsically bad or tainted about the forbidden benefit. It is not some selfish advantage intrinsically opposed to justice.”³⁴

This, of course, raises the question why would the devil choose to forgo a higher good, here justice, and pursue a lower good, here benefit?³⁵ Insofar as we are focused at present on the primal sin, the answer cannot be original sin, a previous sin, or a corrupted moral character. Neither is it simply a result of ignorance, such as being unaware that one’s motivational reasons fail to track the normative reasons, for on Anselm’s view “the devil must know that he ought not to will the inappropriate benefit at that time.”³⁶ Thus he thinks the primal sin is not grounded in ignorance, but in an active and informed choice. But is such a choice explicable? According to Rogers, in one sense it is not:

But why did he will what he ought not?’ asks the student. No cause preceded this will, unless it was that he was able to will.’ But this ability per se is not really the cause, since the good angels were equally able to desert justice. Why then did he will?’ The teacher responds, Only because he willed. For this choice had no other cause by which it was by any means impelled or drawn, but it was its own efficient cause, and

³²Rogers (2008, 67).

³³Rogers (2008, 118).

³⁴Rogers (2008, 67f).

³⁵Augustine puts it this way: “That angel [Lucifer], delighting in himself rather than in God, was unwilling to be subject to Him and swollen with pride: he abandoned the Highest Essence, and he fell” (as quoted in King (2012), page 6 in manuscript.). Similarly, according to the Catechism of the Catholic Church, “In that sin man preferred himself to God and by that very act scorned him. He chose himself over and against God, against the requirements of his creaturely status and therefore against his own good” (398). This raises the parallel question: why would a non-fallen human would choose the self over God.

³⁶Rogers (2008, 96).

effect, if such a thing can be said.' Here we have libertarianism of the self-causation variety stated with brutal clarity and with no attempt to downplay its core problem. . . . Anselm does not go on and try to mitigate the problem of intelligibility.³⁷

Rogers is well aware that this answer will raise questions of the coherence of Anselm's view: Like the student at the end of the *De casu diaboli* we still want to know what explains the beliefs and character of the agent that explains the preference for one option over another. If there is absolutely nothing about the beliefs and character of the agent that causes the preference, then . . . isn't the choice more like an accident that happened to the agent rather than a action for which he can be held responsible? . . . But if nothing about the agent determines that he will make one choice rather than the other, isn't the choice just a piece of luck? How can responsibility be grounded in luck? And remember that in Anselm's universe, Satan's bad luck', if that is what it is, results in eternal damnation.³⁸

Another way to put the same question is to ask for what is often called a contrastive explanation: what is that explains why the devil chose to sin rather than to not sin?³⁹ Furthermore, it seems to many that if a contrastive explanation cannot be given, then the account of free action fails. Rogers is willing to grant that there is no contrastive explanation; as a result, she agrees that there is ultimately something "inexplicable and ultimately mysterious" about the primal sin.⁴⁰ She elaborates as follows:

We can point to the reasons for choosing either option. But there is no antecedent cause or explanation for the preference of the one over the other. It does not really help to add that the reasons for one choice outweigh the reasons for the other by choosing, since the outweighing comes after the choosing. We still want to know why the devil chose sin over justice. And Anselm's answer is, only because he chose. There is no more to be said, and this is an uncomfortable stopping point for it seems to grant that the intelligibility problem is not entirely soluble.⁴¹

But Rogers denies that the inability to give a contrastive explanation undermines the cogency of the account under consideration.

I take it that Anselm sees and is willing to accept that there is a certain mystery at the core of free choice. But perhaps this need not count as a criticism of his theory. Anselm is extremely optimistic about the scope and range of the

³⁷Rogers (2008, 97).

³⁸Rogers (2008, 98f).

³⁹See, for instance, Clarke (1996), especially 192ff. Note that Rogers thinks the call for a contrastive explanation is different from the luck objection to libertarianism; see Rogers (2008) p. 99, footnote 32. Furthermore, on this point the problem facing an incompatibilist account of primal sin is stronger than the objection that incompatibilist accounts of free will cannot account of contrastive reasons in general. A number of incompatibilists have sought to address this latter problem; see, for instance, Kane (1996) and Hitchcock (1999) and Ginet (1989). But even if the libertarian can give an account of acting for reasons in general, there still remains the problem of specifying what reason a non-fallen agent would have for choosing to sin.

⁴⁰Rogers (2008, 87). However, on Rogers' view, this can also be true of other libertarian free choices, and not just the primal sin. See also Rogers (1997, 10)

⁴¹Rogers (2008, 104).

human intellect, but his subject-matter, God and the relationship of the created to the Creator, leads him to assume that he will run up against issues that he cannot divine [or divide?] and conquer. . . . In the *Cur deus homo*, where he argues that if reason has concluded that something is the case, the investigator ought to take it as at least provisionally proven, even if he cannot grasp how it is the case. Created freedom seems to be one of those instances. There is sin. God does not cause it. Therefore it originates in the created will, although this seems a mystery.⁴²

In this passage, Rogers appears to be attributing the mystery involved in why the devil would choose to sin to an epistemic failing on our part; that is, she seems to be suggesting that given our finitude, we are simply unable to discern what the contrastive explanation is. But shortly thereafter, she argues for a stronger conclusion, and that is that there really is no contrastive explanation that can be given: “such a thing is impossible.”⁴³ What it means for a creature to be created in the *imago dei* is for it to be both free and responsible, to have a certain degree of aseity and to be, with respect to its moral character but not its existence, a self-creator. “God has constructed the system so that the rational creature can, in however limited a way, mirror this divine aseity by contributing to its own being. It is a dim reflection of its Creator, but it is a true one in that, through free choice, it participates in its own creation.”⁴⁴ Suppose that Rogers is right about the implications of being created *imago dei*. And suppose that she is also right that the inability to give a successful contrastive explanation is not fatal, in general, to the libertarian’s view of free will. One can agree with her up to this point and still think that the account of the primal sin she develops is wanting because it seems that, at the end of the day, Rogers’ Anselmian inspired answer to the question of why the devil fell in choosing the desire for perceived benefit over the desire for justice is a brute just because. So even if her account can survive some of the challenges it faces, we are still left with an unsatisfactory account of *prima sin* insofar as it is at root, not just unexplained, but inexplicable. According to Robert Brown, it couldn’t be otherwise; an account of the primal sin “must be incomprehensible” and “inexplicable.”⁴⁵ Any attempt to give such an account is “a conceptual blunder.”⁴⁶

§4 Intellectualism: Scott MacDonald on Augustine

In a pair of recent papers, Scott MacDonald develops an account of the primal sin inspired by Augustine. Though Augustine is usually described as a voluntarist

⁴²Rogers (2008, 105).

⁴³Rogers (2008, 107).

⁴⁴Rogers (2008, 106).

⁴⁵Brown (1978) 315 and 27 footnote 1. Brown takes his cue from Paul Ricoeur’s analysis of the fall of Adam: “The first man, in his turn, is summed up in one act: he took the fruit and ate it. About that event there is nothing to say; one can only tell it; it happens and henceforth evil has arrived” (Ricoeur 1967).

⁴⁶Brown (1978, 326).

rather than an intellectualism⁴⁷, I think that MacDonald's reconstruction of Augustine's account of the primal sin is clearly intellectualism, for reasons that will become clear in what follows. In the first of these papers, "Primal Sin," MacDonald takes as his starting point a criticism by William Babcock that the primal sin was "a random outcome, an event of pure happenstance rather than the agent's own act."⁴⁸ More specifically, MacDonald describes his project as follows:

I think that Babcock's assessment of Augustine's account is mistaken. In particular, I reject the inference from Augustine's claims that primal sin can have no cause (or only a deficient cause) to the view that primal sin must be a mere random outcome, an event of pure happenstance,⁴⁹ and so not a manifestation of genuine moral agency. My view is that Augustine's rather abrupt refusal to undertake a search for the causes of evil free choices is misleading, in effect masking his own patient and subtle pathology of sin in general and primal sin in particular. A careful look at Augustine's moral psychology of sin will, I think, provide the materials for constructing a defensible account of the radical voluntary initiatives that, on his view, introduce evil into God's good creation.⁴⁹

MacDonald's reconstruction is guided by the conviction that there are two faces' to morally evil choices. First, since the primal sinlike all choices to sin is a morally evil act of will, the faculty of the will must play a central. But, and this is the second central element of MacDonald's account, this act of will is not inexplicable as both Rogers and Babcock think: "the psychological continuity between primal sinner and primal sin is provided by the other part of the explanation of the sin: insofar as sin is an act of will it is motivated in a perfectly ordinary way by the agent's beliefs and desires."⁵⁰ And for Augustine the choice to sin, like all acts of will, aims at what is perceived to be good by the agent in some way: "If we are to make sense of a person's voluntary actions, we must understand what in or about those actions moves her to view them favorably, what it is in them that she loves or takes to be worth seeking."⁵¹ As we'll

⁴⁷See, for instance, Mendelson (2012).

⁴⁸Babcock (1988, 47), as quoted in MacDonald (1998, 112).

⁴⁹MacDonald (1998, 113). MacDonald makes it clear that his primary aim is not Augustine exegesis, but rather an Augustinian reconstruction of primal sin: "Although I find the materials for this part of my project in Augustine's writing, particularly in *De libero arbitrio* and *City of God*, I extend Augustinian ideas beyond what I can claim to have found explicitly in the texts. For this reason I am not able (and am not particularly concerned) to distinguish clearly this constructive enterprise from the more strictly reconstructive first part of this paper" (MacDonald 1998, 113). For concerns about MacDonald's interpretation of the relevant Augustinian texts, see Rogers (2008, 49f) and Mann (2006). For an account of Augustine as a voluntarist, see King (2012). King's understanding of Augustine's view is very similar to Roger's interpretation of Anselm. This should not be surprising, however, since King writes that "Anselm is clearly at pains to make his account fully compatible with Augustine, cleaning up and extending Augustine's views" (King (2012), p. 20 in manuscript.). King's reading of Augustine's account bears significant similarities to Rogers', discussed above. If King is correct that Augustine ought be interpreted as a voluntarist, then the distinction between voluntarist and intellectualist accounts of primal sin cannot be illustrated by Anselm and Augustine, respectively. However, insofar as I am not primarily concerned with the interpretive issues, I can set aside this potential worry.

⁵⁰MacDonald (1998, 113).

⁵¹MacDonald (2003, 398). See also King (2012) p. 9 in manuscript.

see more fully below, the intellect plays a crucial role in Augustine's view of sin; but this quotation also gives evidence for Augustine's acceptance of the reasons-constraint on free choice.

Augustine's account of primal sin, MacDonald argues, must be taken in the context of the larger context of theodicy of which it is a part. Here, as is well known, Augustine holds that all evil is a privation or corruption of a good created by God, rather than having ontic status of their own. "Primal sin conforms to this general account: the first evils are defective free choices that constitute a corruption in rational nature. Primal sin occurs when, by an act of free will, rational creatures irrationally turn away or defect from the highest good."⁵² MacDonald elaborates this point elsewhere as follows: "We sin, Augustine believes, not because we are motivated by bad things but because we pursue perfectly natural and appropriate delights inordinately, preferring them and the things in which they reside to higher goods."⁵³ But in order for the devil to be blameworthy for this turning away, rather than God, it must also be shown that such a choice is voluntary.⁵⁴ The difficulty, of course, is describing how an otherwise uncorrupted agent could voluntarily choose to sin without that choice being inexplicable.

On Augustine's view, in order for the primal sin (or any sin) to be something for which the agent is responsible, it must be a result of the agent's free choice. And like Anselm will do later, he denies that there is a positively existing previous cause of the devil's choice to sin. But this is not to say that the devil had no reason or motives for his choice:

Augustine holds that sin is essentially a disordered act of will, the turning away from the highest good toward a lesser good. As he sees it, then, the sinner's act of will the choosing of the lesser good is motivated by the fact that the sinner perceives the goodness of the object he comes to choose. . . . So on his view, there is a straightforward sense in which something moves a primal sinner's will: the object toward which the disordered act of will is directed; and a straightforward sense in which that act of will is intelligible: it is directed toward an object that is worth choosing. Augustine's denial that there is any cause of sin other than the will itself is clearly not meant to suggest that sins are bare, utterly unmotivated acts of will. On his own account events of that sort would be nothing more than unintelligible eruptions in the lives of sinners and not voluntary acts at all.⁵⁵

So far, MacDonald's account may seem to be little different from Rogers', for though it provides a motive or some perceived good for the choice to sin, and thus the choice is not completely inexplicable, it does not yet provide a contrastive reason or explanation for why the motive to sin prevailed over the competing motive to choose to align oneself with God. Both accounts also seem to affirm the reasons-constraint on free choice introduced in the previous chapter.

MacDonald continues: "Primal sinners' defection from God, then, cannot be explained simply by the fact that they perceive some created thing as good and

⁵²MacDonald (1998, 115).

⁵³MacDonald (2003, 400).

⁵⁴If Augustine can succeed in this, MacDonald thinks that Augustine will have succeeded in showing that God bears no responsibility in the fall since according to Augustine "the sole source of evil is in the free choice of the will" (MacDonald (1998) 136 footnote 14).

⁵⁵MacDonald (1998, 118f).

so reasonably desire it. Moreover, it is no help simply to add that they desire some created thing more than they desire God, for that irrational preference is just what needs explaining.”⁵⁶ And it is at this point we begin to see why MacDonald’s account is best viewed as a form of intellectualism, rather than voluntarism. The reason why this is that the faculty of the will which chooses to sin does so, on MacDonald’s reconstruction, only as a result of the agent failing to see’ properly:

[Augustine] suggests that the will falls when it fails to guard against sin. It follows from primal sin’s being a sin that it could have guarded against. Its actual occurrence shows that primal sinners failed to guard against it. Would could primal sinners have done to guard against sinning? I think the answer must be that they failed to pay attention to the reason they had for loving God above all things, namely, their knowledge that God is the highest good. . . . Had they attended to the reasons they possessed, they would have seen that rationality required them to love God above all things. . . . Primal sinners, then, must have made their evil choices in some sense without thinking, without deliberating sufficiently, without taking account of relevant information that was nevertheless in their possession.⁵⁷

Noting the language herepay attention to reasons’, knowledge’, what rationality requires’, thinking’, deliberating’ all of these suggests the faculty of the intellect is playing the primary role. And MacDonald seems to admit as much: “the irrationality in primal sin must consist in a kind of carelessness in practical reasoning.”⁵⁸ MacDonald also claims that these failures of practical reasoning are not themselves the result of a disordered choice by the will, for if they were, the case under consideration could not be the primal sin. Furthermore, he explicitly distances himself from a voluntarist reading of Augustine at this point. “On this account, primal sinners are not guilty of naked irrationality, of looking the greater good squarely in the face and at the same time voluntarily and with full knowledge preferring a lesser good. In their case that sort of naked irrationality would be inexplicable because it is impossible.”⁵⁹ So the failure arises not in the will’s volition to act contrary what the agent knows to be the relevant normative reasons, but instead from the agent’s intellect failing to properly grasp or weigh the normative reasons in the first place.

MacDonald is aware that the intellectualist story he’s told so far will likely

⁵⁶MacDonald (1998, 119).

⁵⁷MacDonald (1998, 120f).

⁵⁸MacDonald (1998, 121). See also the following pages where the language also further intellectualist overtones: attention’, neglect of overriding reasons’, failure to hold love to the bounds dictated by reason’, being less than fully informed by my reasons’, etc. . . . On page 127, MacDonald also attributes intellectual, but not volitional, finitude as a necessary condition of primal sinners. And even later in the article, he writes that “the ability to choose otherwise is grounded in the ability to reason otherwise” (132).

⁵⁹MacDonald (1998, 121). Brian Leftow suggests that despite the previous paragraphs, MacDonald’s account is a voluntarist insofar as practical reasoning is under the control of the will. But if that’s the case, then his account differs even less from the kind of voluntarist view discussed above, and it is hard to understand his rejection of the latter. Like myself, Stewart Goetz also reads MacDonald’s account as intellectualist; see Goetz (2009, 31–33).

not be fully persuasive to everyone. As he rightly recognizes, “Augustine sees that arguing that the trail of moral culpability stops at created rational beings requires him to maintain that the first sinners are not created defective in any morally relevant way—that is, that the moral defects constituted by their primal sin are not preceded by any other morally relevant flaw in creation.”⁶⁰ But why would a non-fallen and, thus far, morally perfect being fail with respect to her practical reasoning in the way that MacDonald suggests is required for the primal sin? MacDonald considers an objection of just this sort and replies as follows: It might be objected that we have not yet made any progress on the main task of resolving the paradox of primal sin, namely, showing that irrational free choices that appear necessarily unmotivated are nevertheless intelligible as the choices of morally responsible rational creatures. . . . One might object that he has only pushed the problem one step back, from unmotivated choices to unmotivated failures in practical reasoning. If an essential element in primal sin remains at bottom unmotivated, then primal sin itself must be ultimately inexplicable and therefore unintelligible as an instance of moral agency.

The objection rests on the assumption that, if primal sin is to be an instance of genuine moral agency, it must be explicable right down to the bottom, as it were, in terms of the primal sinner’s reasons and motives. The objection’s point is that Augustine’s account leaves something in primal sin inexplicable in those terms and for that reason leaves primal sin itself unintelligible.⁶¹

MacDonald’s response is to deny the assumption that the objection rests on. He admits that the agent’s act of will has an explanation in terms of the agent’s reasons, thereby endorsing the reasons-constraint on free choice as seen above. But what he denies is that the failure to attend properly to her reasons is something that the agent does. And since it isn’t something that he does, it need not be explicable. “Strictly speaking, then, Augustine’s account leaves no act of the primal sinner unmotivated or unintelligible.”⁶²

I’m willing to grant that omissions are not positive actions, and thus need not always be motivated and chosen by the agent for reasons. But not all omissions are equal in this regard. When a person neglects a relevant consideration that lies outside of his cognitive reach, then his ignorance counts as a moral excuse. For example, consider again gluttonous Gene, whom we encountered earlier. Gene’s favorite ice-cream is homemade pistachio, which he is inclined to eat in substantial quantities on a nightly basis. If Gene has never learned about proper nutrition, if, say, he’s never encountered the Food Pyramid, doesn’t know that his favorite treat has 343 calories per half-cup serving, has never been told of the link between sugar consumption and the risk of developing type-2 diabetes,

⁶⁰MacDonald (1998, 116).

⁶¹MacDonald (1998, 130).

⁶²MacDonald (1998, 131). Since the failing to attend to reasons is not, on MacDonald’s view, something that the agent does, Goetz denies that MacDonald’s account provides a solution to the problem of primal sin: “If the first human sinners failed to attend to reasons they had for delighting in the highest good and only attended to the reason they had to eat the fruit of the tree, then they directly formed by default the intention to take the fruit. They did not first choose to take the fruit” (Goetz 2009, 34).

etc. . . that's one thing. But if, on the other hand, Gene is aware of all of these factors and simply fails to consider the relevant reasons for changing his dietary practices by not paying attention to the relevant normative reasons or by seeing how they compare with his motivation to eat the ice-cream, we would say that Gene's done something: he's failed to consider what he knows. And this kind of failure of practical reasoning, even if not a positive action, is something for which Gene could justifiably be held responsible for if we flesh out the story in the right way. For example, Gene may have engaged in a practice of forming his moral character in such a way that normative reasons simply have no motivational purchase on him and, over time, are no longer noticed by Gene. But insofar as we're talking about primal sin, a parallel story cannot be told regarding the devil's failure of practical reasoning. Another option is that Gene's intellect is simply unable to consider these reasons; perhaps his intellect is systematically insensitive to dietary reasons or the basics of nutrition are for some other reason beyond his intellectual grasp. In this case, Gene wouldn't be blameworthy for this failure of practical reasoning with respect to the normative reasons that should guide his behavior. But this option is not available for MacDonald regarding primal sin, however. For as he notes, it is not open to Augustine to appeal to cognitive deficiencies of this sort . . . because ignorance of or cognitive error with respect to the relevant facts would undermine the primal sinner's responsibility for failing to love God in the appropriate way. Insofar as ignorance and cognitive error mitigate irrationality they also excuse it, provided that the agent is not culpable for the cognitive deficiencies themselves.⁶³

Instead, MacDonald argues that "on certain occasions our exercising or failing to exercise that ability [for practical reasoning] is simply and entirely up to us and so something for which we bear ultimate moral responsibility."⁶⁴ This way of putting it is unfortunate, as it suggests that the failure to exercise practical reasoning is a result of a previous act of the will, contradicting what he's articulated above.

So, what it appears that he means is that agents sometimes experience *de novo* failures of intellect that are themselves without reasons. But this then sounds like it is something that merely happens to the agent, rather than something that the agent controls. And it is hard to see how an agent could be morally responsible for such a failure particularly a failure with the drastic consequences that Augustine thinks primal sin has on the devil and, through him, on the rest of creation! MacDonald claims that "in primal sin we have pure morally culpable wrongdoing laid open to view"⁶⁵, and his main criticism of voluntarist accounts is that they leave the primal sin as something "utterly unintelligible,

⁶³MacDonald (1998, 119f). Similarly, King writes that "if Lucifer did not know that the action he was contemplating was morally wrong, and so ought not be done, then Lucifer would be ignorant rather than blameworthy" (King (2012) 14 in manuscript).

⁶⁴MacDonald (1998, 131).

⁶⁵MacDonald (1998, 133).

that no intelligible motivation can be found that would explain *it+.”⁶⁶ But the intellectual failure that undergirds his entire account of primal sin is no more explicable than is the act of will on the voluntarist picture that Rogers sketches and that MacDonald rejects. As Rogers notes in her own discussion of MacDonald’s account, “the initial failure *of practical reasoning+ is not a voluntary act, or indeed any sort of act at all. But then, contrary to MacDonald, the failure itself cannot be blameworthy.”⁶⁷ The failure of practical reasoning that leads to the primal sin, and thus indirectly to all other moral evil, is simply something that happens to the agent. This failure would seem to be more a design flaw than moral agency in action. Furthermore, one might wonder if understanding the primal sin as primarily a *de novo* failing of practical reasoning can do justice to the seriousness of sin.⁶⁸ Failing to consider the reasons for not engaging in a gluttonous activity is one thing; but failing to consider why one shouldn’t fall from the state of grace is another. It is hard to see how the devil’s fall which would in turn lead to the fall of humanity, original sin, murder, rape, genocide, etc. . . . can satisfactorily be explained by merely failing to consider reasons that one has but isn’t moved by.

§5 Taking Stock

Over half a century ago, C. S. Lewis wrote that “the first sin . . . must be something which a being free from the temptations of fallen man could conceivably have committed.”⁶⁹ In the previous two sections, I’ve explored at significant length what I think are the two best treatments of the primal sin that can be found in the contemporary philosophical literature. One of those treatments, Rogers’ Anselmian understanding, is decidedly voluntarist in nature. MacDonald’s Augustinian treatment, on the other hand, comes out (perhaps surprisingly) as intellectualist in orientation. While there are certainly options for other voluntarist and intellectualist accounts which perhaps differ in some of the details from these accounts, I think that we can take these two accounts as sufficiently representative of their traditions to take comparative stock. Furthermore, insofar as these two traditions are the two guiding paradigms of the interaction of the faculties involved in free agency, whether or not a satisfactory account of the primal sin can be given would seem to depend on one or the success of one or the other of these approaches. Brian Leftow has suggested, for example, that we perhaps could give a better account if we had a more robust angelic psychology, such as including the passions:

The problem with Satan’s fall . . . is that it’s a whopping big case of *akrasia*, but we’re trying to make sense of it while denying ourselves the resources we have in the human case for doing so (e.g. passions). This leads directly to its becoming inexplicable. . . . So why not draw the moral that we’ve got the wrong picture of angels, and add some

⁶⁶MacDonald (2003, 410).

⁶⁷Rogers (2008, 49).

⁶⁸This point was raised by Stephen Boulter.

⁶⁹Lewis (1940, 76).

resources needed to make sense of the story? Why think that angels don't have passions, for instance?... Desire almost paralyzes the will to do otherwise (phenomenologically)... The stronger the passion, the more likely one goes with it, *ceteris paribus*. Given a very strong passion, it can become probable that one goes with it. When what is probable happens, then *ceteris paribus* what made it probable may explain it. Do reasons really give us anything more?⁷⁰

There are two reasons why this line is not persuasive, however. First, one can raise a parallel dilemma to this role of the passions that was raised against the role of the intellect on MacDonald's account above. There it was asked if the devil was incapable of attending, with his intellect, to the relevant motivational reasons for not choosing against God or not. If he wasn't so able to, then it looks to be a design flaw on the part of the Creator rather than a moral failing on behalf of the creature. Similarly, here, if the devil was simply unable to resist a motivational affection reason for a lower good, then that would appear to absolve the devil of moral blame. On the other hand, if the devil was able to attend the relevant reasons he had and compare them to the normative reasons he's aware of, but simply didn't, the charge above was that this does not resolve the arbitrariness worry that the intellectualist raises against the voluntarist. In the case of the passions that Leftow is raising, if the devil could have resisted the disordered desire but simply didn't, this appears to be no less arbitrary, for it seems as if the only difference is if the motivational reasons at issue are intellectual or affective. The second response to Leftow's suggestion is related. He suggests that on the voluntarist account, "reasons *don't+ really give us anything more." But this is not to say that the passions give us any more of an explanation than does the voluntarist account. And if the latter involves an unsatisfactory degree of arbitrariness, then so would Leftow's suggested amendment to angelic moral psychology.

So it looks as if whether or not a satisfactory account of the primal sin can be given would seem to depend on the success of either the intellectualist or voluntarist approach. The chief virtue of MacDonald's intellectualist account is that it seems to avoid the volitional arbitrariness that Rogers' accepts (more on this below); but this is true only if we take that agent's intellect to not be under the control or guidance of the will at the time in question. Remember MacDonald's claim that "on certain occasions our exercising or failing to exercise that ability [for practical reasoning] is simply and entirely up to us and so something for which we bear ultimate moral responsibility."⁷¹ If the exercise (or lack thereof) of practical reason is itself the result of the will, then MacDonald's account fails to differ in this central regard from the kind of voluntarist explanation that he's seeking to avoid. But if, on the other hand, this failure of practical reason is *de novo*, it's not clear that we have any more satisfactory of an answer. For either the intellect could have attended to the reasons for not sinning but didn't, or it could not. The latter option, of course, is problematic for traditional Christian

⁷⁰Leftow relayed this objection in personal correspondence.

⁷¹MacDonald (1998, 131).

views of human nature, insofar as the Creator, rather than the creature, would then be responsible for this inability. The ultimate explanation for primal sin (and for subsequent sins, it would seem) thus would fall to God's creative act rather than the misuse of free will; as such, the attempt to safeguard the goodness of God that lies at the heart of much Christian philosophical and theological reflection regarding evil fails. If, on the other hand, the created intellect could have attended to the reasons it did possess but simply did not, then there is—despite MacDonald's efforts—a brute inexplicability at the heart of his intellectualist account.

Unlike MacDonald, Rogers unapologetically accepts seeing the primal sin as something both “inexplicable and ultimately mysterious.”⁷² Like Brown, she thinks that any further explanation for why the devil would make this choice is to seek an explanation where none can be given. There are, of course, parallels between Rogers' account and other debates in agency theory. As mentioned above, many object to libertarian views of free will in general because of their inability to provide contrastive reasons for the agent's choices. And while a similar kind of inexplicability seems to be at work in cases of weakness of will in general insofar as the agent chooses what she (rightfully) knows is good, but a lesser good, the brute arbitrariness of this choice is significantly starker. This is true both because one cannot appeal to any previous moral corruption in explaining that choice and because the monumental consequences that follow from this choice in Christian theology. Furthermore, this inexplicability is not caused by the rejection of a teleological account of agency or the reasons-constraint on free choice, which both Rogers and Anselm would accept (as would MacDonald and Augustine). It looks then as if a Christian account of primal sin cannot avoid all arbitrariness. And many, even those who are inclined to libertarian accounts of agency, will likely find something unsatisfactory about this arbitrariness. Whether or not this amounts to an insurmountable objection to the philosophical respectability of Christian accounts of free will and sin will depend, among other things, on the positive merits that those accounts can offer. But with respect to the explicability of the primal sin, intellectualist accounts do not offer the advantage over voluntarist accounts that they are sometimes claimed to tender.⁷³

REFERENCES

- Aquinas, Thomas. 1948. “Summa Theologica.” In *Complete English Edition in Five Volumes*, edited by Fathers of the English Dominican Province. Vol. 1. Westminster, MD: Christian Classics.

⁷²Rogers (2008, p. 87).

⁷³Previous versions of this paper were presented at the Joseph Butler Society at Oriel College (Oxford University), Oxford Brookes University, and at the Seventh Annual Philosophy of Religion conference at Baylor University; on these occasions, I benefitted from valuable input and feedback from the audience. In particular, I would like to thank Jon Kvanvig, Jeroen De Ridder, Brian Leftow, Tim Mawson, Bryan Reece, Ross Parker, Stephen Boulter, Jonny Brown, Dan ONiel, Shane Glackin, and Constantine Sandis for their comments on previous versions of this paper.

- Augustine. 1993. *On Free Choice of the Will*. Indianapolis, IN: Hackett Publishing Company, Inc.
- Babcock, William S. 1988. "Augustine on Sin and Moral Agency." *The Journal of Religious Ethics* 16: 28 – 55.
- Boyd, Gregory A. 2001. *Satan and the Problem of Evil*. Downers Grove, IL: InterVarsity Press.
- Brown, Robert F. 1978. "The First Evil Will Must Be Incomprehensible: A Critique of Augustine." *Journal of the American Academy of Religion* 46: 315–29.
- Clarke, Randolph. 1996. "Contrastive Rational Explanation of Free Choice." *Philosophical Quarterly* (183): 185–201.
- Evans, Gillian R. 1999. "Evil." pp. . 340–344. Grand Rapids, MI: Eerdmans.
- Fitzgerald, Allan D. 1999. *Augustine Through the Ages: An Encyclopedia*. Grand Rapids, MI: Eerdmans Publishing Company.
- Ginet, Carl. 1989. "Reasons Explanation of Action: An Incompatibilist Account." *Philosophical Perspectives* 3: 17–46.
- Goetz, Stewart. 2009. *Freedom, Teleology, and Evil*. London: Continuum.
- Hitchcock, Christopher. 1999. "Contrastive Explanations and the Demons of Determinism." *British Journal for the Philosophy of Science* 50: 585–612.
- Kane, Robert. 1996. *The Significance of Free Will*. New York: Oxford University Press.
- King, Peter. 2012. "Angelic Sin in Augustine and Anselm." In *A Companion to Angels and Medieval Philosophy*, edited by Tobias Hoffman. Leiden: Brill.
- Leftow, Brian. 1998. "Voluntarism." *Routledge Encyclopedia of Philosophy* (ed. E.Craig). [Http://www.rep.routledg.com/article/K110](http://www.rep.routledg.com/article/K110).
- Lewis, Clive Staples. 1940. *The Problem of Pain*. New York: HarperCollins.
- MacDonald, Scott. 1998. "Primal Sin." pp. . 110–39. Berkeley: The University of California Press.
- MacDonald, Scott. 2003. "Petit Larceny, the Beginning of All Sin: Augustine's Theft of the Pears." *Faith and Philosophy* 20 (4): 393–415.
- Mann, William. 2006. "Pride and Preference: A Reply to MacDonald." *Faith and Philosophy* 23: 156–68.
- McClusky, Colleen. 2007. "Medieval Theories of Free Will." *Internet Encyclopedia of Philosophy*. [Http://www.iep.utm.edu/freewi-m/](http://www.iep.utm.edu/freewi-m/).
- Mendelson, Michael. 2012. "Saint Augustine." The Stanford Encyclopedia of Philosophy. Forthcoming URL, <http://plato.stanford.edu/archives/win2012/entries/augustine/>.
- Milton, John. 1931. *The Works of John Milton*. New York: Columbia University Press.
- O'Connor, Timothy. 2000. *Persons and Causes: The Metaphysics of Free Will*. Oxford: Oxford University Press.

- O'Connor, Timothy. 2005. "Freedom with a Human Face." *Midwest Studies in Philosophy* 29: 207–27.
- Plantinga, Alvin. 1974. *The Nature of Necessity*. Oxford: Clarendon Press.
- Plantinga, Alvin. 1977. *God, Freedom, and Evil*. Grand Rapids, MI: William B. Eerdmans Publishing Co.
- Quinn, Philip. 1998. "Sin." Routledge Encyclopedia of Philosophy (ed. E. Craig). Retrieved July 12, 2010 from <http://www.rep.routledge.com/article/K095>.
- Rea, Michael C. 2007. "The Metaphysics of Original Sin." pp. . 319–56. Oxford: Oxford University Press.
- Ricoeur, Paul. 1967. *The Symbolism of Evil*. New York: Harper & Row.
- Rogers, Katherin A.. 1997. *The Anselmian Approach to God and Creation*. Lewiston: Edwin Mellen Press.
- Rogers, Katherin A. 2004. "Augustine's Compatibilism." *Religious Studies* 40 (4): 415–35.
- Rogers, Katherine. 2008. *Anselm on Freedom*. Oxford: Oxford University Press.
- Seymour, Charles. 1999. "On Choosing Hell." *Religious Studies* 33: 249–266.
- Sommers, Tamler. 2010. "Experimental Philosophy and Free Will." *Philosophical Compass* 5 (2): 199–212.
- Steadman, John M. 1960. "'Man's First Disobedience': The Causal Structure of the Fall." *Journal for the History of Ideas* 21.2: 180–197.
- Stump, Eleanore. 2003. *Aquinas*. London: Routledge.
- Stump, Eleonore. 1990. "Intellect, Will, and the Principle of Alternative Possibilities." pp. . 254–285. South Bend: University of Notre Dame Press.
- Swinburne, Richard. 1989. *Responsibility and Atonement*. Oxford: Clarendon Press.
- Visser, Sandra and Thomas Williams. 2008. *Anselm*. Oxford: Oxford University Press.
- Wyma, Keith D. 2004. "Innocent Sinfulness, Guilty Sin: Original Sin and Divine Justice." pp. . 263–76. Grand Rapids: Wm. B. Eerdmans Publishing Co.