Arguments from Moral Evil

Many philosophers seem to suppose that the argument of Plantinga (1974)[1]—or a suitably elaborated variant thereof—utterly demolishes the kinds of "logical" arguments from evil developed in Mackie (1955)[2].[3] I am not at all convinced that this is a correct assessment of the current state of play. First, I think that Plantinga's free will defence involves a hitherto undetected inconsistency. Second, I think that even if Plantinga's free will defence is consistent, it relies upon some indefensible metaphysical assumptions. Third, I think that even if the metaphysical assumptions upon which Plantinga's free will defence relies are defensible, there are serious questions to be raised about the moral assumptions which are made in that defence. Finally, I think that, even if Plantinga's free will defence is acceptable, there are arguments closely related to those developed in Mackie (1955) that are not vulnerable to any variant of Plantinga's free will defence, and yet that are clearly deserving of further examination.

The structure of the paper is as follows. In section 1, I present a standard "logical" argument from moral evil, and give Plantinga's reply to it. In section 2, I provide my argument in support of the claim that there is an inconsistency in Plantinga's free will defence. In section 3, I assess those parts of Plantinga (1974) that might be taken to bear on my claim that there is an inconsistency in Plantinga's free will defence. In sections 4 and 5, I identify some of the controversial metaphysical assumptions that are required for Plantinga's reply, and I suggest that at least one of these assumptions is really not acceptable. In section 6, I consider some assumptions about values that are also required for Plantinga's reply, and argue that here Plantinga is probably on

safer ground. In section 7, I consider an alternative formulation of the free will defence that avoids both the inconsistency and the unacceptable metaphysical assumptions, but that is subject to the other kinds of worries that can be raised in connection with Plantinga's reply. In section 8, I turn to consider some probabilistic arguments from moral evil that are natural developments from the standard "logical" argument from moral evil. In the final section of the paper, I consider replies that might be made to these probabilistic arguments.

Throughout, the aim of my discussion is to show that the assumption that there is nothing further to be said on behalf of arguments from moral evil—and, in particular, on behalf of the kind of argument which is developed in Mackie (1955)—is premature. I don't claim to be able to show that there are successful arguments from moral evil; however, I do think that philosophers ought not to be too readily inclined to dismiss these arguments out of hand.[4] In particular—as I shall also go on to argue—I think that it is plausible to claim that arguments from moral evil generate serious constraints on positive arguments that can be mounted for the existence of a perfect being.

(1)

The standard "logical argument" from moral evil—really, an *assertion* rather than an argument—claims that it is logically impossible for the following set of claims to be jointly true:

1. A perfect being exists[5].

2. Any perfect being is omnipotent.
3. Any perfect being is omniscient.
4. Any perfect being is perfectly good.
5. If there is a perfect being, then it is the sole creator of the universe.
6. Moral evil exists[6].

Plantinga's response to this "argument" is, in effect, to describe a logically possible world in which 1-6 are all true.[7] In what follows, I shall try to give a reasonably faithful reconstruction of the kind of possible world that Plantinga envisages (and of the kind of conception of logical space that one must have if one is to suppose that what one has described is, indeed, a logically possible world).

Focus attention on possible worlds in which there are perfect beings. (Perhaps it is not logically possible for there to be any such worlds. However, if we have good reason to believe this, then we don't need to proceed to an examination of the logical argument that is our current target. Perhaps, too, every possible world is one in which there is a perfect being; if so, then our attention is merely directed to all possible worlds.[8])

Some—i.e. at least one—of these worlds will contain nothing but the perfect being (and whatever else, if anything, is necessitated by the existence of the perfect being).[9] Others will contain the perfect being and further things that exist as a result of the free creative activities of the perfect being. Amongst these further possible worlds, there may be some that do not contain any free agents. However, our interest lies in those possible worlds in which there is a perfect being who has created a universe

containing free agents. (As a matter of definition, a "universe" will be that part of the possible world that is left over when the perfect being (and whatever else is necessitated by the existence of the perfect being) is "subtracted" away. Within the perhaps restricted class of possible worlds that we are now considering, universes are the products of the free creative activities of perfect beings.[10])

Suppose that freedom is libertarian, i.e. suppose that if an agent X acts freely in performing action A in circumstances C at time T in world W, then it is not made true by the truth-making core of the world W prior to T that agent X will do A in circumstances C. [11] Compatibilists reject this conception of freedom; they hold that an agent X can act freely in performing action A in circumstances C at time T in a world W in which the truth-making core of the world prior to T is such that, for any world W' with exactly the same truth-making core prior to T, the agent X does action A in circumstances C at time T in W'. That is, compatibilists—unlike libertarians—do not require that a free agent is "able to do otherwise in the circumstances of her actions". Note that it is a consequence of this account that, if there are truths about what agents with libertarian freedom will do in a world W at time T, then those truths do not belong to the truth-making core of the world W prior to the time T. Note, too, that it is a consequence of this account that, if a truth does not belong to the truth-making core of a world W at a time T, then that truth does not constrain the actions of agents with libertarian freedom at time T in world W: if nothing prior to T has made it true that the agent does not do A at T in W, then—provided that there is some possible world W' at which the agent does do A at T[12]—the agent is *able* to do A at T in W.[13]

Suppose that when a perfect being creates a universe containing free agents, we can distinguish two parts of the universe: that part S for which the perfect being has sole responsibility, and that part J whose nature is in part determined by the free choices of the created free agents. Suppose further that, "when" a perfect being "deliberates" about whether or not to create a universe containing free creatures, and—in the circumstances in which it does decide to create a universe containing free creatures—about which universe containing free creatures to create, it begins with a survey of all of the possible universe parts S' for which it has sole responsibility that it could make.

Thus far, there is no difference between the activities of the perfect being in one possible world, and the activities of the perfect being in any other possible world: the possible universe parts for which it has sole responsibility do not vary from one possible world to the next. However, suppose further that, in any given possible world, there are true counterfactuals about the parts of universes whose nature is in part determined by the free choices of free created agents that *would* ensue were the perfect being to create any given universe part that is entirely up to it. That is, suppose that something like the following is true:

In world w_1, if the perfect being were to make S_1, then J_{11} would ensue; if the perfect being were to make S_2, then J_{12} would ensue; if the perfect being were to make S_3, then J_{13} would ensue; and so on.

In world w_2, if the perfect being were to make S_1, then J_{21} would ensue; if the perfect being were to make S_2, then J_{22} would ensue; if the perfect being were to make S_3, then J_{23} would ensue; and so on.

In world w_3, if the perfect being were to make S_1, then J_{31} would ensue; if the perfect being were to make S_2, then J_{32} would ensue; if the perfect being were to make S_3, then J_{33} would ensue; and so on.

And so on, through all of the possible worlds in question. Note that there is no assumption that, say, $J_{11} \neq J_{21}$; rather, what is assumed is that, for any pair of worlds w_i and w_k, there is at least one of the S_j for which $J_{ij} \neq J_{kj}$.

So, in each possible world, there is some—perhaps not necessarily proper—subset of the possible Js that is available to the perfect being as a result of its creative activities. Some of the S+Js are universes in which there are free creatures who always freely choose the good. Some of the S+Js are universes in which there are free creatures who always freely choose the bad. Many of the S+Js are universes in which there are free creatures who sometimes freely choose the good and sometimes freely choose the bad.[14] Moreover—and this is the crucial point—there are some possible worlds in which the S+Js which are available to the perfect being as a result of its creative activities do not include any universes in which there are free creatures who always freely choose the good. As an extreme example, there is at least one possible world in which, no matter which of the S_i the perfect being were to choose, the resulting S+J *would* contain free creatures who *all always* freely choose the bad.

Given this picture of the creative activities of a perfect being—and of the logical space in which that creative activity is embedded—it seems at least *prima facie* plausible to claim that there are possible worlds in which it is true that, if the perfect

being creates any universe which contains free agents, then it creates a universe in which it is not true that all free agents always freely choose the good.

Of course—even given the assumptions which we have already made—this is not enough to show that it is *prima facie* plausible to claim that there are possible worlds in which a perfect being makes a universe in which it is not the case that everyone always freely chooses the good. For plainly one might think that, in circumstances in which a perfect being could not make a universe in which everyone always freely chooses the good, the perfect goodness of that being ensures that it will make no universe at all. However, setting this consideration aside—or, what amounts to the same thing, adding the extra assumption that, in at least some circumstances in which it is not possible for a perfect being to make a universe in which everyone always freely chooses the good, it is possible for a perfect being to make universes in which it is not the case that everyone always freely chooses the good—it seems *prima facie* plausible to claim that we do indeed get to the conclusion that there are possible worlds in which 1.-6. are all true.

(2)

I think that, despite appearances, there is an inconsistency in Plantinga's "possibility proof"; i.e., I think that he has not succeeded in describing a logically possible world. Consider the counterfactuals of freedom that Plantinga supposes constrain the creative activities of a perfect being "when" it is making universes, and yet that are not inconsistent with the libertarian freedom of creatures in those universes. Either these

counterfactuals of freedom are “then” part of the truth-making core for the world, or they are not.

On the one hand, if they are “then” part of the truth-making core for the world, then it follows from the libertarian account of freedom that no-one ever acts with libertarian freedom, for there is no other world with the “then” same truth-making core in which agents do anything other than what they do in the world in question. Allowing that the counterfactuals of freedom are part of the truth-making core of the world “when” the perfect being makes its creative decisions entails that there is no libertarian freedom in the world. So, on the assumption that counterfactuals of freedom are part of the truth-making core “when” the perfect being makes its creative decisions, Plantinga’s free will defence does not go through.

On the other hand, if the counterfactuals of freedom are not “then” part of the truth-making core for the world, then it follows from the earlier noted consequences of the libertarian account of freedom that these counterfactuals *cannot* constrain the choices which a perfect being can make, and nor *can* they constrain the actions which that being can perform. Why not? Because the truth of the counterfactuals of freedom is not *fixed* “when” the perfect being deliberates, but rather somehow depends upon “subsequent” features of the universe in question. “Prior” to the creative decision of the perfect being, there just isn’t anything which makes the counterfactuals of freedom true; hence, which counterfactuals of freedom are “then” true *depends upon* the creative choice which is made by the perfect being. Allowing that the counterfactuals of freedom are not part of the truth-making core “when” the perfect being makes its creative decisions entails that those counterfactuals of freedom cannot

constrain the choices that the perfect being makes. So, again, on the assumption that counterfactuals of freedom are not part of the truth-making core "when" the perfect being makes its creative decisions, Plantinga's free will defence does not go through.

So, no matter how one thinks about the relationship between Plantinga's counterfactuals of freedom and the truth-making core of the world "when" a perfect being deliberates about which universe to make, Plantinga's free will defence does not go through: it *cannot* be that the truth of those counterfactuals *both* constrains the choice that the perfect being makes and yet also allows that there are creatures in the chosen universe that have libertarian freedom.

Before I turn to consider why Plantinga misses this objection, it might be worth explaining the argument on each side of the above dilemma in a little more detail. I shall begin with the horn of the dilemma that assumes that counterfactuals of freedom impose a genuine constraint on the creative activities of a perfect being.

(1) *Counterfactuals of freedom in the truth-making core*:

Recall, again, that one key idea behind Plantinga's free-will defence is that the truth of certain counterfactuals of freedom *constrains* the universe-making activities of a perfect being: given that it is true at w, prior to the creation of any universe, that, if the perfect being were to make S_1, then universe S_1+J_{11} would ensue, then, at w, the perfect being *cannot* bring about any of the worlds S_1+J_{1k}, for $k \neq 1$[15]. So, "when" the perfect being comes to make universes, the counterfactuals of freedom impose

genuine constraints (from which it follows that those counterfactuals of freedom must "then" be part of the truth-making core for the world).

But now, suppose that the perfect being makes S_1, and consider an allegedly free being X's allegedly free choice of an action A within S_1+J_{11}. In order for X's doing A at w to be free—on the libertarian account of freedom—there must be a world w' whose truth-making core is identical to the truth-making core of w up until the time of X's acting, but in which X does something other than A. However, given the membership in the truth-making core of the counterfactual of freedom that constrained the universe-making activities of the perfect being, there is no such possible world: a world in which X does something other than A cannot be a world in which it was part of the truth-making core "when" the perfect being engaged in its creative activities that, were the perfect being to make S_1, then S_1+J_{11} would ensue. If "counterfactuals of freedom" can constrain the universe-making activities of perfect beings, then they cannot fail to constrain the actions of allegedly free agents.

Of course, it is no reply to this argument simply to insist that 'counterfactuals of freedom' are counterfactuals of *freedom*, and hence by definition consistent with the freedom of the agents in question. We have an official account of freedom—the libertarian account mentioned earlier—that holds that nothing in the circumstances of a free choice can *fix* the outcome of that choice. If there is no logically possible world with the same truth-making core in which there is a different outcome for the choice, then the choice is determined by the truth-making core, and hence is not free. In other words, there is no logically possible world in which (1) it is part of the truth-making core prior to T that, were conditions C to obtain, agent A would make choice X at T;

(2) conditions C obtain; and (3) agent A does not make choice X at T. Given that the counterfactual is part of the truth-making core prior to the choice, the agent simply does not have libertarian freedom.

But now, suppose that it is part of the truth-making core prior to the creation of any universe that, were a perfect being to make S_1, then agent A would make choice X at time T in circumstances C. (By hypothesis, this is one consequence of the more general claim that it is part of the truth-making core prior to the creation of any universe that, were a perfect being to make S, then universe S_1+J_{11} would result, given that agent A makes choice X at time T in circumstances C in S_1+J_{11}.) Then, there is no logically possible world in which (1) it is part of the truth-making core prior to the creation of any universe that, were a perfect being to make S_1, agent A would make choice X at time T in circumstance C; (2) a perfect being makes S_1; and (3) agent A makes some choice other than X at time T in circumstance C. So, given that the counterfactual is part of the truth-making core, the agent simply does not have libertarian freedom with respect to this choice.

(2) *No counterfactuals of freedom in the truth-making core*:

If we suppose that, "when" the perfect being is deliberating about which universe to make, there are true counterfactuals of freedom that are not part of the truth-making core, then we are supposing that the truth-values of these counterfactuals depend upon free decisions that are made "after" the creative deliberations of the perfect being have "commenced". Moreover—and consequently—we are also supposing that the truth-values of these counterfactuals depend upon the creative decision that the

perfect being makes: which counterfactuals of freedom are true *depends upon* which of the S_i the perfect being chooses to make.

Consider a relevant possible world w_i. In this world, it is true that if the perfect being were to make S_1, then J_{i1} would result; and it is true that if the perfect being were to make S_2, then J_{i2} would result; and it is true that if the perfect being were to make S_i, then J_{i3} would result; and so on. But it is also true in w_i that the perfect being *does* make S_j, say. Moreover, it is also true in w_i that, at the time when the perfect being chose to make S_j, it *had it within its power* to make some other universe S_k, say. And, crucially, it is also true in w_i that, *had* the perfect being exercised its power to make S_k, then—for at least *some* values of k—a different set of counterfactuals of freedom *would* have obtained in w_i "when" the perfect being made its decision to create. (This is what follows from the assumption that the counterfactuals of freedom are not part of the truth-making core "when" the creative decision is made.)

But, if it is true—as we are supposing—that the perfect being has it within its power to choose which universe to make, and if it is also true that which counterfactuals of freedom are true depends upon which universe the perfect being chooses to make, then it cannot be the case that the choice which the perfect being makes is *constrained* by the "prior" truth of those counterfactuals of freedom. If the perfect being's choice to create universe S_j makes it true—or, at any rate, plays an important role in making it true—that every creaturely essence at the world suffers from transworld depravity, and yet it would not have been true that every creaturely essence suffers from transworld depravity had the perfect being created world S_k instead, then we simply do not end up with a demonstration that there is a possible world in which the perfect

being is *unable* to create a universe in which everyone always freely chooses the good.

Clearly, there is a question to ask about the seriousness of this gap in Plantinga's "possibility proof" (on the assumption that counterfactuals of freedom do not belong to the truth-making core). In particular, one might suspect that there must be some way of reinstating Plantinga's argument using nested counterfactuals: if the perfect being were to make S_1, then such-and-such counterfactuals of freedom would be true; if the perfect being were to make S_2, then such-and-such counterfactuals of freedom would be true; if the perfect being were to make S_3, then such-and-such counterfactuals of freedom would be true; and so forth. I do not think that this can be right. For exactly the same question about the truth-making core can be raised for these nested counterfactuals which was raised in connection with the initial counterfactuals of freedom, and exactly the same difficulties will be seen to arise for the two ways in which that question might be answered. In particular, if these nested counterfactuals do not belong to the truth-making core then they, too, take truth-values that depend upon the creative choice that the perfect being makes (and hence we go down exactly the same argumentative path).

So—unless there is some way of repairing Plantinga's construction that I have overlooked—it seems pretty safe to conclude that Plantinga has *not* managed to describe a possible world in which a perfect being is unable to choose to make a universe in which everyone always freely chooses the good. It may be, for all that I have argued, that there is such a possible world; the point is just that no argument has been offered that can reasonably be said to have settled the case.

(3)

As far as I know, the objection to Plantinga's argument that I have been developing here has not been raised previously. There is a much-discussed related objection that I will consider in the next section. But first, I want to say something about why Plantinga misses this objection, and to comment on passages in Plantinga (1974) which might be thought to be relevant.

The most important point to note, I think, is that Plantinga does not explicitly address the question of the analysis of libertarian freedom at the level of detail that is required in order to answer the question whether the truth of a counterfactual of freedom "determines" the subsequent behaviour of agents whose actions are detailed in the consequent of that counterfactual. All that Plantinga says about an agent with (libertarian) freedom is that "no … antecedent conditions determine either that he will perform the action, or that he will not. It is within his power, at the time in question, to perform the action, and within his power to refrain."[16] But this does nothing at all towards providing a clear possible worlds analysis of the notions of *freedom*, *determination*, and *the possession of powers* (and this despite the fact that the bulk of the book is concerned with possible worlds analyses of modal notions). In my view, it is this shortcoming in the discussion that leads Plantinga to overlook the difficulty upon which I have focussed.

Perhaps it is worth noting that Plantinga does give reasons for denying that the proponent of libertarian freedom is committed to the claim that, if agent A acts freely in performing action X in circumstances C at time T in world W, then there is a world W' which is identical to world W up until the time T at which A chooses to act, but in which A does something other than X in circumstances C at time T. According to Plantinga, it is true in W before T that A will do X, and it is true in W' before T that A will not do X—and so the worlds are not identical before T as claimed. But, of course, while this point is fine as far as it goes, the crucial point to note is that the characterisation of libertarian freedom that Plantinga considers pays no attention to the question whether the worlds W and W' are identical in all respects, or whether they are only identical with respect to their truth-making cores. As we noted earlier, the principle to which the libertarian is committed is that if agent A acts freely in performing action X in circumstances C at time T in world W, then there is a world W' with the same truth-making core as world W up until the time T at which A chooses to act, but in which A does something other than X in circumstances C at time T. And, of course, it is obvious that truths about what the agent *will* do cannot be part of the truth-making core if the agent is to have libertarian freedom.

(4)

Of course, even if there is some problem with the argument presented in Section 2 of this paper, there are *further* questions that can be asked about the assumptions that are required for Plantinga's "possibility proof". As I have already indicated, there is a much-discussed objection to Plantinga's argument that is not entirely unrelated to the argument given in the previous section of this paper. This objection focuses on the

assumption that there *are* true counterfactuals of freedom, i.e. on the assumption that, in any given possible world, there are true counterfactuals about the parts of universes whose nature is in part determined by the free choices of free created agents that *would* ensue were the perfect being to create any given universe part that is entirely up to it. I shall first give the objection in my own terms, and note some consequences that seem to follow. And then I will (briefly) comment on some other discussions of this objection.

The core of the objection is the observation that the assumption that there are true counterfactuals of freedom seems to be in conflict with the plausible metaphysical claim that counterfactual claims that are true at a given possible world require a categorical grounding in that world. If a counterfactual claim is true at a possible world, then there must be something in that possible world that serves as a truthmaker for the counterfactual claim. There are no possible worlds in which counterfactual claims are *bare* truths; there are no pairs of possible worlds with minimal truth-making cores—or minimal supervenience bases—that differ *only* with respect to the truth-values of counterfactual claims.[17]

These requirements are plainly violated by the construction described in the first section of this paper. According to that construction, there are different possible worlds w_1 (= perfect being + S_1 + J_1 + counterfactuals of freedom C_1) and w_2 (= perfect being + S_1 + J_1 + counterfactuals of freedom C_2) with minimal supervenience bases that differ only in the counterfactual claims that were true in that world when the perfect being was deliberating about which universe to make. (In each of these worlds, the perfect being chooses the universe S_1 + J_1 – because, say, that is the best

option that is open to it – but the range of options from which it has to choose differs between the worlds.)

I take it that, even in the absence of the previous argument—and even if one fails to pay any attention to the distinction between counterfactuals which do, and counterfactuals which do not, belong to the truth-making core "when" the perfect being engages in its creative activities—this observation would cast very serious doubt on Plantinga's claim to have described a possible world in which 1–6 are all true. The principle that there are no pairs of possible worlds with minimal supervenience bases that differ *only* with respect to the truth-values of counterfactual claims is, I think, a pretty secure piece of metaphysical doctrine. At the very least, it is worth noting that it hasn't been plucked from the air merely to serve the interests of the current argument. There is a long history—going back at least to the criticisms that Armstrong[18] and Martin make of Ryle's dispositional theory of mind[19]—of reliance upon this principle that is completely independent of the use that might be made of it in the context of discussion of logical arguments from evil. Many people have thought that Plantinga's counterfactuals of freedom are pretty suspicious entities; if I'm right, there is a strongly principled basis to this suspicion.

Of course, Plantinga does have a reply to those who disagree with his claim that there are true counterfactuals of freedom. He gives the example of offering a small bribe to someone in order to get a good recommendation. When the bribe fails miserably, he wonders what would have happened had the bribe been larger. Plantinga claims that there must be some definite, non-probabilistic answer to the question of whether a larger bribe would have succeeded. However, if I am right, then it is only the

determinist who is entitled to this claim: the libertarian about freedom has to allow that the size of the bribe cannot determine the response of the agent, and that the agent has "the power to do otherwise" in the very circumstances which obtained when the bribe was offered.[20] (Of course, there are right answers to the question about what I would have done—e.g. that my mendacity ensures that the chance that I will take the bribe is >99%. But there is no right answer of the kind that Plantinga supposes there to be, if I have libertarian freedom.)

Consider a different case. Suppose that there were a tiny piece of uranium—a single atom—in the pencil sharpener on my desk as I write. Suppose we ask: would that uranium atom decay within the next thirty seconds? On the assumption that radioactive decay is a genuinely chancy process, I take it that intuition supports the view that there is just no answer to this question. But, as we noted earlier, if we have libertarian freedom, then our choices are chancy: for any decision that we make, there is a possible world that is identical to the actual world up until the instant of decision—i.e. involving exactly the same weighing of reasons, exactly the same preferences, etc.—and yet in which a different choice is made. So, we should say exactly the same thing about counterfactuals concerning the choices of agents with libertarian freedom: there is just no right answer to the question of what they would do were they asked to make certain kinds of choices in given circumstances—even though there are right answers to the question of what they would *very likely* do in those circumstances.

Although it is a bit of a digression from the main line of argument, it is perhaps worth pointing out that there is one class of theists who should find the above line of

argument particularly disturbing, namely, those theists who are firmly committed to arguments that rely upon a strong version of the principle of sufficient reason. According to strong versions of this principle, any contingent feature of any possible world has a *complete* explanation in that possible world: that is, for any pair of possible worlds w and w', and any contrasting features S and S' of those worlds, there is an explanation in w of why S rather than S', and there is an explanation in w' of why S' rather than S. Given that counterfactuals of freedom are contingent features of worlds, it follows immediately that there cannot be *bare* true counterfactuals of freedom: the theoretical machinery to which Plantinga is committed requires the falsity of strong versions of the principle of sufficient reason.

There is a similar point that can be made about free will defences more generally. The above criticism turns on the details of Plantinga's chosen method of developing a free will defence. But it is a non-negotiable part of any free will defence that it appeals to a libertarian conception of freedom. According to the libertarian conception of freedom, however, when agents act freely, there can be no complete explanation for their acting in the way that they do. Suppose we ask: why did agent A freely choose to do action S rather than action S' in circumstances C (in which agent A was able to do S and S')? According to the libertarian conception of freedom, there are possible worlds W and W' whose truth-making cores are identical in every respect up until the moment of the agent's choice in circumstances C, but that differ with respect to the chosen actions S and S'. Consequently, there is nothing in either world that can serve as a complete explanation of the choice that is made in that world.[21]

If this line of argument is well-taken, then it suggests that theists who advocate cosmological arguments for the existence of a perfect being cannot consistently appeal to any kind of free will defence in order to reply to logical arguments from evil (except in those cases in which the cosmological arguments make no appeal to strong principles of sufficient reason). Moreover, this same line of argument also suggests that non-theists who respond to cosmological arguments by rejecting strong forms of the principle of sufficient reason cannot consistently object to the claim that there are bare true counterfactuals of freedom on the grounds that this claim violates strong versions of the principle of sufficient reason. Of course, there may well be other reasons for rejecting the claim that there are bare true counterfactuals of freedom—and, indeed, it seems to me that there are such reasons. However, the key point that I wish to make here is just that there is a lot of heavy duty metaphysical machinery that is built into Plantinga's free will defence, and that the use of this machinery has consequences for what one can consistently say and do in other contexts.

Since the digression of the past three paragraphs was fairly lengthy, it may be worth reminding readers of where the main line of argument now stands. I have argued (1) that Plantinga's free will defence fails because it allows no coherent answer to the question whether counterfactuals of freedom belong to the truth-making core of a world "when" a perfect being is engaged in universe creation; and (2) that Plantinga's free will defence is in serious trouble because it relies on the assumption that there are true "ungrounded" counterfactuals of freedom. However, I do not think that the troubles for Plantinga's free will defence end here.

(5)

So far, we have only considered difficulties which follow from the assumption that there *are* counterfactuals of freedom of the kind that Plantinga supposes that there are. But there are more difficulties that follow if we make a definite decision about whether these counterfactuals belong to the truth-making core. In particular, let us suppose that they do not, i.e. let us suppose that the counterfactual in question is not a fixed feature of the circumstances of the agent's choice. [22] Then we must be supposing that there is something outside the truth-making core upon which the truth of the counterfactual depends. But what could this be? Could it be, for example, that the truth of the counterfactual depends upon the choice that the agent makes? No, that can't be right. "When" the perfect being is choosing which universe to make, very many counterfactuals of freedom are supposed to be true, including very many which advert to merely possible universes (and, in many cases, merely possible agents). But it makes no sense at all to say that the truth of these counterfactuals depends upon the choices that the agents make—because, *ex hypothesi*, there *are* no such choices. But what else is there for the truth of these counterfactuals to depend upon?

Perhaps it might be said that the truth of these counterfactuals does depend upon the choices that the agents make, but that these choices take places in possible worlds other than the actual world. But, at least *prima facie*, that can't be right either. The true counterfactuals of freedom are truths about free choices made by agents with libertarian freedom. So we know that, for any given choice, there are possible worlds in which the agent takes each of the options available. Appealing *merely* to what

happens in other possible worlds cannot hope to deliver truthmakers for counterfactuals of freedom.

Perhaps it might be said not only that it is that case that there are other possible worlds, but also that there are fixed relationships of similarity between the worlds that play a crucial role in making counterfactuals of freedom true. What makes a given true counterfactual of freedom true in a given world is what happens in all of the *sufficiently close* worlds to the given world. But now we need to ask: are there fixed relations of similarity between worlds at times, or are there only fixed relations of similarity between worlds? And we also need to ask: are relationships of similarity between worlds primitive, or do they depend upon the intrinsic properties of worlds? If we can suppose that relationships between worlds are primitive—and hence independent of the intrinsic properties of worlds—and if we can suppose that there are only fixed relations of similarity between worlds, then perhaps we *can* claim that we have now found truthmakers for counterfactuals of freedom. But, if we accept—as we surely should!—that relationships of similarity between worlds *depend only upon* intrinsic properties of worlds, then surely we are in no position to claim that we have found truthmakers for counterfactuals of freedom. Truth supervenes upon being. Which counterfactuals are true at a world depends just upon the intrinsic properties of that world. But these intrinsic properties are simply insufficient to make true all of the counterfactuals of freedom that Plantinga supposes to be true "when" the perfect being deliberates about which world to make.[23]

If what I have said here is right, then there is good reason to be suspicious of arguments which claim that analogies between tensed claims and counterfactuals

support the suggestion that there are true counterfactuals of freedom of the kind which Plantinga supposes that there are.[24] There are many good reasons—both physical and metaphysical—to support the claim that there *are* past and future times; hence, there are good reasons for claiming that there *are* truth-makers in the truth-making core for past tense claims, and that there *are* truth-makers outside the truth-making core for future tense claims. Of course, it is controversial to claim that the future exists; but that claim is *nowhere near* as controversial as the battery of assumptions that are required in order to claim that there are truthmakers for counterfactuals of freedom.

(6)

As I noted towards the end of the first section of this paper, even someone who accepts all of the metaphysical machinery that is required for Plantinga's free will defence might not accept the claim that Plantinga succeeds in describing a possible world in which 1.- 6. are all true. For someone who accepts all of Plantinga's metaphysical machinery might perfectly well think that, in circumstances in which a perfect being could not make a universe in which everyone always freely chooses the good, the perfect goodness of that being ensures that it will make no universe at all.

On the picture that we have been given, there will surely be possible worlds in which the perfect being engages in no creative activity involving free moral agents. Consider again a possible world in which, no matter which of the S_i the perfect being were to choose, the resulting S+J *would* contain free creatures who *all always* freely choose the bad. In this possible world, it seems to me that it is just obvious that a perfect being will not make any universe at all containing free creatures: it is just inconsistent

with the perfect goodness of a perfect being that it should knowingly make a world in which there is nothing but unrelieved moral evil.

Perhaps there are sceptical theists who might deny the claim that I have made here. Perhaps there are sceptical theists who will say: given our limited knowledge about possible goods, and possible evils, and the possible connections between them—and given the vastly greater knowledge which a perfect being has of possible goods, and possible evils, and the possible connections between them—how can we have any confidence at all in the judgment that a perfect being could not knowingly make a world in which there is nothing but unrelieved moral evil?[25] However, this response seems to me to be unbelievable: what sense can someone who makes this response be giving to the words "perfect goodness"?

Suppose, then, that we agree that there are some possible worlds in which a perfect being engages in no creative activity involving free moral agents *because* of the moral evils that would inevitably be contained in the universes available to the perfect being for creation. Then the following question naturally arises: in *which* possible worlds is the knowledge afforded by counterfactuals of freedom consistent with the creation of a universe containing free agents? (Given the libertarian conception of freedom, and given the assumption that the perfect being has libertarian freedom with respect to the creation of universes containing free agents, there must be possible worlds in which a perfect being engages in no creative activity involving free moral agents *even though* the perfect being is in a position to bring about universes containing free agents who all always freely choose the good. However, our present question need not lead us to make a detour through these further difficulties concerning the freedom of a perfect

being with respect to creation: all we want to know here is how good a universe containing free moral agents has to be before it is possible for a perfectly good being to make it.)

It seems plausible to claim that, if a perfect being is able to make a universe containing free creatures which all always freely choose the good, then the creation of such a universe does not conflict with the perfect goodness of the perfect being. But what of the following cases:

(a) universes in which there is at least one free creature that freely chooses the bad on at least one occasion?

(b) universes in which there is at least one free creature that freely chooses the bad on every occasion?

(c) universes in which every free creature freely chooses the bad on at least one occasion?

(Plantinga famously uses the expression “transworld depravity” to describe a situation in which every universe that it is open to the perfect being to make is of kind (c): the possible creatures that the perfect being can make suffer from transworld depravity at a given possible world if they go wrong in every possible universe in which they exist that it is open to the perfect being to make at that possible world.)

It seems to me that it is not in the least bit obvious that a perfectly good being *can* make a universe in any of these circumstances. After all, if the perfectly good being does not make a universe, then the possible world is never sullied by any kind of moral evil; on the other hand, if the perfectly good being does make a universe, then moral evil makes an appearance at that possible world. Given the choice between a

possible world in which there is no moral evil, and a possible world in which there is moral evil, why shouldn't we suppose that a morally perfect being will inevitably opt for the world in which there is no moral evil?

But—it will be replied—there are goods that are foregone if a universe of free agents is not created; and those goods outweigh the introduction of moral evil into the world.

How confident should we be that this is so? Suppose, for example, we take seriously the sceptical theist claims: (1) that we have very limited knowledge about possible goods, and possible evils, and the possible connections between them; and (2) that a perfect being would have vastly more knowledge about possible goods, and possible evils, and the possible connections between them. Then, it seems to me, we are in no position to judge whether a perfectly good being is able to make a world in which there is even the slightest amount of moral evil. And, if that's right, then—absent any other relevant considerations—we are in no position to determine whether 1.- 6. are logically consistent. (Perhaps this is a victory of sorts for the sceptical theist, even though it requires a substantial step back from the position that Plantinga defends.[26])

Suppose, instead, that we feel comfortable about our ability to determine what a perfect being would do in circumstances in which every universe that it can make contains some moral evil. What would it do? For what it's worth, my intuitive judgment is that it would make no universe at all. However, since I can see no persuasive argument to back up this judgment, the most that I am prepared to say is that there is room for further thought about this aspect of logical arguments from moral evil. If there were nothing else questionable about Plantinga's free will defence,

then it would probably be necessary to concede that there is decent support for the claim that the defence succeeds (where this decent support lies in the intuitive judgments of those who suppose that a perfect being could make a world that contains moral evil in circumstances in which it can make no other kind of world).

(7)

Even if the argument in the earlier sections of this paper is sufficient to cast doubt on Plantinga's free will defence, it does not follow that there is good reason for theists to be worried about the logical argument from moral evil. Since the difficulty upon which we focussed is due to the metaphysical machinery that is used in constructing Plantinga's response, a natural thought is to look to some other metaphysical framework upon which to hang a response. Perhaps such a framework is not so far to seek.

Focus again on possible worlds in which a perfect being is deliberating about whether or not to create a universe containing free agents. Suppose further that, prior to the actual making of a free decision by a free agent, there is no fact of the matter about what that free agent will do. Suppose, relatedly, that if a perfect being makes a universe in which there are free agents, then the perfect being does not *know* what those free agents will freely choose to do "until" they make the choices in question.[27] Suppose, finally, that there is a universe part S that it is open to the perfect being to make, and that it knows will form part of a universe in which there are free agents who make free choices. (There may well be many such universe parts; however, no harm will come from the pretence that there is only one such universe part.)

Consider all of the possible worlds in which the perfect being exists, and in which it chooses to make a universe that contains S as the part that is entirely up to the creative activities of the perfect being. In some of these possible worlds, all of the free agents will always freely choose the good. In other of these possible worlds, some of the free agents will sometimes freely choose the bad. And in yet other of these possible worlds, the other options described in previous sections of this paper will also be realised. By hypothesis, our perfect being has no way of knowing how the universe that contains S as a part will turn out. Hence, while it is clearly *possible* that, in making a universe which contains S as a part, the perfect being will make a universe in which everyone always freely chooses the good, this is not a matter that the perfect being can decide by fiat. If the perfect being makes a universe that contains S as a part, then it is an objectively chancy matter whether the universe ends up containing moral evil.

Plainly enough, we have here the materials for a variant of the free will defence that avoids the kind of objection that was made against Plantinga's free will defence in the second section of this paper. On the assumption that a perfectly good being is able to make a universe that contains S as a part in the circumstances described, it is clear that the mere unfolding of objectively chancy events can then bring about a world in which 1.- 6. are all true.

However, while this version of the free will defence does not fall foul of the metaphysical principle concerning the grounding of counterfactuals, it still has to face the other objections that were raised in the earlier parts of this paper. First, it is clear

that any proponent of this defence has to give up on strong versions of the principle of sufficient reason (since, of course, proponents of strong versions of the principle of sufficient reason deny that there can be objectively chancy events). And, second, there is a serious question to ask about whether, in the envisaged circumstances, a perfect being can make a universe that contains S as a part. After all, in the circumstances, it is possible that the perfect being will make a universe in which everyone always freely chooses the bad—and it is not clear that a perfect being would take this kind of risk, however unlikely the awful outcome is.

Although it involves a substantial digression from our main argument, perhaps it is also worth noting that, even in the circumstances envisaged, it *might* be that a perfect being could do something very much like choosing that there should be a world in which everyone always freely chooses the good. Suppose that there is nothing to prevent a perfect being in a given possible world from making more than one "physical universe", i.e. more than one maximal spatio-temporally connected sub-part of the larger universe. Then, by making enough physical universes, a perfect being can make it as close to certain as it pleases that there will be at least one physical universe in which everyone always freely chooses the good. Suppose, further—as many theists do—that the perfect being is required to sustain physical universes in existence at every moment at which those physical universes exists. If the perfect being allows any physical universe to pass out of existence as soon as a wrong choice occurs within it, then the perfect being can choose to make an ensemble of physical universes (i.e. a universe)—including, almost certainly, some physical universes that are never allowed to pass out of existence—in which, with vanishingly few exceptions, everyone always freely chooses the good.[28] Perhaps it might be objected

that it is inconsistent with the perfect goodness of a perfect being to allow physical universes to pass out of existence because of the wrong choices made by denizens of those physical universes; it is not *utterly* obvious to me that this is so. In particular, it seems to me that those who are sympathetic to the views which Epicurus and Lucretius take towards the alleged harm of death will be hard pressed to say why a perfect being could not act in the way outlined. Of course, there is much more that could be said here; but this digression would turn into a paper of its own if we tried to pursue it.

(8)

Mackie (1955) provides the materials for the following argument (which sets out the informal argument of the first section of this paper in a more systematic fashion). In order to state the argument, we need to introduce some new vocabulary. We begin with the thought that, amongst possible universes, there is a class of possible universes—the *A-universes*—that are non-arbitrarily better than all of the other universes that contain free agents. The A-universes are some, but by no means all, of the universes in which there are free agents who all always freely chooses the good.

1. Necessarily, a perfect being can just choose to make an A-universe. (Premise)

2. Necessarily, A-universes are better than non-A-universes in which there are free agents. (Premise)

3. Necessarily, if a perfect being chooses between options, and one option is non-arbitrarily better than the other options, then the perfect being chooses that option. (Premise)

4. Hence, necessarily, if a perfect being makes a universe that contains free agents, then it makes an A-universe. (From 1, 2, 3)

5. Our universe contains free agents, but it is not an A-universe. (Premise)

6. Hence, it is not the case that a perfect being made our universe. (From 4, 5)

The strategy behind free will defences is to deny the first premise of this argument: because of the libertarian nature of freedom, it is not necessarily true that a perfect being can just choose to make a world in which everyone always freely chooses the good. (Plausibly, on a compatibilist analysis of freedom, it would be necessarily true that a perfect being can just choose to make a world in which everyone always freely chooses the good. While this claim is clearly in need of argumentative support, I shall not try to provide such support here.[29]) While the remaining premises in the argument—i.e. 2, 3, and 5—are not incontestable, I think that there is likely to be widespread agreement amongst both theists and non-theists concerning their plausibility; at any rate, I don't propose to examine these premises further in the present paper. On the assumption that I am right about the status of these further premises, then the success or failure of this argument turns entirely on the debate about the analysis of freedom. Since it is a controversial matter whether freedom can be given a libertarian analysis, it is a controversial matter whether there is a successful

reply to this variant of Mackie's argument, *when* the argument is supplemented with subsidiary arguments that support premise 1.[30]

I take it that the above considerations are sufficient to establish that—contrary to the claims of many—there is still genuine life in Mackie's argument. Suppose, though, that we are persuaded that premise 1 renders the argument unpersuasive. Does this mean that all arguments of this kind must be abandoned? I don't think so. For, even if one can defend a metaphysical position according to which it is *possible* that a perfect being is not able to choose to make a universe in which everyone always freely chooses the good, it is much less clear that one can defend a metaphysical position according to which it is at all *likely* that there is a perfect being that was not able to choose to make a universe in which everyone always freely chooses the good.

The intuitive idea here is very simple. Clearly enough, agents who are strongly disposed towards doing good can nonetheless act freely (and, moreover, their freedom can be significant). Indeed, agents who are separately and collectively as strongly disposed as you please towards doing good can nonetheless act freely. Among the universe parts that it is open to a perfect being to make, there are universe parts in which the free agents that arise are collectively as strongly disposed as you please towards doing good. But, on the one hand, if we accept Plantinga's metaphysical picture (outlined in section 1 above), then (arguably) it follows that, if there is a perfect being, then it is as close to certain as you please that that being *was* able to choose to make a universe in which everyone always freely chooses the good; and, on the other hand, if we accept the alternative metaphysical picture (outlined in section 4 above), then (arguably) it follows that, if there is a perfect being, then it *was* able to

choose a world in which it was as close to certain as you please that everyone would always freely choose the good and as close to certain as you please that if not everyone always freely chose the good, then these departures from optimal choice would be minimal.[31]

Constructing arguments with these new claims as initial premises is not a straightforward matter. The specimens that I am about to offer are pretty plainly imperfect; however, I am confident that there is something to be learned even from these imperfect attempts.

On the one hand, if we adopt the metaphysical framework which Plantinga defends, then we can construct arguments such as the following (Argument 1):

1. It is as close to certain as you please that a perfect being can choose to make a universe in which everyone always freely chooses the good. (Premise)

2. Necessarily, universes in which everyone always freely chooses the good are non-arbitrarily better than universes in which someone sometimes freely chooses the bad. (Premise)

3. Necessarily, if a perfect being chooses between options, and one option is non-arbitrarily better than the other options, then the perfect being chooses that option. (Premise)

4. Hence, it is as close to certain as you please that, if a perfect being makes a universe, then it makes a universe in which everyone always freely chooses the good. (From 1, 2, 3)

5. It is not the case that everyone always freely chooses the good. (Premise)

6. Hence, it is as close to certain as you please that our universe was not made by a perfect being. (From 4, 5)

On the other hand, if we adopt the alternative metaphysical framework outlined in section 4 of the present paper, then—as foreshadowed in footnote 31—we need to take account of the following consideration. When deliberating about which universe part to make, a perfect being will be concerned not only with the likelihood that the universe part belongs to a universe in which there are free agents who always freely choose the good, but also with the likelihood that, if the free agents do not always freely choose the good, then they depart from perfect goodness rarely and in relatively unimportant ways. That is, when selecting a universe part, a perfect being will select an *A*-part*, i.e. a universe part that is *both* almost certain to result in a universe in which there are free agents who all always freely choose the good *and* almost certain to result in a universe in which there is only minimal departure from universal choice of good if there is such departure.

Given this consideration, we can go on to construct arguments like the following:

(Argument 2):

1. Necessarily, a perfect being can choose to make an A*-part. (Premise)

2. Necessarily, A*-parts are preferable to other parts that give rise to universes in which there are free agents. (Premise)

3. Necessarily, if a perfect being chooses between options, and one option is non-arbitrarily better than the other options, then the perfect being chooses that option. (Premise)

4. Hence, necessarily, if a perfect being chooses a universe part, then it will choose an A*-part. (From 1, 2, 3)

5. It is not the case that our world involves no more than minimal departure from universal choice of good. (Premise)

6. Hence, it is as close to certain as you please that our universe was not made by a perfect being. (From 4, 5)

These arguments are pretty natural developments from the argument of Mackie (1955), and so will likely be heir to whatever difficulties are attached to that argument (apart from considerations involving libertarian freedom). Of course, I have suggested that there are no such further difficulties for Mackie's argument. Nonetheless, these arguments are also plainly subject to difficulties of their own; and it is some of these potential difficulties which will be the focus of the final section of this paper. (This final section of the paper is a digression from the main line of argument of the paper. I

take it that the argument of the present section is already sufficient to vindicate the claim that discussion of the kinds of considerations raised by Mackie is far from exhausted, *even* if the argument set out at the beginning of this section is vitiated by its reliance upon a compatibilist analysis of freedom.)

(9)

It seems to me that there is a clear sense in which claim 4 in Argument 1 is correct: if we arbitrarily select a possible world containing a perfect being, then it is as close to certain as you please that, in that world, the perfect being is able to make a universe in which everyone always freely chooses the good. However, the clear sense in which (I take it that) this claim is correct relies upon the fact that our "arbitrarily selecting" a possible world requires that we do not have any other contingent or *a posteriori* information about that world. But once this is recognised, it seems fairly clear that the inference from 4 and 5 to 6 is no good: even given that it is certain that there is moral evil in the world, the fact that, *on the basis of a priori information alone*, it is as close to certain as you please that, in our world, a perfect being would have been able to make a universe in which everyone always freely chooses the good does not entail that, *on the basis of all relevant information*, it is as close to certain as you please that our universe was not made by a perfect being. Taking all of the relevant evidence into account, one might rather believe that our universe was made by a massively unlucky perfect being.

A similar kind of point can be made in connection with Argument 2. Again, it seems to me that there is a clear sense in which claim 4 is correct: if a perfect being makes a

universe part, then it must make an A*-universe part. However, the clear sense in which (I take it that) this claim is correct relies upon the fact that we are only considering how things stood "when" the creative action occurred. Once this is recognised, it seems fairly clear that the inference from 4 and 5 to 6 is no good: even given that it is now certain that our world involves more than minimal departure from universal choice of good, the fact that, if the creative action occurred, it was *"then"* as close to certain as you please that the universe would involve no more than minimal departure from universal choice of good does not entail that it is *now* as close to certain as you please that our universe was not made by a perfect being. Taking all of the relevant information into account, one might rather believe that our universe was made by a massively unlucky perfect being.

No doubt most readers will now have guessed where this digression is heading. Even if we are prepared to accept that it is quite all right to believe in the kind of unexplained—and, presumably, unexplainable—bad luck which it seems must be invoked by the perfect being theist, there are consequences of this invocation which perfect being theists must face. In particular, perfect being theists can have no truck with any of the fine-tuning arguments for intelligent design which have received so much attention in recent times (since, by parity of reasoning, they shall have to allow that it is perfectly appropriate to respond to these arguments with the claim that it is just a matter of unexplained, and most likely unexplainable good luck that our universe turned out to be hospitable to life, and that it turned out to contain the complex kinds of organisms which it in fact contains[32]). As I noted earlier in connection with the logical argument from moral evil, *defences* against arguments

bring with them commitments that cannot be ignored when one comes to mount positive arguments of one's own.

Doubtless, some theists will claim to be distinctly unimpressed by all of this. Even if it were true that the cost of meeting arguments from moral evil is that all decent *a posteriori* arguments for the existence of a perfect being must be foregone[33], it might well be insisted that it remains open to theists to follow Plantinga (1979)[34] in claiming that rational belief in a perfect being does not require any kind of argumentative support. Suppose that's right, i.e. suppose that there is some sense in which belief in a perfect being can be properly basic. Nonetheless, it seems that the *kinds* of considerations that we have been developing might well suffice to show that there is no good reason for those who are not already convinced of the truth of perfect being theism to become perfect being theists. While arguments from moral evil may not show that perfect being theists are irrational, there is at least some reason to think that these arguments can form an important plank in a case that shows that rational considerations are insufficient to move anyone to *become* a perfect being theist.

Of course, I do not think that the brief argument of this section—and the earlier digression about principles of sufficient reason—conclusively establishes the claim that is mentioned at the end of the previous paragraph. However, as I said at the beginning of the paper, my main aim is a moderate one, namely, to show that there is more to be learned from arguments concerning moral evil than many philosophers are currently prepared to concede. If there is anything at all to the line of thought developed in the current section of this paper, then there is *still* plenty of life left in the kinds of considerations that are appealed to in Mackie (1955).[35]

10

What do I think can be learned from the preceding discussion? *First*, it is highly questionable whether Plantinga (1974) provides a satisfactory response to the standard "logical assertion" about moral evil. *Second*, there is a logical argument from moral evil in Mackie (1955) that (plausibly) stands or falls with a compatibilist analysis of freedom. *Third*, the acceptance of a libertarian analysis of freedom imposes non-trivial constraints on the kinds of principles of sufficient reason that one can endorse (and, hence, on the kinds of cosmological arguments that one can promote). *Fourth*, there are probabilistic analogues of the more powerful logical argument from Mackie (1955) that bear serious comparison with currently popular "fine-tuning" arguments for intelligent design. I do not suppose that this exhausts the important points to be established by a serious reconsideration of the arguments in Mackie (1955); I, for one, am keen to retrieve that paper from "the dustbin of philosophical fashions".

---

[1] Plantinga, A. (1974) The Nature of Necessity Oxford: OUP

[2] Mackie, J. (1955) "Evil and Omnipotence" Mind **64**, 200-12.

[3] Here are a few examples taken from papers collected together in D. Howard-Snyder (ed.) The Evidential Argument from Evil Bloomington: Indiana University Press.:

"It is now acknowledged on (almost) all sides that *the logical argument is bankrupt*. … [The] inductive argument from evil is in no better shape than *its late lamented*

cousin." (Alston, W. (1991/1996) "The Inductive Argument from Evil and the Human Cognitive Condition" pp.97-125, at 97, 121, *my emphasis*)

"Like logical positivism, Mackie's argument has found its way to the dustbin of philosophical fashions." (Howard-Snyder, D. (1996) "Introduction" pp. xi-xx, at xiii)

"It is widely conceded that there is nothing like straightforward contradiction or necessary falsehood in the joint affirmation of God and evil. And (as I see it) rightly so." (Plantinga, A. (1988/1996) "Epistemic Probability" pp.69-96, at 71, n.3)

"It used to be widely held that evil … is incompatible with the existence of God: that no possible world contained both God and evil. So far as I am able to tell, this thesis is no longer defended." (Van Inwagen, P. (1991/1996) "The Problem of Evil, the Problem of Air, and the Problem of Silence", pp.151-174, at p.151)

For a dissenting voice, compare: "I argue that the logical problem posed by moral evil is still with us." (Gale, R. (1996) "Some Difficulties in Theistic Treatments of Evil" pp.206-218, at p.206)

It should be noted that if we are using alethic – as opposed to, say, doxastic— modalities, then there surely are *many* philosophers who continue to maintain that there is no possible universe made by a perfect being that contains moral evil. However, it is now much harder to find philosophers who are prepared to maintain that one cannot consistently believe that there are possible universes containing evil that are made by a perfect being. And it is no easier to find philosophers who are

prepared to maintain that there are demonstrations that establish that there is no possible universe made by a perfect being that contains moral evil. I take it that Alston, Howard-Snyder, Plantinga, van Inwagen *et. al.* think that there are conclusive arguments that show that there are possible universes containing evil that are made by a perfect being (and that it is simply irrational to suppose that the existence of moral evil is logically incompatible with the existence of a perfect being who created our universe). However, nothing that I go on to argue in the present paper depends upon this assumption.

[4] Perhaps it is worth noting here that I certainly do not think that there are extant arguments from moral evil of such strength that perfect being theists who are not persuaded by these arguments to give up on their perfect being theism are *eo ipso* convicted of irrationality. If good arguments are required to be rationally compelling for all rational people not already disposed to accept their conclusions, then no extant argument from moral evil is good. However, arguments that fail this stringent requirement—and hence which are not successful *qua* arguments—*might* have other virtues. (The claim that I have just made might be thought to necessitate some alterations to the argument of Oppy, G (2002) "Arguing about the *Kalam* Cosmological Argument" Philo 5, 1, 34-61. I do not believe that this is the case; however, I shall not try to argue for this belief here.)

[5] The following premises—2-5—can be taken to implicitly define the expression 'perfect being'. Of course, the defining terms—'omnipotent', 'omniscient', 'perfectly good', etc.—also require explanation; but, for the purposes of the present paper, I shall suppose that these terms are sufficiently well understood.

[6] I shall not attempt to provide any analysis of the notion of ‘moral evil’. Doubtless, there are meta-ethical conceptions on which claim 6 fails to express a proposition. However, on those meta-ethical conceptions, the occurrence of ‘perfect goodness’ in 4 ensures that it, too, fails to express a proposition (and, hence, given the previous footnote, the expression ‘perfect being’ fails to be well-defined). An argument for the logical impossibility of the joint truth of 1-6 thus ought to proceed in two stages: first, we consider how things stand if one or more of 1-6 fails to be truth-apt; second, we consider how things stand on the assumption—or, if you prefer, under the pretence—that 1-6 are indeed all truth-apt. For the remainder of this paper, I shall assume—or, perhaps, pretend—that 1-6 are indeed all truth-apt, and that 6 is true.

[7] It might be said that Plantinga actually argues for a stronger claim, viz. that there is a sense in which the existence of moral evil in the actual world may have been necessitated by the creative activities of the perfect being that made our world. However, it remains true that Plantinga does attempt to describe a logically possible world in which 1-6 are all true; and it is important to bear in mind that Plantinga does not deny that it is logically possible for 1-5 to be true in a world in which there is no moral evil.

[8] The “perhaps” in the text should be taken to indicate some kind of doxastic—rather than some kind of alethic—possibility: while the claim in question may not be logically possible, it is a claim which reasonable people can reasonably believe. I do not believe that it is logically possible that every world contains a perfect being; however, I do think that this is something that a reasonable person can believe. There

are obvious difficulties that will arise for my discussion if one believes that every possible world contains a perfect being (though I do think that these difficulties can ultimately be finessed by further appeal to the distinction between alethic and doxastic modalities).

Perhaps I should add here that, unlike many philosophers, I do not believe that there is a genuine distinction to be made between (broadly) logical possibility and metaphysical possibility. Those who suppose that there is such a distinction should disambiguate my writings in whatever way maximises the likelihood of truth!

[9] Perhaps there is only one such world. There are tricky questions here about, e.g., the *thoughts* of a perfect being: given that it has libertarian freedom, there is at least some reason to suppose that a perfect being can have different thoughts in situations in which it is the only existent. To pursue this matter further, we would need to worry about whether a perfect being can have *any* thoughts.

[10] In order to accommodate familiar talk about the possibility that a perfect being might make more than one "universe", we need to introduce a further distinction. The idea that we want to accommodate is that a perfect being might make a universe that consists of more than one (more or less) causally isolated sub-universes. At least roughly speaking, sub-universes are maximal causally connected aggregates of contingent states of affairs. This is rough, in part, because the perfect being's creative activities are excluded from the "maximal" aggregates: the intuitive idea is that there is no causal interaction between sub-universes, and no causal connections other than those that run through the perfect being. For the purposes of the present paper, no

harm will come from adoption of the assumption that perfect beings make no more than one big-bang (sub-)universe.

[11] The intuitive picture is something like this. The libertarian conception of freedom requires that we can make sense of the idea that, at the time at which an agent chooses to act, the outcome of the choice is not already fixed but is rather "up to the agent". So, the libertarian conception of freedom requires a distinction between propositions whose truth-value at T is already fixed by the world prior to T, and propositions whose truth-value at T is not already fixed by the world prior to T. (If, prior to T, it is already *fixed* that the proposition "X does A at T" is true, then—on the libertarian conception of freedom—the agent does not have a genuine choice about whether to do A at T.) Consequently, given that we adopt the libertarian conception of freedom, we must be able to distinguish between those propositions that are true at T that are already fixed—or made true—by the world prior to T, and those propositions that are true at T but whose truth-value is not fixed by the world prior to T. We call the former set of propositions—or perhaps some specially distinguished subset that suffices to generate the whole set, e.g. by closure under entailment—the *truth-making core* of the world prior to T. (Note that we assume that the laws are part of the truth-making core at any time; and that the laws cannot be different at different times. Without these assumptions, it is very hard to make sense of the dispute between compatibilists and libertarians within the established theoretical framework.)

[12] Strictly, the qualification is redundant: if it is impossible for the agent to do A at T in W, then there *is* something that makes it true that the agent does not do A at T in W.

---

[13] For the purposes of this discussion, I am supposing that it is logically impossible for there to be backwards causation. If this assumption is not made, then there are elaborate epicycles that need to be added—but nothing of any importance for the main argument under consideration is required to be changed.

[14] Consider worlds that contain only one free creature X that makes N free choices (each of which has two possible outcomes: RIGHT and WRONG). Since the choices are all free, no one choice determines the results of any of the others. Ignoring other features of these worlds, there are 2^N different worlds: one in which X always goes right, one in which X always goes wrong, and 2^{N-1} in which X sometimes goes right and sometimes goes wrong. Of course, this little argument hardly suffices to show that *most* of the S+J are universes in which there are free creatures who sometimes freely choose the good and sometimes freely choose the bad—but it surely does suffice to suggest that there will be *many* such universes.

[15] In Plantinga's terminology, the perfect being cannot "weakly actualise" any of the S_1+J_{1k}, for $k\neq1$.

[16] Plantinga (1974), pp.166, 171.

[17] Suppose that T is an exhaustive list of the truths about a world w. Any subset S of T such that every member of T is entailed by some collection of the truths in S is a supervenience base for w. Any S for which there is no S' such that $S\supset S'$, where S and S' are both supervenience bases for w, is a minimal supervenience base for w. (There

is an analogous definition of "minimal truth-making core"—cf. the last sentence in footnote 9.) It is a substantive claim that worlds have minimal supervenience bases or minimal truth-making cores; however, we shall not pursue questions about "turtles all the way down" alternatives here. (Again, the effect of such a pursuit is merely to add epicycles to the discussion.)

[18] Armstrong, D. (1989) "C. B. Martin, Counterfactuals, Causality, and Conditionals" in J. Heil (ed.) Cause, Mind and Reality: Essays Honouring C. B. Martin Dordrecht: Kluwer, pp.7-15

[19] See Bigelow, J. (1988) The Reality of Numbers Oxford: OUP, for defence of the related claim that truth supervenes upon being; and Lewis, D. (1999) Papers in Metaphysics and Epistemology Cambridge: CUP, for various approving mentions of this idea.

[20] It is perhaps worth noting at this point that Plantinga assumes that the perfect being is *unable* to make a world in which everyone always freely chooses the good if it is true that for any S that it is open to the perfect being to make, *were* it to make S, not everyone *would* always freely choose the good in the resulting S+J. But if it is true that I *would* choose the bribe *were* I offered it, then surely it follows that I am *unable* to refrain from accepting the bribe in exactly the same sense in which the perfect being is *unable* to make a world in which everyone always freely chooses the good.

[21] Of course, it doesn't follow that there are no senses of "explanation" in which libertarian choices can be "explained". If Jones likes chocolate more than strawberry, then we can appeal to that in "explaining" why she chooses chocolate. But, according to libertarians, there is a possible world in which Jones opts for strawberry, even though she goes through exactly the same process of deliberation, has exactly the same preferences, etc. When we contrast the actual world with this (allegedly) possible world, the libertarian has no resources for explaining why Jones process of deliberation, preferences, etc. actually resulted in a choice of chocolate and not strawberry (since, *ex hypothesi*, there is no relevant difference to which appeal can be made).

[22] The following discussion has some affinity to the discussion in Gale, R. (1991) On The Nature and Existence of God Cambridge: CUP, at 152-68. However, Gale seems content not to challenge the assumption that there can be bare counterfactual truths (see, e.g., p.144). It is worth asking what Gale would now say about this matter, given his new found enthusiasm for strong versions of the principle of sufficient reason (cf. Gale, R. and Preuss, A. (1999) "A New Cosmological Argument" Religious Studies **35**, 461-76.)

[23] Although I have given no argument that there could be no other kinds of truthmakers for counterfactuals of freedom—under the assumptions made in the present section—I do think that it is very hard to find plausible suggestions about where such truthmakers might be found.

[24] See, for example, Flint, T. (1998) Divine Providence Ithaca: Cornell University Press, at 121-137, and references therein.

[25] See, for example, Bergmann, M. (2001) "Sceptical Theism and Rowe's New Evidential Argument from Evil" Nous **35**, 2, 278-296. For a partial critique of sceptical theism, see Almeida, M. and Oppy, G. (2003) "Sceptical Theism and Evidential Arguments from Evil" Australasian Journal of Philosophy 81, 4, 496-516.

[26] Of course, if we are so irremediably ignorant about perfect beings, then it is hard to see how there could be any kind of argument capable of bringing those not already persuaded that there is a perfect being to accept the claim that there is a perfect being. (If almost any evidence is compatible with the existence or non-existence of a perfect being, then almost no evidence can tell in favour of the existence of a perfect being.) But if there is no chain of reasoning which can lead reasonable non-believers to the conclusion that there is a perfect being, then it surely follows that there can be reasonable non-believers (contrary to the doctrinal commitments of a substantial number of theists).

[27] Does this supposition amount to giving up on the assumption that the perfect being is omniscient? No, of course not: at most, omniscience requires only knowledge of that which it is logically possible to know; and, *independent* of the free choices of the free agents, there just isn't anything that it is logically possible to know about these choices.

[28] Of course, there is *some* moral evil in this set up, and so there are still questions about whether the amount of moral evil is too much for a perfect being to countenance. But—to anticipate considerations to be taken up in the later sections of the current paper—this set up certainly seems to have a better *probable* ratio of good to evil than the set up in which just one universe is made.

[29] For some hints about how this claim might be defended, see Nagasawa, Oppy, and Trakakis (forthcoming) "Salvation in Heaven?" Philosophical Papers.

[30] Of course, as things stand, the controversy about Premise 1 is enough to establish that 1.-6. is not, itself, a successful argument for the conclusion that out universe was not made by a perfect being.

[31] The need for the second condition here was impressed upon me by Geoff Brennan, with help from Peter Godfrey-Smith. There is subsequent discussion of the need for this condition in the main text.

[32] We might suppose that a standard 'fine-tuning' argument takes one of the following two forms:

1. It is very close to certain that, if some fundamental parameters in a big bang universe are fixed arbitrarily, then that universe contains no life.
2. Our big bang universe contains life.
3. (Hence) It is very close to certain that those fundamental parameters of our universe were not fixed arbitrarily.

1. If some fundamental parameters in a big bang universe are fixed arbitrarily, then it is very close to certain that that universe contains no life.
2. Our big bang universe contains life.
3. (Hence) It is very close to certain that those fundamental parameters in our universe were not fixed arbitrarily.

We can say against each of these arguments just what was said against each of our probabilistic arguments from moral evil: while the initial premise may very well be plausible when nothing other than *a priori* information is taken into account, it is just illegitimate to infer from this fact that the conclusion is plausible when all relevant information is taken into account. (Of course, I don't say that this is the *only* thing to be said against these fine-tuning arguments. However, it would seem to be a useful response for those non-theists who are unsure what else to say in response to these arguments when these arguments are propounded by perfect being theists.)

[33] Of course, the claim that is being entertained here goes far beyond anything that I have tried to argue for: it would take a much more extended argument to show that all *a posteriori* arguments for the existence of a perfect being are neutralised by the defensive measures that must be taken in order to have adequate replies to arguments from evil. Nonetheless, it does seem to me that there is a question here worth asking: not everyone has noticed that there may be more than one way in which arguments from evil can advance the cause of non-theists. (Plainly enough, theists can make similar kinds of points—about, say, responses to arguments for the inadequacy of naturalism—against those non-theists who are naturalists.)

---

[34] Plantinga, A. (1979) “Is Belief in God Rational?” in C. Delaney (ed.) Rationality and Religious Belief Notre Dame: UND Press, 2-29

[35] Earlier versions of this paper were presented to staff seminars at Monash, Melbourne, Latrobe, and RSSS. I am grateful to all who asked questions at those presentations; the paper is much improved as a result.