

cause of God?” Typical solutions argue that the case of the First Cause is different in some way that is not merely *ad hoc* from the cases to which the explanatory principle was applied.

The final difficulty for cosmological arguments is the Gap Problem.¹ Granted there is a First Cause, but does anything of religious interest follow? There is a gap between the statements that there is a First Cause and that there is a God. Aquinas, in his Five Ways, proves the existence of an unmoved mover and then says: “*et hoc omnes intelligent Deum*” (“and all understand this to be God”). Some critics have taken this to be his way of papering over the difficulty of moving from a First Cause to God; however, that reading is mistaken in light of the fact that succeeding sections of the *Summa Theologiae* give careful and elaborate arguments that the First Cause is wholly actual, unchanging, simple, one, immaterial, perfect, good, and intelligent. Rather, Aquinas is simply marking the fact that the *theist* will recognize the unmoved mover to be God. Aquinas knows that an argument that the First Cause has, at least, some of the attributes of the God of Western monotheism is needed and offers such an argument.

The solutions to the Glendower and Regress problems tend to go hand in hand and, probably, the best way to classify cosmological arguments is by how they address these problems. There are then three basic kinds of cosmological arguments: *kalam*, Thomistic, and Leibnizian. The *kalam* and Thomistic arguments posit an intuitively plausible Causal Principle (CP) that says that every item of some sort – for example, event, contingent being, instance of coming-into-existence, or movement – has a cause. The arguments then split depending on how they handle the Regress Problem. The *kalam* argument proceeds by arguing, on *a priori* or *a posteriori* grounds, that the past is finite and hence, in fact, no infinite regress occurred. The Thomistic argument, exemplified by Aquinas’ first three ways, does not rule out the possibility of an infinite past but uses a variety of methods to argue against the hypothesis that there is an infinite regress of causes with no First Cause. The most distinctive of these methods is an attempt to show that there is an intrinsic distinction between intermediate and nonintermediate causes, where an intermediate cause of *E* is an item *C* that is itself caused by something else to cause *E*, and that this distinction is such that intermediate causes are, of necessity, dependent for their causal activity on nonintermediate causes, which then end the regress.

Leibnizian arguments, on the other hand, invoke a very general explanatory principle, such as the PSR, which is then applied to the cosmos or to some vast cosmic state of affairs, or else a nonlocal CP that can be applied to an infinite chain or the universe as a whole. In the PSR-based versions, the Regress Problem is typically handled by showing that an infinite chain of causes with no First Cause fails to explain why the *whole chain* is there. The main challenge for Leibnizian arguments here is to argue for an explanatory principle or CP that is (a) plausible, (b) applicable to the cosmic state of affairs in question, and (c) not so strong as to lead to implausible conclusions such as the denial of contingency or of free will. In this chapter, I shall defend several Leibnizian arguments.

The basic Leibnizian argument has the following steps:

- (1) Every contingent fact has an explanation.
- (2) There is a contingent fact that includes all other contingent facts.
- (3) Therefore, there is an explanation of this fact.

1. I got the term from Richard Gale.

- (4) This explanation must involve a necessary being.
- (5) This necessary being is God.

We shall see, however, that the first step, the assumption of the PSR, can be modified in various ways, with the resulting argument maintaining the distinctive feature of Leibnizian arguments that the relevant explanatory principle or CP is to be applied to a global state or proposition.

2. The PSR

2.1. *The scope of the PSR*

For simplicity, I shall stipulatively use the term “fact” for a true proposition. The PSR states that every fact, or every contingent fact, has an explanation, and this is the standard tool in Leibnizian arguments for handling the Glendower and Regress problems.

Some authors restrict the PSR to contingent facts. The advantage of a restriction to contingent facts is that we do not know very much about how the explanation of necessary truths works and, hence, may not be in a position to justify the PSR for necessary truths. To explain the Pythagorean Theorem, presumably, I should prove it from the axioms. But which proof counts as explanatory? Which axioms are the right ones to start from? Is there a fact of the matter here?

On the other hand, maybe the case of necessary facts is not a real worry, for it might be that any necessary truth p can be explained by citing its necessity: p holds because p necessarily holds. This leads into a regress since that p necessarily holds will also be a necessary truth by Axiom S4 of modal logic; but perhaps this regress is somehow to be distinguished from vicious ones.

Alternatively, the defender of an unrestricted PSR can say that while we do not yet know how the explanation of necessary truths works, we do know some cases of it. For instance, it might be that the proposition that $1 = 1$ is self-explanatory, namely explained by the very same proposition $1 = 1$, while the proposition that, necessarily, $1 = 1$ is explained by the proposition that $1 = 1$ together with the fact that mathematical truths are necessary truths. The necessary truth that all dogs are mammals, assuming this is indeed metaphysically necessary, is explained by the genetic similarity between dogs and the first mammals, together with some necessary truths about how biological classification works. The necessary truth that making false promises is wrong might be explained by the fact that falsely promising treats the promisee as a mere means. In other words, while we have no general account of the explanation of necessary truths, we do have many examples. And, anyway, the requirement that we have a general account of explanation would also be a problem for a PSR restricted to contingent propositions, since it is not clear that we yet have a general account of explanation of contingent propositions, although we have many clear examples.

2.2. *Why should we believe the PSR?*

2.2.1. Self-evidence

Many of those who accept the PSR do so unreflectively because they take the PSR to be self-evident. I do not think that there is any good argument against the propriety of doing

so. We are perfectly within our epistemic rights to accept the Law of Excluded Middle (LEM), namely the claim that for all p we have p or not- p , because of the self-evidence of LEM, without needing any further argument for it. However, it will be of no use to opponents of the PSR or of the LEM to be told that the claim they deny is self-evident to us. Presumably, the claim is not self-evident to them, and we can all agree that there are many things that people have *claimed* to be self-evident that, in fact, are false, so the fact that the claim is said by us to be self-evident does not provide these opponents with much reason to accept it. There may be a presumption that what people take to be self-evident is, in fact, more likely true than not, but this presumption is often easily defeated.

One might think that philosophical disagreement about the PSR shows that the PSR is *not* self-evident, or at least that those of us who take it as self-evident should not see this as providing any reason to believe it to be true. Otherwise, how could competent philosophers such as David Hume or Graham Oppy fail to see it as self-evident? Or, worse, how is it that some of these philosophers take as self-evident claims *incompatible* with the PSR?

If we think we should accept the LEM because of its self-evidence despite some brilliant intuitionist mathematicians' denials of it, we will be unimpressed by this argument. And it is not clear on what grounds we could accept the LEM other than self-evidence. Is there some inductive argument like: "For many propositions p , we have concluded that the LEM holds. Hence, the LEM holds for all propositions p "? I doubt it. The problem is that an inductive argument of the form "Many F s are G s, thus all F s are G s" is epistemically close to worthless by itself. Many dogs are spotted, thus all dogs are spotted? We would do slightly better if we could show that most F s are G s, although even that would be very weak ("Most humans are female, thus all humans are female"). But how would we check that the LEM holds for most propositions? To check that the LEM holds for a proposition is, presumably, to determine that this proposition is true or to determine that this proposition is false, since in either case, the truth of the LEM follows for the proposition. But most propositions are such that we cannot determine whether they are true or false.

In any case, the argument from philosophical disagreement is weak. It might be that our judgment as to what is or is not self-evident is fallible, and Hume and Oppy have simply judged wrongly. Or it might be that it is possible to be talked out of seeing something as self-evident, just as it is possible to be (rightly or wrongly) talked out of all sorts of commonsensical beliefs. Finally, it could be that the PSR's opponents have failed to grasp one or more of the concepts in it due to their substantive philosophical positions. Thus, Hume's equating constant conjunction with causation suggests that he does not have the same concept of causation as I do – that he is talking of something different – and the fact that he thinks causation *thus understood* yields explanations, as well as his belief that infinite regresses can be explanatory, show that his concept of explanation is different from mine. Differences in views of modality are also relevant. As a result, it is far from clear to me that Hume has even grasped the PSR in the sense that I assign to it. And if not, then his failure to see it as self-evident is irrelevant.

I can give a similar story about Hume's seeing as self-evident propositions that are incompatible with the PSR, such as that no being's existence is necessary.² Hume's concept of the necessity of p is that a contradiction can be proved from the denial of p . If LEM is

2. This is incompatible with the PSR, given the other ingredients in the cosmological argument.

true, this is equivalent to equating necessity with provability. But defenders of the Leibnizian cosmological argument typically use a notion of *broadly logical* necessity when they claim that God is a necessary being, and broadly logical necessity is weaker than provability.

At this point, it may seem as if the defense of the self-evidence of the PSR destroys all possibility of philosophical communication. If philosophers all mean different things by the same terms, how can they even disagree with one another? Two points can be made here. The first is that in many cases, when philosophers use a word such as “cause,” they *both* mean by it what ordinary language does *and* they have an account of what the word says which they think is faithful to the ordinary meaning. And if this is true, then when one philosopher says “A causes B” and the other says “A does not cause B,” there is a genuine disagreement between them even if their analyses of causation are different, since the first philosopher holds that A causes B in the ordinary English sense of “causes” (which he rightly or wrongly thinks is identical with his analysis of the term) and the second denies this. Second, disagreement is possible because even though philosophers may use the term “causes” differently, they will tend to agree on some entailments, such as that if A causes B, then both A and B occurred and B’s occurrence can be explained, at least in part, in terms of A’s occurrence. So differences in meaning do not undercut philosophical communication, but they seriously damage the argument against self-evidence.

Self-evidence might well give those of us to whom the PSR is self-evident a good reason to believe it. But if we want to convince others, we need arguments.

2.2.2. The epistemological argument

This argument is based on the ideas of Robert Koons (1997), although I am simplifying it. Starting with the observation that once we admit that some contingent states of affairs have no explanations, a completely new skeptical scenario becomes possible: no demon is deceiving you, but your perceptual states are occurring for no reason at all, with no prior causes.

Moreover, objective probabilities are tied to laws of nature or objective tendencies, and so if an objective probability attaches to some contingent fact, then that situation can be given an explanation in terms of laws of nature or objective tendencies. Hence, if the PSR is false of some contingent fact, no objective probability attaches to the fact.

Thus, we cannot even say that violations of the PSR are improbable if the PSR is false. Consequently, someone who does not affirm the PSR cannot say that Koons’ skeptical scenario is objectively improbable. It may be taken to follow from this that if the PSR were false or maybe even not known *a priori*, we would not know any empirical truths. But we do know empirical truths. Hence, the PSR is true, and maybe even known *a priori*.

2.2.3. Evolution

One of my graduate students suggested in discussion that if one rejects the PSR, our knowledge of evolution may be undercut. We can use this insight to generate an *ad hominem* argument for the PSR. Most atheists and agnostics (and many theists as well, but it is to atheists and agnostics that the argument is addressed) believe that there is a complete naturalistic evolutionary explanation of the development of the human species from a

single-celled organism. I claim that they are not justified in believing this if they do not accept the PSR.

For consider what could be the argument for thinking that there is such an explanation. We might first try an inductive argument. Some features of some organisms can be given naturalistic evolutionary explanations. Therefore, all features of all organisms can be given naturalistic evolutionary explanations. But this argument is as bad as inductive arguments can come. The error in the argument is that we are reasoning from a biased sample, namely those features for which we already have found an explanation. Such features are only a small portion of the features of organisms in nature – as is always the case in science, what we do not know far exceeds what we know.

Once we admit the selection bias, the argument becomes this: “all the features of organisms for which we know the explanation can be explained through naturalistic evolutionary means, and so all the features of organisms can be explained through naturalistic evolutionary means.” There are at least two things wrong with this argument. The first is that it might just be that naturalistic explanations are easier to find than nonnaturalistic ones; hence, it is no surprise that we first found those explanations that are naturalistic. But even if one could get around this objection, it would not obviate the need for the PSR. For the argument, at most, gives us reason to accept the claim that those features *that have explanations* have naturalistic evolutionary explanations. The inductive data is that all the explanations of biological features that we have found are naturalistic and evolutionary. The only conclusion that can be drawn without the PSR is that all the explanations of biological features *that there are* are naturalistic and evolutionary, not that all biological features have naturalistic evolutionary explanations.

A different approach would be to suppose that natural occurrences have naturalistic explanations, and evolution is the only naturalistic form of explanation of biological features that we know of; therefore, it is likely that the development of the human race has a naturalistic evolutionary explanation. But what plausibility is there in the claim that natural occurrences have naturalistic explanations if one does not accept the PSR for contingent propositions? After all, if it is possible for contingent propositions to simply fail to have an explanation, what reason do we have for confidence that, at least, those contingent propositions that report natural occurrences have explanations? If “natural occurrence” is taken as *entailing* the existence of a naturalistic explanation, the argument for an evolutionary explanation of the development of the human race begs the question in its assumption that the development was a natural occurrence. But if “natural occurrence” is taken more weakly as a physical event or process, whether or not it has a natural explanation, then the naturalness of the occurrence does not give us reason to think that the occurrence has an explanation, much less a naturalistic one, absent the PSR. If we had the PSR in play, we could at least try to use a principle, perhaps defeasible, that the cause is ontologically like the effect, so that if the effect is natural, the cause is likely such as well. (It is interesting that this principle itself could be useful to theists with respect to the Gap Problem – see the perfection axiom in Section 5.4.)

Consider a final way to justify the evolutionary claim. We have good inductive reason to think that everything physical obeys the laws of physics. But everything that is governed by the laws of physics has a naturalistic explanation. Hence, the development of the human race has a naturalistic explanation, and an evolutionary one is the best candidate we have.

The claim that everything that obeys the laws of physics has a naturalistic explanation, however, has not been justified. The claim was more plausible back when we thought that everything could be explained in a Newtonian manner, but even then the claim could be falsified. Consider John Norton's (2003) ball-on-dome example. We have a rigid dome, on the exact top of which there sits a perfectly round ball, and the dome is in a constant downward gravitational field of acceleration g . The dome is rotationally symmetric, and its height as a function of the distance r from its central axis is $h = (2/3g)r^{3/2}$. It turns out to be consistent with Newtonian physics that the ball should either remain still at the top of the dome or start to roll down in any direction whatsoever, in the absence of any external forces. One might wonder how this squares with Newton's second law – how there could be an acceleration without an external force. It turns out, however, that because of the shape of the dome, in the first instant of the ball's movement, its acceleration would be zero, and after that it would have an acceleration given by the gravitational force. The physics would fail to explain the ball's standing still at the top of the dome *or* the ball's moving in one direction or another; it would fail to explain this either deterministically or stochastically. Thus, even Newtonian physics is not sufficient to yield the claim that everything that obeys the laws of physics can be explained in terms of the laws of physics.

And I doubt we do any better with non-Newtonian physics. After all, we do not actually right now know *what* the correct physics is going to be, and in particular we do not know whether the correct physics will make true the claim that everything that obeys the laws of physics can be explained in terms of the laws of physics. Besides, surely it would be an implausible claim that justification for the claim that the human race developed through evolutionary means depends on speculation about what the final physics will be like.

I do not have an argument that there is no other way of arguing for the evolutionary claim absent the PSR. But, intuitively, if one were not confident of something very much like the PSR, it would be hard to be justifiably confident that no biological features of the human species arose *for no reason at all* – say, that an ape walked into a swamp, and out walked a human, with no explanation of why.

2.2.4. Inference to best explanation

Suppose we have a phenomenon and several plausible explanations. We then reasonably assume that the best of these explanations is probably the right one, at least if it is significantly better than the runner-up. How we measure the goodness of an explanation is, of course, controverted: prior probability, simplicity, explanatory power, and so on are all candidates. Or, if we have ruled out all explanations but one, we take the remaining one to be true (White 1979) – this is what the maxim that “when you have eliminated the impossible, whatever remains, *however improbable*, must be the truth” comes down to in Sherlock Holmes's actual practice (Doyle 1890, p. 93; italics in the original).

But suppose we admit, contrary to the PSR, the possibility that the phenomenon has no explanation at all. What reason do we have to suppose that the best or the only explanation is likely to be true? To argue for that explanation, we compared it with its competitors. But the hypothesis that the phenomenon has no explanation at all was not one of these competitors. Indeed, we do not know how to compare this hypothesis with its competitors. The hypothesis that there is no explanation is, in one sense, simpler than any explanatory explanation. On the other hand, it altogether lacks explanatory power. Still, it is unfair to rule it out just because it lacks explanatory power unless one believes in the PSR.

Perhaps the no-explanation hypothesis can be ruled out, not because it is impossible, as the defender of the PSR will say, but because it is simply less probable than its competitors. But does it make any sense to assign a *probability* to the hypothesis that a brick comes to exist *ex nihilo* in midair in front of us for no reason at all, assuming this is possible? We certainly cannot assign a probability grounded in the laws of nature to a brick's coming into existence *ex nihilo*, in the way in which we can to the electron's moving upwards in the Stern–Gerlach experiment, since the brick's entry into existence would arguably not be governed by the laws if it happens “for no reason at all.”

But maybe we can argue that such an arising *ex nihilo* is impossible, since it is contrary to the laws. However, the laws of nature only specify what happens in the absence of external influence. They do not, thus, exclude the possibility of a brick coming into existence by the power of a nonphysical being, say, God. But if the PSR does not hold, intuitively any laws that do not preclude the possibility of a brick coming into existence by the power of a nonphysical being should not exclude the possibility of the brick coming into existence *ex nihilo*. The possibility of a nonphysical being's producing such a brick shows that there is no innate contradiction between the brick's coming into existence and there being such-and-such laws of nature. And it would be odd indeed if the laws of nature entailed that any bricks that come into existence should have causes of some sort or other, *whether natural or not*. Furthermore, if my argument is taken seriously, then we may not have good reason to believe in the laws of nature in the first place (without the PSR, that is) – for the phenomena that we tried to explain in terms of them might just be lacking in explanation.

Suppose, however, that we grant that the laws of nature exist and entail that physical events have causes, natural or not, but continue to balk at the full PSR because we are not sure whether nonphysical facts have to have explanations. Then, at least on probabilistic grounds, we cannot exclude the following explanatory hypothesis, available for any phenomenon *F*: there came into existence, *ex nihilo* and for no reason at all, a nonphysical being whose only basic nonformal property was the disposition to cause *F* as soon as the being is in existence, a property that the being has essentially, and this being came into existence for precisely the amount of time needed for the activation of this disposition. Why did Jones fall asleep? Because a nonphysical being came into existence for no reason at all, a being characterized by an essential *dispositio dormitiva* and by nothing else. No nomic probabilities can be assigned to the hypothesis of such a nonphysical being's coming into existence. (It might be that there is some argument available that only God can create *ex nihilo*, and so such a being cannot create a brick *ex nihilo*. Fine, but at least it should be able to create it out of air.)

One might try to assign nonnomic probabilities to the no-explanation hypothesis and the hypothesis of *ex nihilo* creation by a nonnatural being. But then, the no-explanation hypothesis would be on par with each explanatory explanation. And there would be an infinitude of explanatory hypotheses in terms of nonnatural beings that came into existence *ex nihilo*, for we could suppose that, in addition to the disposition to cause *F*, they do have some other essential property (say, being happy or being beautiful), and they differ in respect of it. Why would we take a “normal” scientific explanation over one of these, then?

It is tempting here to say: “Well, we don't know anything either way about the likelihoods of these weird hypotheses that contradict the PSR. So we should just dismiss them all.” As practical advice for doing our best in finding predictions, this may be fine. But if we are to hope for scientific *knowledge*, that surely will not do. A *complete* inability to estimate the likelihood of an alternate hypothesis is surely a serious problem.

It is easy not to take these odd hypotheses seriously. And that may well be because we do, in fact, have a deep commitment to the PSR and maybe even to a defeasible principle that causes have a resemblance to their effects. If I am right, the PSR is essential to the practice of science, even outside of evolutionary biology.

2.2.5. Why aren't there widespread violations of the PSR all around?

If the PSR were false, we would expect a profusion of events that would not appear to fit into any kind of nomic causal order. After all, for each way that things could go in accordance with the laws of nature, there is an uncountable infinity of ways – of arbitrary cardinality – that things could, for no reason at all, go contrary to the laws of nature. For instance, if we deny the PSR, then for no reason at all, a cloud of photons, \aleph_{9314} in number, could suddenly appear *ex nihilo* just near the moon, heading for San Francisco. (Because the cardinality is so high, some of the photons would have to share the same quantum state; but photons are bosons, so they should be able to do that.) And the number of ways such things could happen seems to have no limit if the PSR fails. Or perhaps, \aleph_{9314} nonnatural beings could come into existence, each of which could then produce one photon.

Our empirical observations suggest that the probability of such events is very low. On the other hand, if we get our probabilities *a priori* from some sort of principle of indifference, supposing all arrangements to be equally likely, the messy PSR-violating arrangements would seem much more probable. How to explain the fact that bricks and photon clouds do not show up in the air for no discernible reason? I suggest that the best explanation is that the PSR holds, and that whatever beings there may be (e.g. God) who are capable of causing bricks and photon clouds to show up in the air for no discernible reason are, in fact, disposed not to do so. We need both parts for the explanation: without the PSR, the possibility of this happening for no reason at all would be impossible to rule out, and without the claim that existing beings are unlikely to cause it, the PSR would be insufficient (this suggests that if the cosmological argument can establish the existence of a First Cause, there is reason to think that the First Cause has a predilection for order, a fact relevant to the Gap Problem).

It may seem that I am caught in a vicious circularity here. I have produced a phenomenon – the lack of weird, apparently causeless, events – and have suggested that its explanation needs to involve the PSR. But am I not invoking the PSR in supposing that there is an explanation here? No. I am only invoking inference to best, or only, explanation, an ampliative principle that we should all accept. Nor am I applying this principle to some strange fact such as the conjunction of all contingent states of affairs. I am applying the principle to the homely fact that bricks and photon clouds do not show up in the air *ex nihilo*. And the best explanation of this fact is that they, simply, cannot do that, absent some cause, and that there does not, in fact, exist a cause likely to produce such effects.

One might think that some physical law, say, a conservation law, would do the explanatory work here, a principle other than the PSR. But the logical possibility of miracles shows that it should be possible for a supernatural being to cause photon clouds to show up *ex nihilo*, and if the PSR is false, such supernatural beings could be coming into existence all the time, causing the weird effects. Our best explanation for why *this* is not happening is that there is nothing in existence that would be likely to cause such supernatural beings to come into existence, and by the PSR they cannot come into existence uncaused.

2.2.6. An argument from the nature of modality

2.2.6.1. Alethic modality

Alethic modality is a deeply puzzling phenomenon. Whence comes the difference between a golden mountain and a square circle? Why is it necessary that $2 + 2 = 4$, but merely contingent that horses exist? I could become a biologist, but I could never be a number or a point in space. What makes that so?

The question here is as to the ground of truth of these kinds of facts. I am not asking the explanatory question of *why* these facts obtain. That is easy to find in at least some cases. A square circle is contradictory, for instance, and had evolution gone somewhat differently, the niche occupied by horses would have been occupied by medium-sized and fast reptiles. But what features of reality make these alethic modal facts hold?

Five main kinds of nonrevisionist theories have been offered here: narrowly logical, Lewisian, Platonic, Aristotelian-essentialist, and Aristotelian-causal. The first three will be seen to be unsatisfactory, and only the Aristotelian theories will remain. Of these, the Aristotelian-essentialist account will have some serious problems with it and, moreover, seems to require theism, so the agnostic or atheist cannot embrace it as an alternative to the Aristotelian-causal one. The remaining theory, the Aristotelian-causal one, turns out to entail a PSR sufficiently strong to run a cosmological argument, given some plausible auxiliary assumptions. Hence, we should accept the PSR, unless we have a better account of alethic modality.

I shall now argue for the unsatisfactoriness of the first four theories. I have no argument that there is no better story possible than the Aristotelian-causal one. But until a good competitor is found, we should accept this account, and hence the PSR.

2.2.6.2. Narrowly logical account of modality

In a number of other early modern thinkers, we have the following “narrowly logical” account of modality, probably best developed in Leibniz. A proposition p is necessary if and only if a contradiction can be proved from its negation. Assuming classical logic, as these thinkers did, it follows that *necessity* is equivalent to *provability*. And a proposition is possible if and only if no contradiction can be proved from it.

There are counterexamples to this account.

First, we learn from Gödel that for any axiomatization within our reach (any set of axioms we can generate recursively), there will be truths of arithmetic that we cannot prove from the axiomatization. On the narrowly logical account, thus, there are contingent truths of arithmetic. This seems absurd. (For one, what kind of truthmakers would they have?)

Second, necessarily, all horses are mammals. But this is an empirical discovery. We cannot *prove* it by narrowly logical means. *A posteriori* necessities such as this provide a large family of counterexamples.

Third, it is impossible for anything to cause itself. (If, like Descartes, you disagree, choose another example – maybe, the claim that it is necessarily possible for something to cause itself.) But how would we go about proving this? We might start with some partial analysis of causation. Perhaps a cause has to temporally precede the effect (a dubious thesis in my opinion, but what I say will apply to any story we could fill in here). And nothing can temporally precede itself. But how could we *prove* that a cause has to temporally precede the effect, and how do we prove that nothing can temporally precede itself?

In two ways, I suppose. First, we might derive these claims from some *definitions*, say of causation or temporal priority. But, leaving aside the somewhat implausible suggestion that “causation” and “temporal priority” can *both* be defined, how do we *prove* that this definition is in fact the right way to define the terms? To show that a definition is correct is beyond the powers of logic narrowly conceived, unless the definitions are stipulative, in which case the proof is trivial. But a stipulative route is unsatisfactory for two reasons. First, the claim that nothing can cause itself is *not* just a claim involving a stipulative concept of “cause.” Second, even if I have a stipulative definition, I need the principle that if *D* is stipulatively defined as *E* (where *E* is some linguistic expression), then necessarily anything that satisfies *D* satisfies *E*. But what grounds the latter necessity? If I say that I can prove it from the definition of “stipulated,” then I go around in a circle – for either the definition of “stipulative” is nonstipulative, in which case it seems we need to go beyond logic narrowly conceived to prove the definition of “stipulative” correct, or else we have a stipulative definition of “stipulative,” and to prove that anything that satisfies *D* must satisfy *E* whenever *E* is the stipulative definition of *D*, I need to know that, necessarily, whatever is stipulative has the properties in terms of which the word has been defined.

So the stipulative route to proving that nothing can cause itself will not work. The only other route is that among our *axioms* there are substantive axioms about the nature of causation or that there are substantive *rules of inference* in our logic. Without such axioms or rules of inference, we get nowhere when dealing with a nonstipulative concept. But now note that any axiom gets to be necessary *for free* on the narrowly logical account. So what would it be that would make it be the case that among our axioms is the claim that, say, *causes temporally precede their effects*, or whatever other truth it would be from which we were going to prove that nothing can cause itself, while the equally true claim that *there are horses* is not among the axioms? The intuitive answer is that the claim about causation is more plausibly a necessary truth, while the claim about horses is plainly contingent; but that would be viciously circular. Similarly, if there are substantive rules of inference in our logic, say, ones that allow us to infer from *x causes y* and *y causes z* that *x is not identical with z*, the question of what makes these but not other substantive rules of inference (say, the rule that one can derive *there are horses* from every statement) appropriate is equally problematic as the question of what gets to count as an axiom.

And so the narrowly logical account is of little help – a part of what makes a proposition an axiom seems to be that it is necessary, and a part of what makes a proposition be a rule of inference is that it embodies a necessary implication. Moreover, the necessity here is the same sort of necessity we were trying to explicate, so there is really very little gain. Alethic modality remains ungrounded.

Our last example has shown the general problem with narrowly logical accounts of modality: the grounding burden simply shifts to the question of the choice of the axioms and/or rules of inference and that question we cannot answer with the resources of the view in question.

An early modern answer one might try is this: we take as axioms all and only the claims that are clear and distinct. An anachronistic objection is that this does not solve the Gödelian problem. A counterexample-based answer is that the claim that I exist seems to be as clear and distinct as anything can be, and yet is contingent. Moreover, plausibly, there are necessary truths that are far beyond our ken and cannot be derived from clear and distinct truths within our ken. (If we assume the existence of God, this is very plausible: there surely are many such facts about him.) Besides, we no longer have much of a handle on the notion

of clear and distinct claims, and to use them to ground necessity would be to confuse facts about our doxastic faculties with metaphysics.

The narrowly logical view is distinctly unsatisfactory. Let us thus continue our brief survey.

2.2.6.3. *Lewisian account of modality*

The Lewisian account, also known as Extreme Modal Realism (EMR), says that a proposition is possible if and only if it holds in some possible world, and necessary if and only if it holds in all possible worlds. This is only going to be of help if we have an independent account of possible worlds, and indeed EMR supplies one. A possible world is a maximal spatiotemporally interconnected aggregate of things. (We can also stipulate that abstract entities count as existing in every world.) We live in one of these worlds, the *actual* world, and there are infinitely many others. Every way that things could have been is a way that things are in some world. We then make a distinction between *existence* and *actuality*. Something *exists* provided it exists in some world or other. Something is *actual* provided it exists in the actual world.

EMR has a number of problematic consequences. For instance, if EMR holds, consequentialistic moral reasoning breaks down completely because no matter what I do, the overall consequences in reality are the same, since reality always already contains all possible worlds. Lewis thinks that we can restrict our concern to those who exist in our world and only count what happens to them as relevant. But this neglects the importance of overall consequences. Even deontologists need consequentialistic moral reasoning. If I am to give money to one of two charities, and everything is otherwise morally on par, I should choose the one giving to which will produce better consequences.

Lewis, however, thinks that what matters ethically is not just the consequences but that *I have produced them* (Lewis 1986, p. 127). I cannot affect what happens in other worlds, but I can be the cause of goods in our world. Of course, this makes no difference in the space of all possible worlds – in infinitely many of them, people very much like me are causes of goods and in infinitely many of them, people very much like me are not causes of goods, and the distribution of worlds is not affected by my action. But my relationship to the goods *is* affected.

However, this unacceptably reduces the moral weight of consequences. Suppose that either you or I can operate on a patient. The operation is perfectly safe, but I am better than you at this particular operation, and so the patient will recover somewhat faster after the surgery if I do it. I thus have good reason, when we are deciding which of us will perform the operation, to volunteer to do it. And if I do perform the operation, then I additionally gain the agent-centered good of *my* being the cause of the patient's improvement. However, the latter consideration is surely of very little moral weight. After all, the same kind of consideration would also give you reason to do the surgery, but this consideration should be trumped by the good of the patient. Even if my skill at this operation is only *slightly* better than yours, so that the patient will likely recover *slightly* better, all other things being equal this fact should trump your reason to be the cause of the patient's improvement. Thus, the agent-centered reason of wanting to be the cause of good is, in a case like this, of very low weight – the consequences are the *main* consideration.

This is not so in every case. When there is a close relationship between me and someone else, then it may matter very much that I be the one to benefit that person. However, when

there is no particularly morally important relationship – and merely being spatiotemporally connected is very low on the scale of moral importance – it should not matter or at least matter much.

On Lewis's view, however, my reason to help strangers is *only* the agent-centered reason to be the cause of goods because the consequences are always the same. But since the agent-centered reason to be the cause of goods has extremely low weight, it follows that EMR radically lowers the weight of reasons to help strangers. If we accept a more traditional assessment of the weight of these reasons, we shall have to reject EMR.

Instead of cataloging further problems entailed by EMR, I shall give what I take to be one of the deepest criticisms, which I believe is due to van Inwagen. The criticism is simply that the existence of infinitely many maximally spatiotemporally interconnected aggregates has nothing to do with modality. If we found out that reality contains infinitely many maximally spatiotemporally interconnected aggregates, we would simply have learned that the *actual* world is richer than we thought – that it contains all of these island universes – rather than learning something about the space of possibilities.

Here is a variant on the objection. Suppose that there exist infinitely many maximally spatiotemporally interconnected aggregates, and some of them contain golden mountains but none contains unicorns.³ It would follow that golden mountains are possible, simply because what is actual is also possible, but surely it would not follow from this fact that unicorns are impossible. And if there were only *one* spatiotemporally interconnected aggregate, namely ours, it would not follow that modal fatalism is true – that every actual truth is necessary. Yet on Lewis's view, if no unicorns were found in any island universe, it would follow that unicorns are impossible, and if there were only one island universe, it would follow that every actual truth is necessary since things could not be otherwise than they are then.

Now Lewis, of course, thought there was more than one universe, and indeed that there was a universe that contained unicorns. He believed this because he accepted a recombination principle that said that one can cut up the ingredients of one world and rearrange them in any geometrically available way, and the resulting rearrangement would be exemplified in some world or other. However, while he accepted the recombination principle, the recombination principle is not, on his view, a part of what makes alethic modal claims true. What makes alethic modal claims true on his view are just the facts about universes, and we have seen that that is not correct.

We should thus reject EMR and keep on searching for a good account of modality.

2.2.6.4. *Platonic account of modality*

The most promising contemporary realist alternative to Lewis's account of possible worlds are the abstract worlds accounts promoted by Robert M. Adams (1974) and Alvin Plantinga (1974). On their accounts, worlds turn out to be abstract Platonic entities, exactly one of which is instantiated by the universe, where “the universe” is defined to be the aggregate of all existing or occurring concrete entities, and this is the world that is absolutely actual. I will focus primarily on the Adams permutation of this account.

3. To avoid Kripkean worries as to what precise species a unicorn would belong to, we can stipulatively define a unicorn as any horselike mammal with one horn.

We thus start off by introducing *propositions* as theoretical abstract entities that are the bearers of truth-values and are needed to explain what it is that sentences express, what the objects of beliefs and propositional attitudes are, and what paraphrases preserve, somewhat as electrons are needed to explain various physical phenomena. Some propositions, namely the true ones, are related to things and events in the universe, with the relation being one of the propositions *being made true by* or *representing* these things and events in the universe. If things in the universe were otherwise than they are, then different propositions would stand in these relations to things in the universe – if there were unicorns, then the proposition that there are unicorns would stand in the relation of *being made true by* to some things, namely, the unicorns in the universe.⁴

Note that the theoretical reason for believing in these Platonic propositions is largely independent of issues of modality. Adams then constructs a possible world as a maximal consistent collection of propositions. (An argument is needed that such collections exist, but let that pass.) Exactly one world is then absolutely actual: it is the one all of whose propositions are true. A proposition can be said to be true *at* a world, providing it is one of the propositions that are members of the collection of propositions that the world is identical with. Note that because the worlds are Platonic entities, I had to distinguish between the concrete *universe*, which we physically inhabit, and the actual *world*, which is the collection of all true propositions.

One might object to the Platonic approaches on the grounds that they all involve queer entities. Not only are we required to believe in Platonic beings, but, as Lewis notes, we are to believe that there is a magical relation of representation holding between Platonic beings such as propositions and the concrete entities that make them true, with it being contingent which propositions enter into those relations since it is contingent which propositions are true. What is it, then, that picks out one relation in the Platonic heaven rather than another as *the* relation of representation?

The proponents of these Platonic worlds can argue, however, that they have no need to answer this question. The relation of representation is one of the primitive terms in their theory, and it is not a primitive chosen *ad hoc* to explain possible worlds but a primitive needed for other explanatory purposes, such as for making sense of our practices of claiming, believing, and paraphrasing. Nonetheless, if we had some way of pointing out this relation within the Platonic universe of all relations, we would be happier as theorists.

These Platonic theories are expressly nonreductive as accounts of possibility, unlike Lewis's theory. For Adams, a possible world is a maximal consistent collection of propositions, which is just the same as saying it is a maximal *compossible* collection of propositions. On this theory, there is a primitive abstract property of possibility or consistency that applies to individual propositions and to collections of them. One could also take necessity to be the primitive concept, but this would not change anything substantially.

That the Platonic accounts are nonreductive is only a problem if a reductive account of possibility is available. However, the most plausible account claiming to be reductive is Lewis's, which is too paradoxical to accept. But while a complete reduction is probably impossible, it could be desirable to give at least a partial reduction, on which the whole realm of

4. Lewis (1986) worries that the relation between the propositions and the things they are about is magical, but as van Inwagen (1986) notes, it is no more magical (although no less) than the relation between sets and their members, a relation that Lewis accepts.

alethic possibility would be seen to have its root in some more comprehensible subclass. An example of an otherwise implausible theory that would provide such a reduction would be an account on which a proposition is possible if and only if Alvin Plantinga could conceive its being true: all of modality would then be reduced to Alvin Plantinga's considerable powers of imagination. Claims about Plantinga's powers are still modal claims, but of a more comprehensible sort than claims about the possibilities of unicorns and zombies. However, these Platonic accounts do not succeed in performing this more limited reduction either.

Adams's theory is an *actualist* one. His possible worlds are built up out of things that are actual. These abstracta actually exist – indeed, necessarily so – and an actualist theory is one that grounds possibility in actually existent realities. On the other hand, Lewis's other worlds are not actual entities by Lewis's indexical criterion, as they are not the world in which my tokening of the word “actual” in this sentence occurred. If we think of possible worlds as possibilities for our universe, then there is a sense in which Adams and Plantinga have grounded possibilities in actuality, thereby answering to the Aristotelian maxim that actuality is prior to possibility.

However, in a deeper way, the Platonic approach is not faithful to what the Aristotelian maxim affirms. When Aristotelians say that a possibility is grounded in an actuality, they mean that actuality includes some powers, capacities, or dispositions capable of producing that possibility, which of course once produced would no longer be a mere possibility. This is clearest in the paradigm case where the actuality is temporally prior to the possibility. Aristotle's favorite illustration is how the actuality of one man makes possible the existence of a future man through the first man's capability for begetting a descendant. If we find attractive the idea that possibilities should be grounded in actuality in the stronger Aristotelian sense, then the Platonic approach will be unsatisfactory because Platonic entities, in virtue of their abstractness, are usually taken to be categorially barred from entering into causal relations, and hence cannot make possibilities possible by being capable of producing them. And if they make possibilities possible by being capable of producing them, then what we have is a variant on the Aristotelian-causal account.

Moreover, an Aristotelian can argue that in fact there *are* capabilities and dispositions sufficient to ground the truth of at least *some* possibility claims. That I could have been a biologist is very plausibly made true by my capacities and dispositions and those of various persons and things in my environment. These capacities and dispositions are concrete real-worldly things, albeit ones having modal force. Hence, in fact, we do not need a Platonic realm to make at least some possibility claims true. Indeed, the facts about the Platonic realm – about propositions' having or not having some primitive property – are interlopers here. Just as the statement that I could have been a biologist was not made true by what my Lewisian counterparts in other worlds do, so too it is not made true by abstract properties of Platonic abstracta. The common intuition behind both cases is that it is something in me and my concrete environment that makes the statement true.

This, however, creates a major problem for the Platonic approach. On the Platonic approach, what makes it possible that I have been a biologist is that the abstract proposition (an entity in the Platonic heaven) *that I have been a biologist* has the abstract property of possibility. But we have just seen that there are concrete capacities and dispositions in the universe that are by themselves sufficient to make it possible that I have been a biologist. We thus have two different ways of characterizing possibility: one is via the concrete this-worldly Aristotelian properties of concreta, which really do exist – the Platonist should not deny this – and the other is via the abstract Platonic primitive properties of abstracta.

Moreover, anything that is possible on Aristotelian grounds will be physically possible, and hence also logically possible, and thus possible on Platonist grounds (though *prima facie* perhaps not conversely). But now we can ask: Why is this so? Why is there this apparent coincidence that anything made possible by this-worldly powers and capacities and dispositions happens to correspond to a proposition in the Platonic realm that has a certain abstract property? The Platonist is unable to explain this coincidence between powers in our universe and abstract facts about the Platonic realm, given the lack of causal interaction between the two realms.

2.2.6.5. Aristotelian-essentialist account of modality

Aristotle's own account of modality seems to have been based on the idea that a sentence is necessarily true if and only if it holds always. Then, a sentence is possibly true if it holds at some time. I shall not consider this account further. It is not clear that in characterizing "necessarily" in this way, one is really talking of the same thing as we are when we say that *necessarily* there are no square circles. We certainly mean more by saying that there can be no square circle than that just that there have never been, nor are, nor ever will be any square circles. Granted, if we adopt some kind of principle of variety, on which given infinite time every possibility is realized, we might get out of this Aristotelian story an account that is extensionally acceptable. However, that account would still face many of the same problems Lewis's account faces – indeed, it would be just like Lewis's account, but with time-slices replacing universes. In particular, the objection that we are not talking about modality at all would be to the point. If it should turn out that the past, present, and future of our world contain no golden mountains, that would say nothing about whether golden mountains are possible.

But while Aristotle's own account of modality was flawed, two somewhat different accounts have been derived from ingredients of Aristotelian ontology. One of these grounds modality in the *essences* of things and takes necessity to be the primitive notion. The other account grounds modality in *causal powers* and takes *possibility* to be more primitive. I shall begin by discussing the account based on essences (cf. O'Connor 2008).

Things that exist have essences. These essences, on this account, constrain what properties these things can have. Thus, a horse cannot be immaterial, and a dog cannot become a cat. A proposition is impossible provided that it affirms something contrary to the essences of things.

There are several objections to this rough sketch of a view. First, maybe it is plausible that the essence of a horse encodes that a horse must occupy space. But what makes it necessary that horses must occupy space or be green? Do we really want to suppose that for every property *P*, the essence of a horse contains in itself the specification that a horse occupies space or has *P*? An affirmative answer appears implausible. Why should the essence of a horse include the specification that horses occupy space or are cats?

This objection is not just an incredulous stare. Horses could surely exist without any cats in existence. But the essence of a horse, on this view, in some way presupposes catness. It follows that it makes sense to talk of catness – the essence of cats – apart from cats, since horses could exist apart from cats, and hence the essence of a horse could exist apart from cats. The Aristotelian, however, cannot tolerate this, *unless* the Aristotelian is a theistic Aristotelian who accepts that all essences have some kind of an existence in the mind of God. Thus, unless one accepts theism, the theory seems to be unsatisfactory.

But maybe I was too fast. Perhaps it is not that the essence of a horse contains all the necessary truths about horses, but that all the necessary truths about horses can be *derived* from the essence of a horse as combined with all other essences there are. That every horse occupies space or is a cat can be derived from the essence of a horse and the essence of a cat.

But “derived” surely means “logically derived.” And so it turns out that the Aristotelian-essentialist needs elements of the narrowly logical view. Once again, the same question comes up: what grounds the choice of axioms or rules of inference? However, the Aristotelian is better off here than the proponent of just the narrowly logical view because the truths contained in the essences of things provide a rich set of nonarbitrary axioms.

Aristotelian-essentialists might then be able just to specify, say, some plausible version of logic (e.g. some second-order quantified modal logic), and claim that our thought and language presupposes the truth of this logic. They could then say one of two things about the status of this logic. First, they could say that the basic rules of this logic are grounded in some or all essences. For instance, maybe every essence encodes the rules of logic, or maybe one could make the theistic move of saying that the essence of God encodes these rules. In this way, the rules of logic would be on par with other truths within the essences of things, such as the truth that horses occupy space that is encoded within the essence of a horse. This construal of the rules of logic would be to make the rules of inference effectively into facts or propositions written into essences, such as:

- (6) For all p and q , if it is the case that if p then q , and if it is the case that p , then it is the case that q .

But the rules of logic cannot be construed in this way without losing what is essential to them, namely their *applicability*. If *modus ponens* is just the fact (6) or maybe the necessary truth of (6), then how do you apply *modus ponens*? You have p , you have *if p then q* , and then you know that the antecedent of the big conditional in (6) is satisfied. But how do you know that the consequent of the big conditional in (6) holds, namely that it is the case that q ? You know it by *modus ponens*. But *modus ponens* is just the truth (6), so you need to go back once more to (6) to apply it to the case where you have (6) and the antecedent of (6). In other words, you need *modus ponens* to apply *modus ponens* if *modus ponens* is just a truth like (6), and a vicious regress ensues. Applicability requires that the truths of logic be more than just statements.⁵

A better solution for advocates of the Aristotelian-essentialist account of modality would be to say that logic narrowly construed is something deeper than the necessities they are grounded in essences. One could, for instance, take the Tractarian line that narrowly logical impossibilities cannot even be thought.

But we have not exhausted all the objections to the Aristotelian-essentialist view. Consider truths that hold of all things no matter what essence they might have. No entity has a shape that is both a square and a circle (at the same time and in the same respect), and no entity is the cause of itself. What makes *these* be necessary truths? Granted, it may be encoded in the essence of every actually existing thing that nothing having that essence is a square circle, or is *causa sui*, or exists in a world where some (actually true) Gödelian

5. This argument goes back, at least, to Sextus Empiricus (1993, sec. II.11, para. 114).

unprovable arithmetical claim fails to hold, but a seemingly stronger claim is true: there could not be *anything*, whether with one of these essences or with some other essence, that is a square circle or that is *causa sui* or that exists in a world where some particular (actually true) Gödelian unprovable arithmetical claim fails to hold. Maybe the square circle case can be handled through a narrowly logical move as described earlier, but it may not be plausible that this can be done with the *causa sui* case, although perhaps there is some Tractarian line that one can take that self-causation cannot even be thought. But in any case, the Tractarian line does not seem to help much with the Gödelian worry.

Moreover, consider the question of what essences can possibly exist (in mind or reality). The story we have so far is that something is possible provided its existing is not contradictory to the truths encoded in those essences that exist. This, however, seems to let in so many essences that a certain amount of skepticism is engendered. For instance, it seems that there will be a possible world w that is just like this one, with this exception. The essence of human beings does not exist at w , but instead there are entities that physically behave just like human beings, except that instead of being a single natural kind, they are divided up into two natural kinds defined by different essences: there are those who have an even number of hairs on their bodies and there are those who have an odd number of hairs on their bodies. As soon as one of these beings gains or loses a hair, it perishes, and a new being comes to exist, physically and psychologically just like it, apart from that hair. Otherwise, everything is as it is in our world. After all the existence of such being and such essences does not seem to contradict the truths encoded in any of the essences that exist, such as the essence of the live oak or the photon. But once we allow that w is possible, do we have good reason to suppose that it is not *our* world, that in our world there are no different essences for people with even numbers of hairs and for people with odd numbers of hairs?

The problem, thus, is with what constrains what essences there could be. One answer, inspired by the static character of Aristotle's universe, would be that all the essences that can exist in fact do exist, or at least existed, exist, or will exist. However, a crucial difficulty remains as to what "can" could mean here. What constrains which essences *can* exist?

Some of these problems can be solved by going a theistic route. Perhaps there is a God whose essence encodes necessary truths not just about himself but about others, such as that there can be no square circles, and that certain weird essences cannot exist.

In fact, I think one can argue that only a necessarily exemplified essence can solve the difficulties here; for, on the present account, it seems very likely that an essence cannot in any way constrain what happens in any worlds in which that essence is not exemplified. An essence E can exclude some worlds containing an exemplification of it from including something incompatible with E , but it does not have anything to say about what things are like in worlds where there is no exemplification of E .

Suppose now that none of the essences that are exemplified in our world is necessarily exemplified. We should then be able to describe a world full of really, really weird things – beings with essences that make their kinds be defined by the number of hairs, self-caused beings, and the like – as long as we do not transgress narrowly logical norms and as long as we take care to include none of the beings of our world. And such a world will be possible since the essences that exist in our world will be irrelevant to what goes on in that world as our world's essences will be unexemplified there. Likewise, a completely empty world would be possible then – a world with no essences exemplified. In *that world*, it will

be true that everything that is narrowly logically possible is metaphysically possible, since there will be no constraining essences at all. In particular, in that world it will be possible that Gödelian claims of arithmetic that are true at our world are false. And, of course, it would then be the case that S5 is false, but the Aristotelian-essentialist may not mind that consequence.

If we think that the space of all possible worlds is not such a slum as to include all such worlds, we have to think that at least one of the beings that exist in our world is such that its essence is *necessarily* exemplified, and that the essences of the necessary beings place constraints on what sorts of essences there can be, what sorts of arithmetical truths there can be, and so on.

There are now two difficulties. First, what does it mean that the essence of some being is *necessarily* exemplified? If an essence *E* cannot constrain what happens in worlds where it does not exist, it is unclear how *E* could prevent the actuality of worlds that do not contain an exemplification of *E*. Second, just *how* does an essence place such global constraints on worlds and on what essences are exemplified in them?

The first difficulty forces us, I think, to modify the account. Let *N* be one of the necessarily exemplified essences. Even if some necessities are grounded in essences, the necessity of *N*s being exemplified cannot be grounded in an essence, at least not in the sense in which essences exclude their being exemplified together with something incompatible, since by doing so, the essences do not exclude their *not* being exemplified. So there is some other kind of necessity that the exemplification of *N* has. This is, in general, not going to be narrowly logical necessity, since unprovable arithmetical truths will follow from the exemplification of all the necessarily exemplified essences.

The account now becomes rather less attractive. It posits *three* kinds of modality as together yielding metaphysical alethic modality: the necessity of the exemplification of certain essences, the necessities encoded in essences, and narrowly logical necessity. Moreover, our best story as to what a necessarily exemplified essence that constrains reality outside of itself is like is that it is the essence of God, so this is not an escape an atheist is likely to want to take. And we have no story yet about what necessary exemplification is grounded in.

There is a way of making something similar to this story work. If we posit that all contingently exemplified essences must originate from something, then we might get the idea of an essence that does not itself originate from anywhere, an essence that is necessarily exemplified, so that the contingently exemplified essences get their reality from at least one necessarily exemplified essence or from the exemplifier of such an essence (in the case of God, if divine simplicity holds, the two options will come to the same thing). It will also be plausible that just as the essences originate from something, so do their exemplifications; on an Aristotelian view, essences are not completely independent of their exemplifications. All of this focuses the attention, however, on causation, and leads us to the last account of modality – the causal one.

Another thing that leads us away from the Aristotelian-essentialist account of modality is the intuition that I used against the Platonic view. One can give a simple account of why I could be a biologist in terms of my abilities and the powers of various entities in my environment. On the Platonic side, I wondered why there is this coincidence between what happens in the Platonic realm and earthly powers and capacities. Now one can wonder why there is a coincidence between powers and essences. Why is it that I cannot do anything

that contradicts the essence of any entity in existence? Perhaps this question is somewhat less pressing than on the Platonic side. After all, maybe my powers are grounded in my essence. But it is still not clear why something could not have the power to act contrary to its essence.⁶

2.2.6.6. *Aristotelian-causal account of modality*

The critiques of the Platonic and Aristotelian-essentialist accounts point the way toward an account where causation is central. Here is a sketch of an account that does this. Say that a nonactual state of affairs *S* is merely possible provided that something – an event or substance or collection of events or substances, say – exists (in the tenseless sense: existed, exists presently, exists eternally, or will exist) with a causal power of bringing about *S*, or with a causal power of bringing about something with a causal power of bringing about *S*, or with a causal power of bringing about something with a causal power of bringing about something with a causal power of bringing about *S*, or more generally provided that something exists capable of originating a chain of exercises of causal power capable of leading to *S*. We then say that a state of affairs is possible if it is either actual or merely possible, and that it is necessary when its nonoccurrence is impossible. A proposition, then, is possible provided it describes a possible state of affairs, and necessary if it describes a necessary state of affairs.

This account has the advantage of reducing metaphysical possibility to causal possibility. One might think this is not much of a gain – we are still stuck with some primitive modality. Yes, but the primitive modality we are left with is a modality that we have a better handle on and a better epistemological story about. We ourselves exercise causal powers all day long and run up against the causal powers of other entities. Our scientific observation of the world gives us information as to what is and what is not within the powers of things. For instance, we know that unicorns⁷ are possible because we know that it would be within the powers of natural selection and variation processes to have produced unicorns.

Moreover, we are probably going to need causal powers, or something like them, in our metaphysics even if we have an independent story about metaphysical alethic modality. It does, after all, seem to be a feature of the world that entities can produce effects. So by reducing metaphysical to causal modality, we seem to make a real gain in elegance and simplicity.

Furthermore, this account lets us handle a spectrum of modalities in a uniform framework by restricting the entities in the causal chains that define mere possibility and the

6. It is also worth noting, by the way, that essences function differently in the Aristotelian-essentialist account than they did in medieval Aristotelian views, and perhaps even in Aristotle himself, although this is not an objection to the Aristotelian-essentialist account. The essences that medieval Aristotelians have talked about were not understood as having the modal implications that the Aristotelian-essentialist accounts needs them to have. For instance, the Christian West took it for granted that it was possible for the second person of the Trinity to take on a human nature. But a human nature is an essence. So in one case at least – that of the second person of the Trinity – the possession of the essence of humanity was not an “essential property” in the modern, Kripkean sense of a property that the entity could not possibly lack, and it is this modern sense that would be needed to get the Aristotelian-essentialist account going.

7. See n. 3 above.

causal relations between them. For instance, a nonactual state of affairs is physically causally merely possible provided that it can be produced by a causal chain consisting purely of physical entities and starting with something physical. A state of affairs is temporally merely possible provided that it is not actual but can be produced by a chain of exercises of causal power starting with something in the present or future.

But what is of most relevance to this chapter is that, given some plausible assumptions, the Aristotelian-causal account, perhaps surprisingly, entails a version of the PSR: every contingent state of affairs has a causal explanation, that is, an explanation based on facts about contingent exercises of causal powers, perhaps combined with some necessary truths.

For the argument, I need a *prima facie* weaker version of the Brouwer Axiom. The Brouwer Axiom, in general, states that if p holds, then it is a necessary truth that p is possible. The weaker version of it that I need is:

- (7) If p holds contingently, then it is possible for p to be both possible and false.

This follows from the full Brouwer Axiom, since if p holds contingently, then p is possible, and so it is necessarily possible, but since it is contingent, it is possibly false, so possibly it is both false and possible. And Brouwer, in turn, follows from S5.

Suppose for a *reductio* that a contingent state of affairs E has no causal explanation. Let E^* be the state of affairs of E 's obtaining without causal explanation. Then E^* is a contingent state of affairs. By the weaker version of the Brouwer Axiom, it is possible that E^* does not obtain but is nonetheless possible. Let us suppose a possible world w where that happens. Here, the use of possible worlds is inessential, but it helps make the argument clear. In w , E^* does not obtain but is possible. Thus, there is a cause C in w that could initiate a chain of exercises of causal powers capable of leading to E^* 's obtaining. But that is absurd, since in doing so, the chain would give a causal explanation of E as well as leading to E 's not having a causal explanation!

One might deny Brouwer, as well as Brouwer's weaker cousin (7), and hold on to the Aristotelian-causal account in the absence of the PSR.⁸ But the Brouwer Axiom is intuitively plausible: however else things might have gone than they did, it would still be true that they could have gone as they actually did.

Without the Brouwer Axiom, we can give an alternate argument for the PSR based on the following highly plausible material conditional:

- (8) If the PSR is true in all possible worlds with the possible exception of the actual world, then the PSR is in fact true in all possible worlds.

It would be incredibly bad luck for us to inhabit the one world where the PSR is false, if there were one. Moreover, if (8) is false, the following absurdity ensues: the PSR is false, but had I skipped breakfast this morning, it would have been true (since it is true in all possible worlds in which I skip breakfast this morning, as it is true in all possible worlds but the nonactual one, and in the actual one I had breakfast). And even someone who is

8. In earlier work (e.g. Pruss 2006, sec. 19.5.2), I said that S5, and hence Brouwer, can be proved from the Aristotelian-causal account of possibility. The sketch of an argument that I gave there does not seem to work, however, unless one assumes something like the PSR.

willing to embrace this absurdity should still accept the cosmological argument, since the cosmological argument could be run in the world where I skip breakfast this morning, and it would be an even greater absurdity to suppose that God does not in fact exist but would have existed had I skipped breakfast this morning. And this would in fact be a contradiction, not just an absurdity, if God is a necessary being.

To show the PSR to be true given (8), for a *reductio* suppose that there is a possible world w , distinct from the actual world, but in which the PSR does not hold. Let E be a state of affairs in w that has no causal explanation. If E does not obtain in the actual world, let $F = E$. Otherwise, let F be the conjunction of E with some other state of affairs obtaining in w that does not obtain in the actual world – there must be such, since w is not the actual world, and hence different states of affairs obtain in w than in the actual world. In either case, F is a state of affairs in w that has no causal explanation. Let F^* be the state of affairs of F 's obtaining with no causal explanation. Then F^* is a possible state of affairs but is not actual, since F does not obtain in the actual world. But then there is something that can initiate a chain of causes leading to F^* , which, as in the Brouwer-based argument, is absurd, since the chain of causes will lead to F 's obtaining, as well as to F 's not having a causal explanation.

Thus, the Aristotelian-causal account of modality leads to the PSR, while the main alternatives to this account of modality are unsatisfactory and/or require something like theism anyway. This gives us a powerful reason to accept the PSR.

2.2.7. Philosophical argumentation

It is morally acceptable to redirect a speeding trolley from a track on which there are five people onto a track with only one person. On the other hand, it is not right to shoot one innocent person to save five. What is the morally relevant difference between the two cases? If we denied the PSR, then we could simply say: "Who cares? Both of these moral facts are just brute facts, with no explanation." Why, indeed, suppose that there should be some explanation of the difference in moral evaluation if we accept the denial of the PSR, and hence accept that there can be facts with no explanation at all?

Almost all moral theorists accept the supervenience of the moral on the nonmoral. But without the PSR, would we really have reason to accept that? We could simply suppose brute contingent facts. In this world, torture is wrong. In that world, exactly alike in every other respect, torture is a duty. Why? No reason, just contingent brute fact.

The denial of the PSR, thus, would bring much philosophical argumentation to a standstill.

An interesting thing about this argument is that it yields a PSR not just for contingent truths but also for necessary ones.

2.2.8. Justification via the sense of deity

If God exists, then the PSR for contingent propositions is true. Why? Because God's activity ultimately explains everything. This is going to be clearest on views on which God's activity *alone* explains everything, and that is going to be most plausible on Calvinist-type views but also seems correct on any theological account that has a strong view of divine concurrence with creaturely activity. Moreover, the inference from God's being the creator and sustainer of everything to the claim that divine activity provides the

explanation of everything contingent, or at least of everything contingent that is otherwise unexplained (this variant might be needed to handle creaturely free will), is a highly plausible one. Thus, someone who has good reason to accept theism has good reason to accept the PSR.

Now one might think that this is a useless justification for the PSR if we are going to use the PSR to run a cosmological argument, since then the cosmological argument will be viciously circular: the conclusion will justify the PSR, whereas the PSR is a premise in the argument.

However, recently, Daniel Johnson (forthcoming) has come up with a very clever account showing that a cosmological argument based on the PSR could still be epistemically useful, even if the PSR is accepted because of the existence of God (he also applies the view to the possibility premise in the ontological argument). Suppose that, as Calvin and Plantinga think, there is a *sensus divinitatis* (SD), which, noninferentially, induces in people the knowledge that God exists – at least absent defeaters – and tells them something about God’s power and nature.

Suppose that Smith knows by means of the SD that God exists. From this, Smith concludes that the PSR is true – this conclusion may not involve explicit reasoning, and it is one well within the abilities of the average believer. Smith then *knows* that the PSR is true. Next, Smith sinfully and without epistemic justification suppresses the SD in himself and suppresses the belief that God exists. If Calvin’s reading of Romans 1 is correct, this kind of thing does indeed happen, and it is why nontheists are responsible for their lack of theism. However, the story continues, the suppression is not complete. For instance, Smith’s worshipful attitude toward God turns into an idolatrous attitude toward some part of creation. It may very well happen, likewise, that Smith does not in fact suppress his belief in the PSR, although he forgets that he had accepted the PSR in the first place because he believed in God. Indeed, this situation may be common for all we know.

Johnson then claims that Smith remains justified in believing the PSR, just as we remain justified in believing the Pythagorean theorem even after we have forgotten from whom we have learned it and how it is proved. Thus, Smith continues to know the PSR. The cosmological argument then lets Smith argue to the existence of God from the PSR, and so Smith then can justifiably conclude that God exists. Of course, unless Smith has some additional source of justification for believing the PSR, Smith has no more justification for believing that God exists than he did when he learned about God from his SD. So the argument has not provided *additional* evidence, but it has restored the knowledge that he had lost.

We have a circularity, then, but not one that vitiates the epistemic usefulness of the argument. Irrational suppression of a part of one’s network of belief can be incomplete, leaving in place sufficient beliefs allowing the reconstruction of the suppressed belief. A similar thing happens not uncommonly with memory. Suppose I am trying to commit to memory my hotel room number of 314. I note to myself that my hotel room number is the first three digits of π . Later I will forget the hotel room number, but remember that it is identical to the first three digits of π , from which I will be able to conclude that the number is 314. My reason for believing the number to be identical to the first three digits of π was that the number is 314, but then, after I lose, through a nonrational process of forgetting, the knowledge that the number was 314, I will be able to recover the knowledge by using a logical consequence of that very piece of knowledge. In doing so, I do not end

up with any more justification for my belief about the room number than I had started out with, but still if I started out with knowledge, I end up with knowledge again.

This means that an argument where a premise was justified in terms of the conclusion *can* be useful in counteracting the effects of nonrational or irrational loss of knowledge. This means that the cosmological argument could be useful even if none of the arguments for the PSR given earlier worked, and even if the PSR were not self-evident, for some people may know that the PSR is true because they once knew that God exists. They lost the knowledge that God exists but retained its shadow, the entailed belief that the PSR is true.

2.3. *Objections to the PSR*

2.3.1. **Modal imagination argument**

One can, arguably, imagine that a brick pops into existence uncaused. Therefore, one might conclude that it is possible that a brick pops into existence uncaused, and hence that the PSR is not a necessary truth. This is a popular Humean argument against the PSR.

The defender of the PSR can, of course, simply insist that the inference from imaginability to possibility is defeasible. After all, someone might imagine that a certain straightedge and compass construction trisects an angle,⁹ and if the inference from imaginability to possibility were indefeasible, it would follow that the construction *possibly* trisects an angle. But a mathematical construction *possibly* (in the metaphysical sense) trisects an angle if and only if it *actually* does so, and in fact we know that angles cannot be trisected with straightedge and compass. So the inference had better be defeasible. The defender of the PSR can then claim that the arguments for the PSR are so strong that the argument from imaginability of PSR failure, being defeasible, does little to shake our confidence in the PSR.

However, there is a better solution for the defender of the PSR, and this is to question the claim that the opponent has actually imagined a brick popping into existence uncaused. It is one thing to imagine something without simultaneously imagining its cause, and another to imagine something along with the absence of a cause. In fact, the task of imagining absences *as such* is a difficult one. If I tell an ordinary person to imagine a completely empty room, the subject is likely to imagine an ordinary room, with walls but no furniture. But has the subject really imagined an *empty* room? Likely not. Most likely the imagined room is conceptualized in a way that implies that it has *air* in it. For instance, we could ask our subject what it would be like to sit in that empty room for 8 hours, and our subject is unlikely to respond: “You’d be dead since the room has nothing in it, and hence no oxygen either.”

Could one with more directed effort imagine a room without any air in it? I am not at all sure of that. While we have the *concept* of vacuum as the absence of anything, it is not at all clear that we can imagine vacuum. Our language may itself be a giveaway of what we imagine when we imagine, as we say, a room “filled” with vacuum – perhaps we are not really imagining an empty room, but one filled with some colorless, frictionless, zero-pressure substance. Moreover, most likely, we are imagining the room as embedded in a universe like ours. But a room in a universe like ours will be pervaded with quantum

9. In fact, many people have imagined just that (see Dudley 1987).

vacuum as well as with electromagnetic and other fields, and perhaps even with spatial or spatiotemporal points. Whether these “items” count as things or not is controversial, of course, but at least it is far from clear that we have really imagined a truly empty room.

It is true that philosophers sometimes claim that they can imagine a world that, say, consists only of two iron balls (Black 1952). But a claim to imagine that is surely open to question. First of all, the typical sighted person’s imagination is visual. The balls are, almost surely, imagined visible. But if so, then it is an implicit part of what one is imagining that there are photons bouncing off the balls. Furthermore, unless one takes care to specify – and I do not know how one exactly one specifies this in the imagination – that the balls obey laws very different from those of our world, there will constantly be occasional atoms coming off the edges of the balls, and hence there will be a highly diffuse gas around the balls. Suppose all of this physics is taken care of by our careful imaginer. Still, have we really imagined a world containing only two balls? What about the proper *parts* of the billiard balls – does the world not contain those? What about properties such as *roundness*, or at least tropes such as *this ball’s roundness*? And are there no, perhaps, spatial or other relations between the balls? We see that unless one is a most determined nominalist, the content of the imagined world is going to be rather richer than we initially said. There are details implicit in the imagined situation that we have omitted.

There may, however, be a way we can imagine an absence. We can probably imagine absences of particular kinds of things in a particular area of space-time. Certainly, I can imagine a room free of talking donkeys, or even of donkeys in general. Moreover, I can probably imagine a room with no particles or electromagnetic fields in it. But that is not the same as imagining a truly empty room. A truly empty room does not have any other kinds of fields in it, at least if fields are things; there are no points of space or space-time in it; and it certainly has no ghosts, angels, or demons in it. But no list of kinds of things that we imagine as absent from the room will assure us of the literal and complete emptiness of the room, for there may always be a different kind of being, one utterly beyond the powers of our imagination, whose absence from the room we have failed to imagine. Nor will it do to imagine “unimaginables” as missing since “unimaginables” are not a genuine kind of thing but, surely, a mix of very different kinds of possibilities – it seems highly plausible that there are many kinds of possible things beyond our wildest imagination.

Similarly, we can imagine a brick coming into existence in the absence of a brickmaker, a brick not resulting from the baking of clay, a brick not made by an angel, demon, or ghost. But that is not the same thing as imagining a brick that comes into existence completely causelessly. To imagine that, we would need to imagine every possible kind of cause – including the unimaginable ones – as absent. That seems to be a feat beyond our abilities. We can, of course, *say* the words “This is causeless” both with our lips and with our minds while imagining the brick, but the claim that whenever one can imagine an *F* and say of it, with lips or minds, that it is a *G*, then, possibly, there is an *F* that is a *G*, would not only be highly defeasible but would also surely be a nonstarter. I can imagine a circle and say the words “This is a square” while imagining it.

Moreover, in general, when we imagine a situation, we imagine not a whole possible world, but a part of one, and our imagination is neutral on whether there are further support structures. I imagine three billiard balls on a billiard table. Probably, it is part of my imagining that there is gravity. Something, then, has to hold the table up, but what it is is not a part of the imagined situation. But I am not imagining a table miraculously

suspended in a gravitational field – I am simply not imagining what the outside of the situation has to be like to support the part I care about.

Maybe, with a lot of work, one can imagine a situation involving a brick and involving enough imagined detail that one can, with confidence, say that the situation is not only one where the ordinary causes of bricks are not present near the brick, but where nowhere in the universe are there any causes of the brick and where there are no nonphysical causes of the brick either. But now we see that the situation imagined took rather more effort, and the given examples of how there may be more to an imagined situation than one initially thought should severely reduce one's confidence that one has been successful at the task of imagining a causeless brick. And even if one has been successful at it, the inference to the possibility of a causeless brick is still defeasible.

I want to end this discussion by comparing the imaginability argument for a causeless brick with the imaginability argument against Platonism. One might claim that it is possible to imagine a brick that does not stand in an instantiation relation to any other entities. If one can, then defeasibly it follows that possibly a brick does not stand in an instantiation relation to any other entities. But that, of course, contradicts Platonism, which holds that, necessarily, all bricks instantiate brickness. While I am not a Platonist, this argument against Platonism strikes me as weak. The Platonist can answer as I did earlier: have we really imagined a brick that does not stand in an instantiation relation to another entity, or have we merely imagined a brick without imagining its standing in an instantiation relation?

But there is also a further answer the Platonist can make. The Platonist can say: "For all you know, by imagining it as a *brick* you have implicitly imagined a situation where it is related to brickness, although your description of the contents of what you imagined contradicts this." Compare this to the point one should make against someone who claims that to have imagined a cube without any space or spatial relations – surely, by imagining it as a cube, you have implicitly imagined it as occupying space or as involving spatial relations (say, between the vertices).

Can the defender of the PSR make this point too? Perhaps. The brick we allegedly imagine coming into existence *ex nihilo* is a contingent brick. But it might be that the nature of contingency involves being caused (cf. Section 2.2.6.6, above). Moreover, the brick has existence. But it seems implausible to claim that we have plumbed the depths of the nature of existence. It could, for instance, be that *to be* is either *to be necessary* or *to be caused* – that the *esse*, the *existence*, of a contingent being is its being caused (it may be that Thomas Aquinas thought this; I explore this kind of a view in Pruss 2006, chap. 12). It could even be that the *esse* of a contingent being is its being caused by that particular set of causes by which it is caused – that would cohere neatly with and explain the essentiality of origins.

A variant of the argument from modal imagination is to say that one can without overt logical contradiction state the claim that a brick exists without a cause:

$$(9) \quad \exists x(\text{brick}(x) \ \& \ \sim\exists y(\text{causes}(y,x))).$$

However, that is a bad argument. That one can state something without *overt* contradiction does not imply that there is no *hidden* contradiction. After all, compare (9) with:

$$(10) \quad \exists x(\text{sculpture}(x) \ \& \ \sim\exists y(\text{causes}(y,x))).$$

This claim is impossible since it is a necessary truth that sculptures have sculptors – that is what makes them be sculptures. In the case of (10) the contradiction lies pretty close to the surface. But how do we know that in (9), there is no contradiction somewhat further from the surface? Maybe there is even a hidden complexity in the concept represented by the existential quantifier.

2.3.2. Van Inwagen’s modal fatalism argument

2.3.2.1. *The basic argument*

Peter van Inwagen (1983, pp. 202–4) has formulated an influential and elegant *reductio ad absurdum* of the PSR. Let p be the conjunction of all contingent truths. If p has an explanation, say q , then q will itself be a contingent truth, and hence a conjunct of p . But then q will end up explaining itself, which is absurd. We can formulate this precisely as follows:

- (11) No necessary proposition explains a contingent proposition. (Premise)
- (12) No contingent proposition explains itself. (Premise)
- (13) If a proposition explains a conjunction, it explains every conjunct. (Premise)
- (14) A proposition q only explains a proposition p if q is true. (Premise)
- (15) There is a Big Conjunctive Contingent Fact (BCCF), which is the conjunction of all true contingent propositions, perhaps with logical redundancies removed, and the BCCF is contingent. (Premise)
- (16) Suppose the PSR holds. (for *reductio*)
- (17) Then, the BCCF has an explanation, q . (by (15) and (16))
- (18) The proposition q is not necessary. (by (11) and (15) and as the conjunction of true contingent propositions is contingent)
- (19) Therefore, q is a contingent true proposition. (by (14) and (18))
- (20) Thus, q is a conjunct in the BCCF. (by (15) and (19))
- (21) Thus, q explains itself. (by (13), (15), (17), and (19))
- (22) But q does not explain itself. (by (12) and (19))
- (23) Thus, q does and does not explain itself, which is absurd. Hence, the PSR is false.

Versions of this argument has been defended by James Ross (1969, pp. 295–304), William Rowe (1975, 1984), and, more recently, Francken and Geirsson (1999).

The argument is plainly valid. Thus, the only question is whether the premises are true. Premise (14) is unimpeachable.¹⁰

Premise (13) bears some discussion. In favor of it, one might note that the explanation of the conjunction might have more information in it than is needed to explain just one of the conjuncts, but if it has enough information to explain the conjunction it also has enough information to explain the conjuncts. We may, however, worry about Salmon’s remark that irrelevancies spoil explanations (Salmon 1990, p. 102). If we are worried about this, however, we can replace “explains” with “provides material sufficient for an explanation” throughout the argument, and whatever was plausible before, will remain plausible.

10. That aliens shot John F. Kennedy *would* be a good explanation of JFK’s death *were it true*, but since it is false, it is not an explanation. False propositions can be *putative* explainers, but not *actual* explainers.

Alternately, we may say that if q explains a conjunction, then the only reason it might fail to explain a conjunct r is because q might contain irrelevant information. But, surely, when the conjunct r is equal to q itself, this worry will not be real – how could q contain information irrelevant *to itself*? So even if (13) is questioned, (21) still very plausibly follows from (15), (17), and (19).

This leaves the technical Premise (15) about the existence of a BCCF, and two substantive claims, (11) and (12), about explanation. Leibnizian cosmological arguments based on the PSR need something like a BCCF, so questioning (15) is probably not a fruitful avenue for questioning for a defender of the Leibnizian cosmological argument (see further discussion in Section 4.1.1.3, below, as well as in Pruss 2006, sec. 6.1).

But we should *not* accept (11). We shall see that the main reason for believing (11) rests on a misunderstanding of how explanation works. Moreover, I shall argue that someone who accepts the logical possibility of libertarian free will should deny at least one of (11) and (12).

2.3.2.2. *Is (11) true?*

Premise (11), that no necessary proposition can explain a contingent one, needs some justification. The main reason to accept (11) is the idea that if a necessary proposition q explained a contingent proposition p , then there would be worlds where q is true but p is false, and so q cannot give the reason why p is true. This sketch of the argument can be formalized as follows:

- (24) If it is possible for q to be true with p false, then q does not explain p . (Premise)
- (25) If q is necessary and p is contingent, then it is possible for q to be true with p false.
(a theorem in any plausible modal logic)
- (26) Therefore, if q is necessary and p is contingent, then q does not explain p .

Instead of attacking (11) directly, I shall focus my attack on (24). Without (24), Premise (11) in the modal fatalism argument does not appear to be justified. Now, granted, someone might one day find a powerful argument for (11) not dependent on (24), in which case more work will need to be done, but (24) seems to capture just about all the intuition behind (11).

By contraposition, (24) is equivalent to

- (27) If q explains p , then q entails p .

Let me start with a quick *ad hominem* argument against (27). It seems a perfectly good explanation of why the dog did not bark that neither a stranger came by the dog nor did any other potential cause of the dog's barking occur. But the explanans here only entails the explanandum if we suppose that it is a necessary truth that if the dog barked, its barking had a cause. But opponents of the PSR are unlikely to grant that this is a necessary truth, unless they have some principled reason to argue that dogs' barkings metaphysically require causes, but some other things do not need any explanation, whether causal or not. But I doubt that there is a good way of drawing the line between barkings and other states of affairs.

Now, (27) does seem to hold in the case of many conceptual explanations. These explain a state of affairs by saying what the state of affairs is constituted by or consists in. For instance, in *Metaphysics Z*, Aristotle suggests explaining an eclipse of the sun by noting that an eclipse of the sun is *identical with* the earth's entry into the moon's shadow. Likewise, one might explain a knife's being hot by noting that its being hot consists in, or maybe is constituted by, its molecules having high kinetic energy.

However, (27) is falsified by just about every modern scientific *nonconceptual* explanation that I know of. Scientific causal explanations, in general, simply do not give conditions that *entail* the explanandum. This is obvious in the case of statistical explanations, since in these, the explanans gives laws of nature and states of affairs that do not entail the explanandum but either render the explanandum more probable than it would otherwise be, or at least are explanatorily relevant to the explanandum. Why did the cream spread throughout the coffee cup? Because it is *very likely* that random molecular motion would disperse cream in this way.

But the falsity of (27) also follows in the case of nonstatistical explanations. Why are the planets moving in approximately elliptical orbits? Because the main gravitational influence on them is that of an approximate point mass (the sun), the secondary gravitational influences on them from other objects, including other planets, being weak. But what I just said does not entail that the planets move in approximately elliptical orbits. It only entails that *absent* other influences, the planets move in approximately elliptical orbits.

Perhaps we can build that proviso into the explanation. Why do the planets move in approximately elliptical orbits? Because the main gravitational influence is that of an approximate point mass, *and there are no other relevant influences*. But the word "relevant" is crucial here, for, of course, there are many other influences, such as electromagnetic ones. The word "relevant" here seems to mean "relevant to affecting the approximate shape of the planets' orbits." But that is not quite right. Electromagnetic influences of the sort that the planets undergo are, in general, relevant to affecting the approximate shape of the planets' orbits. They are just not relevant *in this case* because the gravitational influence of the sun swamps the other effects.

So what we are really saying when we give the proviso is that the main gravitational influence is that of an approximate point mass, *and no other influence prevents the orbits from having elliptical shape*. Perhaps now we have entailment? Actually, the objector to the PSR should not say so. For if the PSR is false, then surely things can come into existence for no reason at all and can likewise pop out of existence for no reason at all. Thus, it is quite possible for all of the aforementioned to be true, and yet for the planets to pop out of existence, thereby preventing them from having any orbits at all.

Do we add, then, to our explanans the conjunct "and the planets remain in existence"? Perhaps then we will get entailment, although even that is not at all clear. For if the PSR is false, and if the laws of nature are of such a nature as not to rule out the possibility of external influence, then it seems likely that the laws of nature cannot rule out the possibility of a brute, unexplained departure from the laws of nature.

Or perhaps objectors to the PSR will admit that by their own lights (27) is false, but insist that the defenders of the PSR are committed to (27). Then the argument against the PSR becomes *ad hominem*. I have no objection against *ad hominem* arguments, but one would need to give an argument that while the opponents of the PSR can reasonably reject (27), for some reason the proponents of the PSR should accept (27). But then an argument

is needed as to why the proponents of the PSR must accept a theory of the nature of explanation that requires (27) just because they happen to think that all contingent facts have explanations. Yet since the PSR is incompatible with (27) given some plausible other assumptions, this would be a hard case to make!

In any case, suppose that through a lot of careful work we have somehow managed to come up with an explanans that entails that the planets have approximately elliptical orbits. An easy point I can make here is that the resulting explanation is unlike *standard* scientific explanations in more than one way.

First, note that with the provisos we have loaded into the explanans, the explanation becomes logically odd. What we end up saying is essentially that the planets move in approximately elliptical orbits because the gravitational influence of an approximate point mass moves them in approximately elliptical orbits. The provisos all add up to saying that the gravitational influence of the sun *succeeds* in moving the planets in elliptical orbits. But now the explanandum is in effect there in the explanans, and our explanation is like “He died because he died of being stabbed.” But that is not how we give explanations. He died *because he was stabbed*. The planets move approximately elliptically *because the sun gravitationally influences them*. He did not die because he died of being stabbed, and the planets do not move approximately elliptically because the sun moves them approximately elliptically.

Second, the proponents of the PSR have an epistemic right to reject the loading up of the explanation of provisos, a right grounded in reasons apparently independent of their need to reject the van Inwagen argument. Our provisoed explanation basically was: “The planets move approximately elliptically because the sun gravitationally influences them as an approximate point source and nothing prevents them from moving approximately elliptically.” But if the PSR is a necessary truth, then that nothing prevents the planets from moving approximately elliptically *entails* that they in fact move approximately elliptically, since if they did not, there would have to be a reason why they do not. However, it is an odd sort of explanation where one of the conjuncts in the explanans is sufficient by itself to entail the explanandum. One wonders why one bothers mentioning the sun’s gravitational influence at all! In fact, this worry may be there even for the PSR’s opponent, if one of the provisos has to be something like “and the PSR is not relevantly violated in this case.”

Third, the claim that *nothing* prevents the planets from moving approximately elliptically involves universal quantification over all entities in existence, whether natural or not. It is a claim that each of these entities is a nonpreverter of the planets’ elliptical motion. But while scientific claims have *entailments* about nonnatural entities (e.g. that the planets are moving approximately elliptically entails that God is not making them move along logarithmic spirals), they should not quantify over nonnatural entities. Thus, our heavily provisoed explanation does not appear any longer to be a scientific one.

Let me end this section with the following argument for (27). The PSR had better understand “explains” as “gives a sufficient reason for.” But a sufficient reason is, surely, a *logically* sufficient reason, that is, an entailing reason. And, indeed, Leibniz thought that the reasons said by the PSR to exist would be entailing.

A simple answer to this is to say that I am not defending Leibniz’s PSR, but a PSR sufficient for the cosmological argument. We do not, in fact, need entailing reasons for the cosmological argument, as shall be clear when we discuss cosmological arguments. A fuller

answer is that when I talk of the PSR, by “sufficient reasons” I mean reasons that are sufficient *to explain* the explanandum. Leibniz may have erroneously thought that a reason is only sufficient to explain something that it entails, but we do not need to follow him in his error – and should not, since that route leads to modal fatalism. But if the reader is not convinced, I can just rename the principle I am defending the “Principle of Good-Enough Explanation.”

2.3.2.3. *Libertarian free choices*

Let me now offer an argument that someone who accepts the possibility of libertarian free will *must* reject the van Inwagen argument. Since van Inwagen is a libertarian, he too must reject his own argument. To make this more than an *ad hominem*, I would need to argue for the possibility of libertarian free will (or for its actuality), for which, of course, there is no space here.

Libertarian free will is nondeterministic. From the condition of the mind of the chooser prior to the choice, one cannot deduce what choice will be made. This has given rise to the *randomness* objection to libertarianism: libertarian free choices are not really caused by the person, but are merely *random* blips, as some people think quantum events are. We would not account a person free if acts of will occurred *randomly* in a person’s mind or brain.

Libertarians are, of course, committed to a denial of the randomness objection. However they manage it, they must reject the claim that libertarian free actions are random – they may, for instance, insist that they are not random because they are caused by agent causation. Now suppose that a libertarian allowed that in the case of a libertarian free choice between options *A* and *B*, where in fact *A* was chosen, there is no sufficient explanation of why *A* was chosen. Such a libertarian has succumbed to the randomness objection. If there is no explanation for why option *A* was chosen, then that *A* was chosen is a brute, unexplained, uncaused fact – a *random* fact. Thus, the libertarian cannot allow that there is no explanation of why *A* was chosen.

Look at this from another direction. Suppose someone is externally determined to choose *A* instead of *B*, so that the explanation for why *A* was chosen was that some external puppet master has caused the agent to choose *A* rather than *B*. In that case, there would indeed be an explanation for why *A* was chosen rather than *B* – the causal efficacy of the puppet master. The libertarian will insist that in that case, there is no free will. Now take this situation and *subtract* the puppet master, without adding anything. We get a situation where there is no explanation for the choice of *A* rather than of *B*. We get a genuine case of randomness, having replaced the puppet master by nothing at all. And this mere removal of the puppet master does nothing to give freedom to the agent. Libertarian freedom is not supposed to be something purely negative, the lack of a puppet master, but something positive like self-determination. To go from the choice determined by the puppet master to a genuine libertarian free choice, we cannot merely delete the explanation of the action in terms of the puppet master: we must *add* something to the situation. It is plausible that what needs to be done is to substitute the free agent and/or her free will for the puppet master: the action must be explained in terms of the agent now, instead of in terms of something external. The basic intuition of a libertarian is that determinism places the ultimate point of decision outside the agent, in the

environment that forms and influences the agent. This external determinism, to produce freedom, must not only be removed but must be replaced by something causal, though still indeterministic in the case of agents that have a cause of their existence,¹¹ within the agent.

Thus, the libertarian should hold that there is an explanation for why one rather than another choice was freely made. Otherwise, the randomness objection to libertarianism succeeds. This either forces the libertarian to say (a) that a description of a mind in a state that is equally compatible with either of two actions, *A* or *B*, can be used to explain why *A* was in fact freely chosen – a denial of (27) – or (b) that the claim that action *A* was freely chosen, or perhaps freely chosen for reason *R*, is “almost” a self-explanatory claim, despite its contingency, with the only thing unexplained being why the agent existed and was free and perhaps impressed by *R*. If the agent were a being that necessarily exists and is necessarily freely and omnisciently, then in case (b), nothing would be left unexplained, and we would have a counterexample to (12).

Nonetheless, how there can be explanation of exercises of libertarian free will is mysterious. I shall here defend option (a), that a choice of *A* can be explained in terms of a state that was compatible with choosing *B*. I shall defend this by offering a hypothesis about how libertarian free will works. If this hypothesis is false, perhaps another can do the same job, but I find this one plausible. For simplicity, I will assume a binary choice between *A* and *B*. On my hypothesis, free choices are made on the basis of reasons that one is “impressed by,” that is, that one takes into consideration in making the decision. Some of the reasons are in favor of one choice, and others are in favor of another choice. Reasons that are neutral between the options are not taken into account by the agent in the choice between *A* or *B*.

I now suppose that when the agent *x* chooses *A*, there is a subset *S* of the reasons that favor *A* over *B* that the agent is impressed by, such that *x* freely chooses *A* on account of *S*. My explanatory hypothesis, then, is that *x* freely chooses *A* because *x* is making a free choice between *A* and *B* while impressed by the reasons in *S*. On my hypothesis, further, had the agent chosen *B*, the agent would still have been impressed by the reasons in *S*, but the choice of *B* would have been explained by *x*'s freely choosing between *A* and *B* while impressed by the reasons in *T*, where *T* is a set of reasons that favor *B* over *A*. Moreover, in the actual world where *A* is chosen, the agent is also impressed by *T*. However, in the actual world, the agent does not act on the impressive reasons in *T*, but on the reasons in *S*.

This explanation fits well with how agents in fact describe their choices. They say things like: “I chose this graduate school because it was important to me that my spouse be able to study at the same institution.” Sure, another school might have better fit with their academic interests, and that may also be important to them. But while the latter consideration is one they are also impressed by, they did not in fact choose on the basis of it, and hence it does not enter into the explanation.

11. The reason for the proviso is this: If agent *x* is caused by *y* to exist, and is *internally* determined to do *A*, then *y* by causing *x* to exist has caused *x* to do *A*. But if agent *x* has no cause of its existence, then this argument no longer works, and internal determinism may be compatible with freedom. This is important for the question whether a God who cannot choose evil can be free (see, for instance, Pruss 2003).

Note that I am not claiming that the same thing explains the choice of *A* as would have explained the choice of *B*. That sameness of explanation claim might seem absurd, so an opponent might try to push me to admit that there is a single explanation in both cases. Thus, one might say that if *p* (a proposition about the choice being made while impressed by the reasons in *S*) explains the choice of *A*, and *q* (which is like *p* but with *T* in place of *S*) is true, then *p*&*q* also explains the choice of *A*. However, “irrelevancies [are] harmless in arguments but fatal in explanations” (cf. Salmon 1990, p. 102). Thus, even though *p* explains the choice of *A*, and *q* is true, one can coherently deny that *p*&*q* explains the choice of *A*. If, on the other hand, this point is denied, I will regroup by saying that the idea that the same proposition should explain *A* in our world and an incompatible *B* in another world is defensible. Salmon (1990, pp. 178–9) argues that one *must* accept the possibility that the same kinds of circumstances can explain one event on one occasion and an incompatible kind of event on another occasion if one is to have any hope of explaining stochastic outcomes.¹² For instance, if a carcinogen causes cancer 12 percent of the time, with 60 percent of the time its being type *A* cancer and 40 percent of the time its being type *B* cancer, these statistical facts can explain both an occurrence of type *A* cancer in one patient and an occurrence of type *B* cancer in another.

For the cosmological argument, the most important case of libertarian choice is God’s choice what world to create. In this case, I actually think it is a *necessary* truth that God is impressed by the reasons *S* on account of which he created the actual world, just as it is a necessary truth that God was impressed by a different set of reasons on account of which he might have created another world. After all, necessarily, an omniscient and morally perfect God is impressed by all and only the good reasons. What the reasons on the basis of which God created this world are is something largely beyond my ken, although we can say a few standard things about the value of beings that participate in God’s life.

As a modification of my hypothesis, I should note that it might be that what matters explanatorily is not only the fact of the agent’s being impressed by the reasons, but also the degree to which the agent is impressed by them. It is easy to modify the account to take this into account, by explaining not just in terms of a set of reasons but in terms of a set of reason–weight pairs.

There is, still, something uncomfortable about the proposed explanation of libertarian action. I think a reader is likely to have the sense that while it is correct to say that the choice of graduate school might be explained by what is better for a spouse, even though this reason would have equally been present had a choice not supported by this reason been made instead, this kind of explanation is explanatorily inferior to, say, deterministic causal explanation or explanation in terms of a necessitating metaphysical principle. That may be. But there is no need to take the PSR to say that there is always *the best* kind of explanation – the PSR I am defending merely says that there is *an* explanation of every contingent proposition. And that is all I need for the cosmological argument.

12. Note that Salmon’s definition (Salmon 1990, p. 67) of a statistical relevance explanation of a fact as simply being an assemblage of statistically relevant facts implies the claim that the explanation of *p* in one world will be an explanation of $\sim p$ in another if we add the observation that sometimes, maybe even always, whatever is relevant to *p* will also be relevant to $\sim p$.

2.3.3. A probabilistic version of van Inwagen's argument

But even if van Inwagen's argument fails, there is a probabilistic variant that does not rely on (11). This argument is inspired by some remarks I got from Peter Forrest. Instead of (11), the argument uses the following claim:

(28) If q explains p , then $P(p|q) > 1/2$.

Instead of concluding that the BCCF *in fact* does not have an explanation, the argument will conclude that some worlds have a BCCF that does not have an explanation. We proceed as follows. Making use of all the other premises of van Inwagen's argument, generalized to hold in all worlds, we get the claim that in every possible world there is an explanation of the BCCF, and the explanans is a necessary proposition. Now, if q is a necessary truth, then $P(p|q) = P(p)$. Conditioning on necessary truth gets us no new probabilistic information beyond prior probabilities. Hence, in any world w , if the BCCF p of w is explained by a necessary truth, then $P(p) > 1/2$ by (28). Therefore, the BCCF of every possible world has probability greater than $1/2$. But the BCCFs of different worlds are mutually exclusive, since any two worlds differ in the truth-value of some contingent proposition, and then the BCCF of one of the worlds will contain that proposition and that of the other will contain its denial. Hence, if p_1 and p_2 are the BCCFs of two distinct worlds, we have $P(p_1 + p_2) = P(p_1) + P(p_2) > 1/2 + 1/2 = 1$. But no probability can be bigger than 1, and absurdity ensues again.

A defender of the PSR could, of course, deny (12), which this version of the argument presupposes, since otherwise we could have self-explanatory contingent explanations of the BCCF. A desperate, but not entirely unjustified, alternate measure would be to deny the assumption that if q is necessary, then $P(p|q) = P(p)$, perhaps allowing that this is true if q is a tautology, and maybe even any narrowly logically necessary truth, but not if it is a substantive necessary truth, such as that horses are mammals, that water is H_2O or that God values unity and happiness. It could, then, be the case that p_1 has probability greater than $1/2$ given one necessary truth q_1 , while a proposition p_2 incompatible with p_1 has probability greater than $1/2$ given another necessary truth q_2 . For instance, perhaps that the universe consisting only of a single particle has high probability given that God values unity, and that the universe containing infinitely many happy persons has high probability given that God values happiness, even though it is a necessary truth that God values *both* unity *and* happiness.

The best way out for the PSR's defender, however, seems to be to oppose (28). First of all, statistical relevance theories of explanation deny (28), and, more broadly, (28) may be a manifestation of the mistaken conflation of explanation with prediction that plagued both the deductive–nomological (Hempel & Oppenheim 1948) and inductive–statistical (Hempel 1962) models of explanation. A standard counterexample to these models is the syphilis/paresis case (Scriven 1959; see also the discussion in Salmon 1990, sec. 2.3), which is also a counterexample to (28). We can explain why a person has paresis in terms of the earlier having of latent untreated syphilis, even though latent untreated syphilis leads to paresis only in a minority of cases.

Second, it is plausible that citing the relevant actual cause of an event explains the event. Indeed, to give causes is a paradigmatic way of explaining. But causation can filter through indeterministic events of probability less than $1/2$. This is particularly clear in the case of

forensic explanations. George murderously pushes Maurice off a very high cliff. Maurice falls and drowns. Unbeknownst to George, Maurice is actually a cliff diver, and had a 75 percent chance of survival for the fall. Nonetheless, George's murderous push killed Maurice, and George's having pushed Maurice explains why Maurice died. Granted, in this case it does not explain *everything* about why Maurice died. It does not explain, for instance, why in this case Maurice did not manage to swim out or why lack of oxygen kills earthly vertebrates. But it is still a fine explanation. In any case, it could well be that even after we answered all of these questions, it would be that the explanans made the explanandum less than 50 percent probable – there could be indeterministic quantum events in Maurice's brain behind Maurice's inability to swim out.

Third, if libertarianism holds, and if a plausible account of action requires one to say that free choices are explained by the agent's reasons, we have reason to deny (28). For it seems likely that libertarian-free agents can act on reasons that they had probability less than 1/2 of acting on.

2.3.4. Quantum mechanics

A common objection to the PSR is that indeterministic quantum effects lack sufficient reasons. However, the PSR that I am defending concerns explanation, which is the giving of reasons sufficient *to explain* the explanandum, not the giving of reasons logically sufficient for entailing the explanandum.

Quantum mechanical events do, however, have explanations. The experimental setup in which they happen has the property of giving rise to emissions with certain probabilities (John Haldane makes this suggestion in Smart & Haldane 2003, p. 126). Granted, on indeterministic accounts of quantum mechanics, this explanation does not entail the outcome, and will only be a statistical explanation, perhaps involving small probabilities, but as the syphilis/paresis case in the previous section showed, that should not be a problem.

Still, one might be somewhat dissatisfied with quantum mechanical explanations. One might say that, yes, they are explanatory, but they lack some feature that better explanations have. But that is fine for the defense of the PSR. The PSR does not say that for every contingent proposition there is the best possible kind of explanation, but just that there is *an* explanation, "an 'explanation enough'" in Haldane's words (Smart & Haldane 2003, p. 126). And the kind of explanation that our PSR provides will be, as we shall see, enough to yield a cosmological argument – and that is the point here.

2.3.5. Contrastive explanation and the PSR

Perhaps, though, we can formulate the dissatisfaction with statistical quantum mechanical and libertarian explanations as follows. Suppose we are dealing with an electron in a mixed $|\text{up}\rangle + |\text{down}\rangle$ state, which in an appropriate magnetic field will either go up or down, with equal probability. Suppose it goes up. Why did it go up? Because of its state, the experimental setup, and the laws of nature. Maybe this is a fine explanation, but it does not seem to be a *contrastive* explanation. It does not explain why the electron went up *rather than* down.

The simplest move at this point is just to deny this intuition, and say that the same facts can explain why it went up rather than down, as would have explained why it went down rather than up. Alternatively, one might distinguish the quantum mechanical and

libertarian cases. Perhaps one can take a deterministic interpretation of quantum mechanics, and in libertarian cases give contrastive explanations in terms of different sets of reasons, as in Section 2.3.2.3, above.

Another move available to the defender of the PSR is to admit the failure of contrastive explanations. But the PSR says that every contingently true proposition p has an explanation, not that for every pair of propositions p and q where p is contingently true and q is a relevant alternative to p , there is an explanation of why p rather than q holds. There may well be such a notion of explanation that would make explanation be a ternary relation, but there is also a perfectly fine notion of explanation that makes explanation a binary relation, and it is the latter that the PSR concerns.

Some do, however, believe that all explanation is contrastive (cf. Dretske 1972; van Fraassen 1980). The standard example is something like this. George ate a banana rather than eating an orange because he liked bananas. George ate a banana rather than putting it in his backpack because he was hungry. Without specifying a contrast, we cannot tell which explanation we are after.

Arguments like this do not, however, establish that explanation is *always* contrastive. If we do not specify a contrast, we can give an explanation along either set of lines. George ate a banana because he liked bananas and chose to eat. George ate a banana because he was hungry and chose a banana. Neither explanation tells the whole story. But we can elicit more of the story by applying the PSR again. Why did George like bananas and choose to eat? Granted, we might say that this is because he likes nonjuicy sweet fruit and chose to eat, leaving that second conjunct as yet unexplained. But if explanation comes to an end in an ultimate explanation, we cannot *just* keep on furthering the explanation of the first conjunct – eventually, we will be done with that side, and a further demand for explanation will force us to tackle the question why George chose to eat.

A different move is that *on its own terms* the PSR that I am defending requires contrastive explanation. After all, it requires an explanation of every contingent proposition and that the electron went up rather than going down, or that George ate a banana rather than eating an orange, is a perfectly good proposition.

There is room to be quite unsure here, though. For it might be argued that when we make a contrastive claim, we are doing two things. We are asserting a proposition with an “and . . . not” truth-functional connective, for example, that the electron went up and did not go down, and drawing the listener’s attention to the contrast between the two claims joined by the truth-functional connective. The proposition asserted, however, is not contrastive in nature and can be explained straightforwardly. We can just give the statistical explanation of why the electron went up and explain that if it went up, it could not have gone down at the same time, so it went up and did not go down.

There is reason to think this is the right way to understand contrastive claims. First, note that whatever proposition is asserted by saying “ p rather than q holds,” necessarily, it is a proposition that is true if and only if p is true and q is false. To see this, begin by observing that if p is not true or q is not false, then whatever “ p rather than q holds” expresses must be false.

The converse is more difficult to establish. There certainly are cases where the sentence “ p rather than q holds” is not assertible even though “ p holds” and “ q does not hold” are assertible. These will be cases when there is no relevant contrast between p and q . Thus, in typical contexts, “The moon is spherical rather than Jupiter being cubical” is not assertible. However, the failure of assertibility is not due to facts about the objective situation being

talked about, but due to one's concerns, interests, and epistemic position. There will be epistemic contexts involving no mistakes but where "The moon is spherical rather than Jupiter being cubical" is assertible. For instance, suppose that George has seen neither the moon or Jupiter, and nobody has told him anything about them, except that an epistemic authority testified to him that the moon is spherical or Jupiter is cubical. One day, George learns that Jupiter is spherical. He then correctly sums up his conclusions: "The moon is spherical rather than Jupiter being cubical!" Given knowledge that p and that not- q , the assertibility of " p rather than q holds" depends on nonalethic matters, and hence all we need for *truth* is p and not q .

One might object that " p rather than q holds" *asserts* something about the state of mind of the speaker – that it is a mind-dependent proposition. But that is completely mistaken, since, then, every "rather than" claim would entail the existence of a speaker saying that claim, but that the moon is spherical rather than Jupiter being cubical entails nothing about a speaker who is saying that the moon is spherical rather than Jupiter being cubical.

So if proposition r is expressed by " p rather than q holds," then, necessarily, r holds if and only if $p \& \sim q$. I think it is simplest to suppose that r is actually the same proposition as $p \& \sim q$.

But suppose this is denied, and it is said that there is "something more" in the proposition that p rather than q holds than in $p \& \sim q$ (for surely there is nothing less). Nonetheless, the contrastive explanation argument can be questioned. It is no coincidence that p rather than q holds if and only if $p \& \sim q$ holds – it is a necessary truth, in light of the said argument, that this is always the case. In fact, it seems right to say that what makes it be true that p rather than q holds is simply that p holds and q does not hold. The fact that $p \& \sim q$ seems to be the more basic, the more primitive, since the fact that p rather than q holds contains it and that mysterious "something more." But then, this provides a conceptual explanation of why it is the case that p holds rather than q : p holds rather than q because $p \& \sim q$ holds, and $p \& \sim q$ is more ontologically basic, and necessarily whenever $a \& \sim b$ holds, a rather than b holds. Granted, this is not a contrastive explanation; but that only shows that the attempt to assimilate contrastive explanations to explanations of contrastive propositions failed.

3. Nonlocal CPs

3.1. From local to nonlocal CPs

A local CP is a principle that every localized contingent item of a certain sort has a cause. Thus, a local CP about contingent substances holds that every *substance* has a cause. A cosmological argument making use of a local CP needs to rule out infinite regresses. On the other hand, a nonlocal CP lacks the restriction that the items be localized in the way substances and events are, and this allows one to get out of infinite regresses. Using a CP instead of the PSR has the advantage that avoiding the van Inwagen problem and its relatives is easier.

I shall argue that the intuitions that typically make local CPs plausible apply just as well to nonlocal CPs. The locality restrictions are objectionably *ad hoc*, and if we should accept a local CP, we should accept a nonlocal CP. I will then give an argument for a powerful CP.

Consider first a restriction of CPs to (localized) substances as opposed to substance-like aggregates such as heaps of sand or the mereological sum of all the contingent substances now existing. The basic intuition behind such CPs is that bricks and other objects cannot come into existence without cause. But suppose that we learned from the correct metaphysics that bricks are not actually substances but are heaplike (as is indeed what Aristotle's metaphysics says about bricks). That would not affect our commitment to the impossibility of bricks coming into existence *ex nihilo*.

Now, maybe, we could argue that a CP restricted to substances would suffice to show that nonsubstantial items such as bricks that are made up of a finite number of substances (maybe elementary particles are substances even if bricks are not) have causes. For we could just apply the CP separately to each of the component substances, and while some of them could be causes of others, it could not be true of all the component substances that they are caused by other component substances, since that would require either a causal loop or an infinity of component substances. Thus, the restricted CP is sufficient to do justice to our intuition that bricks do not causelessly pop into existence even if bricks are heaps.

But this argument only works if bricks are made up of a finite number of substances. However, suppose we found out that a brick was, in fact, made up of an infinite number of particles. It does not look right now as if physics is heading in the direction of positing an infinite number of particles in ordinary material objects, but unless there is some logical problem with actual infinities – which problem would then be grist for the *kalam* arguer's mill, so an atheist will probably not want to embrace that option – the possibility is not absurd. Finding this out would not, I think, shake our conviction that a brick cannot pop into existence. Should it not be, if anything, harder for more particles to pop into existence? Nor would we be impressed by being told that the brick made of infinitely many particles popped into existence by the following pattern. At time $t_0 + 1$ second, particle number 1 was caused to exist by particle number 2; at time $t_0 + 1/2$ second, particle number 2 was caused to exist by particle number 3; at time $t_0 + 1/3$ second, particle number 3 was caused to exist by particle number 4, and so on, with none of the particles existing at time t_0 . That would still count as an objectionable causeless popping into existence of the brick. But if bricks are not substances, this possibility cannot be ruled out by a CP restricted to substances. However, our intuitions call for this to be ruled out.

Perhaps we can restrict the CP to entities that consist of less than the sum total of all contingent beings. But that will gain the opponent of global CPs nothing. For instance, let S_0 be any simple particle that has a contingent cause (there are in fact many such) and let S_1 be the aggregate of all other contingent beings now in existence. Let C_1 be a cause of S_1 , by the restricted CP. As the contingent cause of S_0 is outside of S_0 (since a particle cannot be caused by itself), this cause must be a part of S_1 , and hence is caused by C_1 . By transitivity, C_1 will also be the cause of S_0 . If there are no simple particles, the argument is slightly more elaborate and is left as an exercise to the reader (hint: just let S_0 be a cat, and note that the cat surely has a cause that is outside of it).

And, certainly, it will not do to restrict the CP based on size, absurdly as if objects that are less than 10 m in diameter needed causes, but larger objects like the universe did not. Here, it is worth recalling Taylor's example of the universe being like a walnut (Taylor 1974, chap. 10). If we accept that *then* it should have a cause, we should also accept it when it is much larger.

A different kind of restriction is diachronic in nature. Perhaps, the CP can only be applied to entities that exist all at one time and cannot be applied to causal chains of

entities, or at least to infinite such chains. However, once again, such a CP will fail to rule out a brick's doing what intuitively should count as popping into existence causelessly. Suppose we saw a brick pop into existence in midair. We would be deeply puzzled. To allay our puzzlement, a scientist tells us that study of the phenomenon reveals that what actually happened was this. The brick popped into existence at $t_0 + 1$ millisecond. There was no cause at t_0 . However, at $t_0 + 1/2$ milliseconds, the particles of the brick were caused by a set of earlier particles making up a brick, which then immediately annihilated themselves (or perhaps underwent substantial change into the new ones). At $t_0 + 1/3$ milliseconds, the earlier particles were caused by a yet earlier set. And so on, *ad infinitum*. The whole infinite sequence took 1 millisecond to complete, but nonetheless each synchronic collection of particles had an earlier cause.¹³ Surely, this would still be as objectionable as a brick popping into existence causelessly. That there was an infinite sequence of bricks, or of sets of particles, in that millisecond does not seem to affect the idea that this cannot happen.

Thus, to rule out the popping of bricks into existence *ex nihilo*, we need a CP not restricted in a way that rules out infinite chains. Similar considerations rule out CPs concerning events that do not generalize to infinite chains of events. A fire could start for no cause via an infinite chain of events, each temporal part of the fire being caused by an earlier temporal part of the fire, and so on, with the whole infinite chain only taking a second, and the temporally extended conflagration having no cause. That is absurd, and CPs for events should rule it out.

But perhaps there is a difference between infinite chains that take an infinite amount of time and infinite chains, like the one in the previous examples, that take a finite amount of time. Maybe we can restrict CPs to chains of causes that take a finite amount of time?

I do not think this is plausible because an interval from minus infinity to a finite number is order-isomorphic to a finite interval. For instance, the function $f(t) = -1/(t - 1)$ maps the half-infinite interval¹⁴ $(-\infty, 0]$ onto the half-open finite interval $(0, 1]$, while preserving order relations so that $t_1 < t_2$ if and only if $f(t_1) < f(t_2)$ for t_1 and t_2 in $(-\infty, 0]$.

But perhaps there is something relevantly metaphysically disanalogous about infinite amounts of time as opposed to finite ones. One difference would be if infinite amounts of time were impossible. But if so, then the *kalam* argument again shows up, and in any case, if an infinite amount of time is impossible, then a restriction of the CP to chains that take a finite amount of time is no restriction at all.

Another potential difference is that one might argue that a finite temporal interval either is preceded by a time or at least *could* be preceded by a time (if time has a beginning at the start of the interval), but a temporal interval infinite at its lower end *could not* be preceded by a time. This would be in support of a restriction of the CP to causal chains that are not temporally infinite in the direction of the past.

Consider first the following version of this disanalogy: the interval $(-\infty, 0]$ is not preceded by an earlier time, while a finite interval $(0, 1]$ is. But suppose that we are dealing with a causal chain spread over a half-open finite interval that is *not* preceded by an earlier time, since time starts with this half-open finite interval. In that case, we need a cause for the chain as a whole just as much as in the case where the half-open finite interval *is*

13. Examples like this go back at least to Łukasiewicz (1961), who tried to use them to show how free will could be reconciled with determinism (see also Shapiro 2001).

14. Here I use the standard notation where $(a, b] = \{x: a < x \leq b\}$.

preceded by an earlier time. Suppose that we have a chain of causes spread over $(0, 1]$, tending to being temporally positioned at 0 in the backwards limit, and suppose that there is no “time 0.” Surely, the nonexistence of a time prior to the interval makes it, if anything, “harder” for a chain of causes to arise without an external cause. After all, if even time does not exist prior to the chain of causes, then the chain is even more a case of coming into existence *ex nihilo*, since there is even less there. The absence of an earlier time does nothing to make it easier for things to arise causelessly.

Maybe, though, the idea is that we should require causes where causes can be reasonably demanded. But, the argument continues, a cause can only be reasonably demanded if there is a prior time, since causes must be temporally prior to their effects. However, the latter thesis is dubious. Kant’s example of a metal ball continually causing a depression in a soft material shows that simultaneous causation is conceivable. And apart from full or partial reductions of the notion of causation to something like Humean regularity and temporal precedence, I do not think there is much reason to suppose that the cause of a temporal effect must even be in time.

Now consider the second version: even if there is no “time 0,” there *could* be one. The finite interval $(0, 1]$ could be preceded by a time, while the interval $(-\infty, 0]$ could not. But it is quite unclear why this alleged modal difference is at all relevant to the existence of a cause for the chain. The absence of the possibility of an earlier time does not seem relevant, unless perhaps one thinks that causation requires temporal priority, a thesis one should, I think, reject.

There is also another, more controversial, objection to such a restricted CP. An infinite interval $(-\infty, 0]$ can be embedded in a larger temporal sequence $[-\infty, 0]$ obtained by appending a point, which we may call $-\infty$, and which stands in the relation $-\infty < x$ to every point x of $(-\infty, 0]$. It may well be that such a sequence could be the temporal sequence of some possible world. And, if so, then the interval $(-\infty, 0]$ *could* be preceded by an earlier time, and the disanalogy disappears.

It appears, thus, that local CPs are restricted in an *ad hoc* way. If we have strong intuitions in favor of local CPs, then we likewise should accept unrestricted CPs that can be applied to infinite chains of global aggregates of entities.

3.2. A modal argument for the CP

3.2.1. The basic argument

3.2.1.1. The effect would not have occurred without the cause

I will formulate this argument in terms of the cause of a contingent state of affairs. “States of affairs” here are understood as concrete things that can stand in causal relations, rather than as abstracta, and that exist if and only if they obtain. Moreover, I shall assume that states of affairs are individuated in such a way that in every world where Socrates is sitting at t_0 , his sitting at t_0 is the very same state of affairs, just as the proposition that he is sitting at t_0 is the same proposition in every world. States of affairs are, thus, fine-grained, and Kripkean essentiality of origins does not hold for them – the state of affairs of Socrates’ sitting at t_0 is the same state of affairs regardless of what caused it.

The argument now bootstraps its way from the *possibility* of a cause to the *actuality* of a cause. Thomas Sullivan (1994) has tried to find an argument for the CP based on the idea that a cause was a necessary condition for the effect. While this requirement is too strong, something in the vicinity of the following fact should be true:

- (29) That *C* causes *E* entails that were *C* not to exist or take place, *E* would not have taken place.

Claim (29) is not meant to be a complete analysis of causation, and, anyway, it requires that in cases of causal overdetermination we describe *C* carefully, for instance, as a disjunctive state of affairs. But something like this counterfactual claim is certainly a part of our notion of causation. David Lewis thought that this counterfactual claim was at the root of a complete analysis of causation, but this further controversial claim will not be needed.¹⁵

Suppose now that an airplane crashes due to metal fatigue in the ailerons. Then the following nested counterfactual is true:

- (30) Were the plane earlier hit by a surface-to-air missile, then the plane would have crashed and it would have been the case that were the plane not hit by a surface-to-air missile, the plane would, or at least might, still have crashed.

The plane would or might still have crashed because of the metal fatigue in the ailerons. Analogously, one might say this. Suppose that an airplane crashes for no reason at all. Then the following nested counterfactual should be true by parallel to (30):

- (31) Were the plane hit by a surface-to-air missile, then the plane would have crashed and it would have been the case that were the plane not hit by a surface-to-air missile, the plane would, or at least might, still have crashed.

Presumably, the consequent of the inner counterfactual can be taken to say that the plane would or might still have crashed for no reason at all. But this results in the absurdity that in the counterfactual world *w* where the plane is hit by a surface-to-air missile, and where no other crash-inducing causes are available (since the counterfactual that moved us to that world presupposed only one added cause – the surface-to-air missile), it is the case that were the missile not to have hit, the plane would or might still have crashed, contradicting the fact that the missile is the cause in *w* of the plane's crashing. Therefore, we should reject the possibility of the assumption that an airplane crashes for no reason at all.

As it stands, the argument may be thought to rest on improperly assimilating the case where the plane crashes because of no reason at all to the case where the plane crashes for some specific reason. In the latter case, when we move to a counterfactual world by positing

15. One might argue as follows against Lewis's more general claim. The recently shown failure of Lewis's own semantics for counterfactuals to properly exclude absurd cases of backtracking counterfactuals where the consequent is in the antecedent's past (Elga 2001; Pruss 2004b) strongly suggests that a semantics for counterfactuals will have to presuppose an asymmetry between the past and the future. One might further argue that there are no scientific asymmetries to sufficiently ground an asymmetry of such philosophical significance, and this might lead one to the Kantian view that the asymmetry in time supervenes on the asymmetry of causation: the past is just that region of time where (at least most of) the causes of present events are situated and the future is just that region of time where (at least most of) the effects of present events are situated. But if the asymmetry of time is presupposed in a semantics for counterfactuals, and the asymmetry between cause and effect is presupposed in the asymmetry of time, then at the pain of circularity one cannot analyze causation in terms of counterfactuals.

a new cause, we generate a case of overdetermination, and hence a case where the effect would still happen even without the new overdetermining cause. But in the case where the plane crashes because of no reason at all, the counterfactual world where a cause is posited is a world where there is only one cause, and hence the counterfactual that were the cause not to have occurred the effect would not have taken place is intact.

3.2.1.2. *What can have a cause, does*

We will see, however, that we can make a variant of the said argument into a valid and plausible argument for a CP. We will need a certain precise version of the observation that were the cause to have taken place, the effect would not have. This version says that if a state of affairs E is in fact caused by C , then E would not have occurred were no cause of E to exist:

$$(32) \quad (C \text{ causes } E) \Rightarrow ((\sim \exists D (D \text{ causes } E)) \Box \rightarrow E \text{ did not occur}),$$

where “ $p \Box \rightarrow q$ ” stands for “were p to hold, q would hold,” and where “ \Rightarrow ” marks entailment. We will also need a *might* operator: “ $p \Diamond \rightarrow q$ ” will stand for “were p to hold, q might hold.” The two operators are related as follows: $(p \Box \rightarrow q) \Leftrightarrow \sim(p \Diamond \rightarrow \sim q)$.

Premise (32) takes into account the possibility of overdetermination, where more than one state of affairs takes place, each of which is sufficient to cause E . It also takes into account the possibility that perhaps, were C not to have occurred, some other state of affairs D would have caused E . For instance, if members of some group are asked to volunteer to execute a traitor, then it might well be that Jones’s shooting the traitor causes the death of the traitor, although were Jones not to have shot the traitor, someone else would have, and hence the traitor would still have died.

I shall now argue that if E is a state of affairs that *can* have a cause, then E is a state of affairs that *does* have a cause. Since every step in the argument will be a conceptual truth if the argument works, it will follow that if E has a cause in one possible world, then in every world in which E takes place, E has a cause.

David Lewis proposed the following analysis of counterfactuals for a possible proposition p : $p \Box \rightarrow q$ holds providing there is a $(p \& q)$ -satisfying world that is more similar to the actual world than any $(p \& \sim q)$ -satisfying world is (Lewis 1986, sec. 1.3). While this analysis is, doubtless, not correct in all its details,¹⁶ the intuitive idea of a connection between counterfactuals and possible worlds should remain. When we try to see whether $p \Box \rightarrow q$ is true, we move to worlds relevantly similar to our world, but in which p holds, and see whether q holds in all such worlds. What features we must carry over from the actual world to the counterfactual world for it to count as “relevantly similar” is a difficult question. One might well say that, to the extent that p allows, one needs to carry over laws of nature and the past of p , while Lewis insists that “relevant similarity” has to do with being as similar as possible to the actual world. If, on the other hand, we think that there is some world relevantly similar to our world in which p holds but q does not, then we say that were p to hold, q might not hold.

16. See, for instance, Edgington (1995), Elga (2001), and Pruss (2004b, 2007). See also Section 3.2.2.1, below.

In modal logic, the Brouwer Axiom, which is entailed by S5, says that if a proposition p is actually true, then necessarily that proposition is possible. In terms of accessibility, this says that if we were to move to a world accessible from the actual world, the actual world would be accessible from that world: the accessibility relation is symmetric. But perhaps the best intuitive way to think about the Brouwer Axiom is to think of it as encapsulating the observation that in any nonactual situation we might consider, the events of the actual world remain relevant as alternative possibilities.

There is an analogue of this observation in the case of counterfactuals:

$$(33) \quad (q \ \& \ p \ \& \ M\sim p) \supset (\sim p \ \Box \rightarrow (p \ \Diamond \rightarrow q)),$$

where M indicates metaphysical possibility. If we actually have both p and q holding, and then move to a relevantly similar world w in which p does not hold, so as to evaluate a counterfactual with antecedent $\sim p$, the events of the actual world are going to be relevant for the evaluation of counterfactuals in w . Hence, if we ask in w what would happen were p to hold, we need to say that q might happen, since q in fact happens in the actual world.

Consider how (33) plays out in some paradigmatic cases. Suppose p claims that Jones freely chose to set fire to a barn and q claims that Jones was arrested. Then, were Jones not to have set fire to the barn, it would certainly have been true that were he to have set fire to the barn, he at least *might* have been arrested. In the case where p reports the occurrence or nonoccurrence of some punctual event in time, we can think of the space of possibilities as a branching structure. Were p not to have occurred, we would have gone on a different branch from the one we had in fact gone on. But were we to have gone on that branch, it would have been true that were p to have occurred, things *might* have gone just as they have *actually* gone. The fact that things *have* gone a certain way witnesses to the relevant possibility of them going this way. In this sense, (33) is an analogue to the Brouwer Axiom.

We also need two further obvious axioms dealing with counterfactuals, where “ \Rightarrow ” is entailment:

$$(34) \quad (p \Rightarrow q) \Rightarrow (p \ \Box \rightarrow q)$$

$$(35) \quad ((p \ \Box \rightarrow q) \ \& \ (p \ \Box \rightarrow \sim q)) \Rightarrow \sim Mp.$$

Entailment relations are stronger than counterfactual conditionals, and it cannot be that both q would hold were p to hold and $\sim q$ would happen were p to hold, unless p is itself impossible.

But now (33)–(35) imply that anything that can have a cause does have a cause. Let q be the true proposition that event E occurs, and suppose that E can have a cause. For a *reductio*, let p be the true proposition that there is nothing that causes E , that is, $\sim \exists D$ (D causes E). However, since E can have a cause, $M\sim p$. Thus, by the Brouwer analogue (33), we have:

$$(36) \quad \sim p \ \Box \rightarrow (p \ \Diamond \rightarrow q).$$

Let w be any possible world at which $\sim p$ holds. Then, w is a world at which E has a cause. Since nonexistent and nonoccurrent things can neither cause nor be caused, E occurs at w ,

as does a cause, call it C . Applying (32), we see that it is true at w that were no cause of E to have existed, E would not have occurred, that is, it is true at w that $p \Box \rightarrow \sim q$. Since this is true at every world at which E has a cause, that is, at every world at which $\sim p$ holds, it follows that:

$$(37) \quad \sim p \Rightarrow (p \Box \rightarrow \sim q).$$

But $p \Box \rightarrow \sim q$ is equivalent to $\sim(p \Diamond \rightarrow q)$. Thus, by (34):

$$(38) \quad \sim p \Box \rightarrow \sim(p \Diamond \rightarrow q).$$

By (35) and (36) it follows that $\sim Mp$. But p was assumed to be true, and true propositions are possible, and hence absurdly $\sim Mp$ and Mp .

Thus, the assumption for the *reductio* is false, and so p is false. Hence, there is a cause of E .

This is enough to show that Humeans are wrong to think that a brick could come into existence for no cause at all. For it is plain that there *can* be a cause of the state of affairs of a brick's coming into existence at t , and hence by the argument, there *is* such a cause.

It is plausible that for any physical kind of object identified *de dicto* and in a positive way, such as *a galaxy containing exactly n stars and having total mass M* , the state of affairs of that kind of object existing can have a cause, and hence does. Similarly, if we have a positive *de dicto* description D of all the physical stuff in the universe, it seems that it ought to be possible for there to be a cause of the state of affairs described by D . For instance, we could imagine D being satisfied in a larger world w^* where D describes a proper part P of the contents of w^* , a proper part that has a cause in another proper part Q of the contents of w^* , where Q might not actually exist in the actual world. That the description D is positive is important, since a nonpositive description could rule out the existence of Q , for example, by saying that there is nothing outside of what D describes.

From such considerations, we get a CP for physical objects, and by the same reasoning for causal chains of physical objects (surely there could be a cause of the whole chain). And this can yield a cosmological argument for a nonphysical being (see Section 4.2, below).

But let us slow down for a moment, and try for a more expansive result.

3.2.1.3. Which contingent states of affairs can have a cause?

If I could argue that all contingent states of affairs can have causes, then a CP for contingent states of affairs would follow from the conclusion of the previous section. However, there are several concerns about this idea.

A contingent state of affairs that contains a part that obtains necessarily perhaps cannot be expected to possibly have a cause. We do not expect the state of Socrates' having existed in a world without square circles to have a cause. Consider now the notion of a wholly contingent state of affairs, that is, one that has no component part that is necessary. Thus, the state of affairs of Socrates' having existed is wholly contingent, but the state of affairs of Socrates' having existed in a world that has no square circles is not wholly contingent. On a plausible set of mereological axioms for states of affairs, one can establish that

every contingent state of affairs S contains a maximal wholly contingent part S^* such that, necessarily, S obtains if and only if S^* does.¹⁷ We can now reasonably expect the possibility of causes for the wholly contingent substates.

A second problem is that if essentiality of origins holds, then that the state of affairs of Socrates' existing *can* have a cause *immediately* implies that the state of affairs *does* have a cause, since a cause of the state of affairs will presumably have to be a cause of Socrates, and hence will have to exist in every world where Socrates does. So if essentiality of origins holds, the atheist is likely not to grant that all contingent states of affairs *can* have causes. (Note that essentiality of origins could, in principle, hold for an uncaused being – such a being would then be *essentially* uncaused.)

Likewise, the smart atheist is likely not to grant that, in general, nonpositive contingent states of affairs can have causes, since granting that would yield the existence of a necessarily existent, causally efficacious being too quickly. For instance, the atheist is likely to think that there is a possible world that consists of only a single photon, and no necessarily existent, causally efficacious beings. But then, consider the state of affairs of there being one photon and nothing else. That state of affairs *cannot* have a cause, since that cause could not be the photon on pain of circularity and could not be anything else on pain of contradiction.

Consider now the following pair of claims:

- (39) If all wholly contingent, positive states of affairs that do not *de re* involve entities for which essentiality of origins holds have causes, then all wholly contingent, positive states of affairs have causes.
- (40) Every wholly contingent, positive state of affairs that does not *de re* involve contingent entities for which essentiality of origins holds *can* have a cause.

Claim (40) is an extension of the observation that states of affairs of the existence of *de dicto* described physical entities all can have causes. There is no reason to limit that observation to physical entities. If there can be a ghost that is 7-feet tall, then there can be a 7-foot tall ghost with a cause.

17. Koons (1997, Lemma 2) shows that every contingent state of affairs ("fact" in his terminology) contains a wholly contingent part. Let S^* be the aggregate of all wholly contingent parts of S . Note that S^* must itself be wholly contingent. For suppose, for a *reductio*, that it has a necessary part N . Then N has to overlap at least one of the wholly contingent parts of S , since every part of an aggregate must overlap at least one of the aggregated things. Thus, N will have a part in common with a wholly contingent part P of S . Thus, there will be a part Q that N and P will have in common. Any part of a necessary state of affairs is necessary (Koons 1997, Lemma 1), and so Q is necessary, contrary to the claim that P is wholly contingent, which is absurd. So S^* is wholly contingent. Moreover, it is a maximal, wholly contingent part of S . The only remaining question is whether it is the case that, necessarily, S obtains if and only if S^* does. One direction is clear: necessarily, if a state of affairs obtains, so do its parts, so, necessarily, if S obtains, so does S^* . For the converse, let N be the aggregate of all necessary parts of S . Clearly, N is itself necessary. Let S^{**} be the aggregate of N and S^* . Since N is necessary, necessarily if S^* obtains, so does S^{**} . If we can show that $S^{**} = S$, it will follow that, necessarily, if S^* obtains, so does S . For a *reductio*, suppose that S^{**} is not equal to S ; then there is a U that overlaps one but not the other (Koons 1997, Axiom 3). Since S^{**} is a part of S , it must be that U overlaps S but not S^{**} . But then, let V be a part that S and U have in common. If V is contingent, it will have a wholly contingent part (Koons 1997, Lemma 2), and this part will then be a part of S^{**} , and so V will overlap S^{**} , and hence so will U , which contradicts what was already said. So V must be necessary. But then V is a part of N , and hence overlaps S^{**} , and hence U overlaps S^{**} , which again contradicts what was already said.

I now argue for (39). Say that a kind of entity is “essentially originated” if essentiality of origins holds for that kind of entity. I claim that any contingent state of affairs S that does *de re* involve essentially originated entities has an associated state of affairs S^\dagger that does not. We obtain S^\dagger by taking a canonical description D of S in some ideal language and Ramseyfying it as follows. If the description D made reference to essentially originated entities e_1, e_2, \dots , so that $D = D(e_1, e_2, \dots)$, then let E_i be the maximally specific, positive description of e_i that does not involve the *de re* occurrence of any essentially originated entities (I shall assume there is a unique maximal description, since we can just conjoin any descriptions that meet the criteria other than maximality). A positive description is one such that the state of affairs of its being satisfied is a positive state of affairs. Descriptions that use words such as “unique” are not positive. And now we can Ramseyfy by letting D^\dagger be:

$$(41) \quad \exists x_1 \exists x_2 \dots (D(x_1, x_2, \dots) \ \& \ E_1(x_1) \ \& \ E_2(x_2) \ \& \ \dots).$$

Finally, let S^\dagger be the state of affairs described by D^\dagger .

Say that a world is “nice” if every pair of distinct essentially originated entities in that world differs in the maximally specific, positive, definitive descriptions that do not involve the *de re* occurrence of any essentially originated entities. On a plausible way of understanding Leibniz’s doctrine of identity of indiscernibles, any world for which identity of indiscernibles holds is a nice world. It is very plausible that our world is nice – it seems very likely that our world, in fact, lacks indiscernibles.

Now, plausibly, in a nice world, a cause C of S^\dagger is also going to be a cause of S . First of all, C will be the cause of everything in S except maybe of the numerical identities of the satisfiers of D being what they are since perhaps different individuals could play the same roles and satisfy $D(x_1, x_2, \dots)$. But, plausibly, there is no *further* step in causing particular individuals to occupy roles. Sophroniscus and Phainarete were causes of the existence of a philosopher executed by hemlock. There was nothing further that they did to cause the existence of *Socrates*. Moreover, if we include *all* of the causes of S^\dagger in C , then the numerical identities of the essentially originated individuals will also be taken care of, since it is plausible that for essentially originated entities, once their full causes have been given, their identity is thereby explained.

Therefore, (39) holds in nice worlds, and our world seems to be nice.

To amplify on the argument, observe that there are three plausible kinds of entities, with “entity” broadly understood, that might be essentially originated: substances, events, and some natural kinds. We might be a bit more worried about natural kinds. If *all* natural kinds were essentially originated, then the maximal descriptions introduced in the Ramseyfication would be unable to include reference to natural kinds, and that might make the descriptions not be specific enough to ensure niceness of our world. However, plausibly, only *some* natural kinds are essentially originated. The thesis of essentiality of origins for natural kinds is highly implausible for basic kinds such as *electron*, *star*, and *organism*. Suppose the first electron, star, or organism could have arisen from a different cause. It would perhaps be a numerically different electron, star, or organism (respectively) from the first one in our world, but it would nonetheless still be an electron, star, or organism (respectively). Let us suppose that electrons arose from collisions between certain other particles. Then even had these collisions happened earlier or later, and had different individuals been involved in the collisions, it would still be *electrons* that arose. The only natural

kinds for which essentiality of origins is plausible are biological taxa defined in evolutionary terms. It is somewhat plausible that had an animal with a different evolutionary history had the same DNA as the first horses, that animal would not have been a horse. But the limitation on descriptions that they do not involve taxa defined in evolutionary terms is not much of a limitation for our purposes – we can use phenotypic or genotypic descriptions instead, and if these are maximally specific, we will capture sufficient detail for the purposes of niceness.

And it seems that typical substances and events of our world can be captured by positive *de dicto* descriptions quite well. This may not capture their numerical identity, but it gives a maximal description strong enough that we would say that the cause of that description's being satisfied is the cause of the entity.

Now, given the conclusion of the previous section, together with (39) in nice worlds as well as (40), we get the claim that all wholly contingent, positive states of affairs in nice worlds have causes. But it is highly plausible that if the CP holds in nice worlds for wholly contingent, positive states of affairs, it also holds in non-nice worlds. The niceness condition is a version of the identity of indiscernibles. It would be odd indeed if there could not be a world consisting of a single uncaused brick, but there could be a world consisting of two indiscernible uncaused bricks. Hence, plausibly, the CP holds in all worlds for all wholly contingent, positive states of affairs.

One objection to this line of argument is that if libertarianism holds, it seems that states of affairs such as of George's freely choosing *A* are wholly contingent and positive, but cannot have a cause. One may worry whether *freely* choosing can be part of a positive state of affairs, since perhaps it entails the absence of external compulsion; but whether that worry is a good answer to the objection is unclear because many libertarians may accept that freedom is an *intrinsic* property of an action, and the absence of external compulsion is only relevant insofar as external compulsion would remove something from the intrinsic character of the action. However, the libertarian *can* say that the state of affairs of George freely choosing *A* has a cause. Maybe George is the cause. Or maybe George's making a choice between *A* and *B* while impressed by reasons *R* is the cause. Whether this cause provides a sufficient *explanation* of George's freely choosing *A* is a further question (see Section 2.3.2.3, above), but the mere claim of the existence of a cause is plausible.

3.2.2. Back to the PSR

If we do grant that causes always yield explanations, we can do even better than a CP on the given assumptions – we can get the PSR, supposing that the arguments previously outlined have successfully established the following claim:

- (42) Necessarily, all wholly contingent, positive states of affairs have causes.

This is not only a CP, but it seems to entail the necessary truth of a PSR for positive propositions, that is, propositions that report positive states of affairs. For if *p* is a proposition reporting a positive state of affairs *S*, we can let *S** be the maximal, wholly contingent part of *S*. Recall that, necessarily, *S* obtains if and only if *S** does, and by (42), *S** has a cause *C*. Hence, we can explain the obtaining of *S** as follows: *S* obtains because *S** has a cause *C*

and because contingent states of affairs obtain if and only if their maximal, wholly contingent substates do. If one objects that the noncontingent substates of S have not been explained, we can instead say that their obtaining is explained by the necessity of their obtaining or that they are self-explanatory, or stipulate that we are talking about explaining things *modulo* necessary truths, or perhaps hope that there is some way in which ultimately even the necessary truths all have explanations in terms of self-explanatory necessary truths (such as that each thing is identical with itself).

But now the necessary truth of a PSR for positive contingent propositions entails the necessary truth of a PSR for negative contingent propositions, where a negative proposition is the denial of a positive one. For if p is a negative contingent proposition, we can explain why p holds as follows: p holds because (a) there is nothing to explain why not- p holds, and (b) not- p is a positive contingent proposition, while (c) necessarily all positive contingent propositions that hold have explanations. This explanatory scheme is a variant of the scheme: E did not take place because no cause of E took place (see the discussion of the dog that did not bark in Section 2.3.2.2).

Finally, it is plausible that once we have explained all the positive contingent propositions and all the negative ones, then *all* contingent propositions will thereby have been explained, since their truth-values should supervene, in an explanation-preserving way, on the truth-values of the positive and negative ones.

This argument has an interesting consequence. I have argued (Pruss 2004a) that if we reject the PSR because we think that it has some counterexamples, such as the BCCF according to the van Inwagen argument, we should instead accept the restricted PSR (R-PSR):

(R-PSR) Every proposition that possibly has an explanation actually has an explanation.

Now, since the R-PSR claims to be a metaphysical principle, we should take it to be a necessary truth. We are now, however, in a position to see that the necessity of the R-PSR actually entails the PSR, *if* the arguments of the preceding section are successful. The argument is easy. The previous section shows, independently of any CP, that every wholly contingent, positive state of affairs can have a cause. Hence:

(43) Necessarily, every proposition reporting a wholly contingent, positive state of affairs can have an explanation.

The R-PSR then entails that it does have an explanation. But the same argument that shows that (42) entails the PSR also shows that (43) does so as well.

There is good news and bad news here for the cosmological arguer. The bad news is that if there are counterexamples to the PSR, there will be counterexamples to the R-PSR, and so the R-PSR does not make possible a cosmological argument that would work even if the PSR were false. The good news is that those whose intuitions lead them to accept that whatever is explainable is explained need to also accept the PSR for all contingent propositions.

If, on the other hand, the arguments of the preceding section are *not* successful, then it seems to be possible to accept the R-PSR without being committed to the PSR. In Section 4.4, I will show how to run a cosmological argument for a First Cause based only on the R-PSR.

3.2.2.1. *The Brouwer analogue*

The greatest difficulties in the given modal argument for the CP are with (33). The first difficulty is that (33) cannot be a conceptual truth on Lewis's semantics for counterfactuals. According to David Lewis, $p \Box \rightarrow q$ is true if and only if either p is necessarily false, or there is a $p \& q$ -world (i.e. a world where $p \& q$ holds) closer to the actual world than any $p \& \sim q$ -world (i.e. a world where $p \& \sim q$ holds).

Write Aw for a proposition true at w and only at w . We might take Aw to be the BCCF of w , or we might take Aw to be the proposition that w is actual. Let $q = Aw_0$, where w_0 is the actual world. Let w_1 be any other world, and let $p = \sim Aw_1$. Then, $q \& p \& M \sim p$ holds. Consider the consequent of (33). This says that there is a $\sim p$ -world w at which $p \Diamond \rightarrow q$ and which is closer than any $\sim p$ -world at which $\sim(p \Diamond \rightarrow q)$. In fact, there is only one $\sim p$ world, namely w_1 . Thus, the consequent of (33) says simply that $p \Diamond \rightarrow q$ holds at w_1 . Now, $p \Diamond \rightarrow q$ is equivalent to $\sim(p \Box \rightarrow \sim q)$. The proposition $p \Box \rightarrow \sim q$ holds at w_1 if and only if there is a $p \& \sim q$ -world that is closer to w_1 than any $p \& q$ -world is. Now, there is only one $p \& q$ -world, namely w_0 , and a $p \& \sim q$ -world is just a world different from w_0 and w_1 . Thus, $p \Box \rightarrow \sim q$ holds at w_1 if and only if there is a world different from w_0 and w_1 that is closer to w_1 than w_0 is. Thus, $\sim(p \Box \rightarrow \sim q)$ holds if and only if no other world is closer to w_1 than w_0 .

What we have shown is that if (33) holds, then for any world w_1 other than the actual world w_0 , the closest world to w_1 is w_0 . But this is most unlikely. Moreover, (33) is presented as a conceptual truth. If it is such, then the said argument should work in all possible worlds. It follows that for every pair of worlds w and w_1 , no other world is closer to w than w_1 . This is equivalent to the claim that one never has a chain of three distinct worlds w_1 , w_2 and w_3 such that w_2 is closer to w_1 than w_3 is. But surely there are such chains, and thus the consequence is absurd. Hence, the assumption that (33) is a conceptual truth leads to absurdity on Lewis's semantics.

However, all we need (33) for is the special case where q reports a wholly contingent, positive state of affairs and p reports the nonexistence of a state of affairs under a certain description (namely, under the description of being a cause of the state of affairs reported by q), and it might well be that in those cases (33) could still hold on Lewis's semantics. The given counterexample was generated using very special propositions – the proposition q was taken to be true at exactly one world and the proposition p was taken to be false at exactly one world. Ordinary language counterfactuals do not deal with such special propositions, and hence it might be that the intuitions supporting (33) do not require us to make (33) hold for these propositions, and hence these intuitions are not refuted in the relevant case by the counterexample.

This, however, is thin ice. One might perhaps more reasonably take (33) to entail a refutation of Lewis's semantics. In any case, Lewis's semantics are known to be flawed, especially when applied to propositions like the ones in the given counterexample. To see one flaw in them, suppose that w_0 is the actual world, and we have an infinite sequence of worlds w_1, w_2, w_3, \dots such that w_{n+1} is closer to the actual world than w_n is. For instance, these worlds could be just like the actual world except in the level of the background radiation in the universe, with this level approaching closer and closer to the actual level as n goes to infinity. Let p be the infinite disjunction of the Aw_n for $n > 0$. Fix any $n > 0$. On Lewis's semantics, we then have:

$$(44) \quad p \Box \rightarrow \sim Aw_n.$$

For w_{n+1} is a $p \& \sim Aw_n$ -world that is closer than any $p \& Aw_n$ -world, since there is only one $p \& Aw_n$ -world, namely w_n , and w_{n+1} is closer than it. This implies that it is true for every disjunct of p that were p true, that disjunct would be false! But, surely, there has to be some disjunct of p such that were p true, that disjunct might also be true.

Like the counterexample to (33), this counterexample deals with propositions specified as true at a small (in the case of p here, infinite, but still only countably infinite and hence much “smaller” than the collection of possible worlds, which is not only not countably infinite but is not even a set¹⁸) set of worlds. This shows that there is something wrong with Lewis’s semantics, either in general or in handling such propositions (see also Pruss 2007).

To see even more clearly, although making use of a slightly stronger assumption about closeness series, that there is a commonality between a problem with Lewis’s semantics and the Lewisian counterexample to (33), suppose the following principle of density: for any nonactual world w , there is a nonactual world w^* closer to the actual world than w is. This should at least be an epistemic possibility: our semantics for counterfactuals should not rule it out. Let w_0 be the actual world and put $p = \sim Aw_0$. Then, by the principle of density, on Lewis’s semantics, there is no possible world w such that were p true, w might be actual, that is, such that $p \diamond \rightarrow Aw$. To see this, suppose that we are given a w . First, note that it is hopeless to start with the case where w is w_0 since p and Aw_0 are logically incompatible. Next, observe that if w is not w_0 , then we have $p \square \rightarrow \sim Aw$. For let w^* be any world closer than w . Then, w^* is a $p \& \sim Aw$ -world that is closer than any $p \& Aw$ world, there being only one of the latter, namely w . But if we have $p \square \rightarrow \sim Aw$ and p is possible, then we do not have $p \diamond \rightarrow Aw$.

But, intuitively, if p is possible, then there is *some* world which is such that it might be actual were p to hold. Lewis’s semantics fails because of its incompatibility with this claim, on the aforementioned not implausible principle of density, which should not be ruled out of court by a semantics of possible worlds. Note further that the failure here is precisely a failure in the case of a might-conditional $p \diamond \rightarrow q$ with p of the form $\sim Aw_1$ and q of the form Aw_2 , which is precisely the kind of might-conditional that appeared in the analysis of the putative counterexample to (33). Lewis’s semantics makes too few might-conditionals of this sort true, and it is precisely through failing to make a might-conditional of this sort true that it gave a counterexample to (33).

Thus, rather than having run my argument within Lewisian possible worlds semantics, it was run on an intuitive understanding of counterfactuals, which intuitions do support (33). It would be nice to have a complete satisfactory semantics for counterfactuals. Lewisian semantics are sometimes indeed helpful: they are an appropriate model in many cases. But as we have seen, they do not always work. Other forms of semantics meet with other difficulties. We may, at least for now, be stuck with a more intuitive approach.

If we want some more precision here, we might speak as follows. To evaluate $p \square \rightarrow q$ and $p \diamond \rightarrow q$ at a world w , we need to look at some set $R(w,p,q)$ of “ q -relevant p -worlds relative to w ” and check whether q holds at none, some, or all of these. If q holds at all of them, then $p \square \rightarrow q$ and $p \diamond \rightarrow q$. If it holds at none of them, then neither conditional is true. If it holds at some but not all of them, then $\sim(p \square \rightarrow q)$ and $p \diamond \rightarrow q$. The difficulty is

18. Pruss (2001).

with specifying the q -relevant p -worlds. Proposition (33) then follows from the claim that the actual world is a q -relevant p -world relative to every world w which is a relevant $\sim p$ -world relative to the actual world. This is plausible, and somewhat analogous to the Brouwer Axiom. However, this does not let us embed the discussion in a precise semantics because we do not have an account of what $R(w, p, q)$ is.

David Manley (2002) has come up with the following apparent counterexample to (33), which I modify slightly. Suppose our soccer team wins 20 to 0. Then, it is true that the team won overwhelmingly in the actual world w_0 . What would have happened had our team not won? Presumably, the score would have been rather different, say 20 to 20, or 0 to 5, or something like that. Suppose the score is one of these – that we are in a possible world w_1 where our team has lost. Then, it is *not* true that were our team to have won, it would have won overwhelmingly. If our team in fact failed to win, as at w_1 , then worlds where the team wins *overwhelmingly* are much more distant from our world than worlds where it wins by a bit. Thus, it is true at w_1 to say that were our team to have won, it would have won by a tiny amount. Putting this together, we conclude that were our team not to have won, then were it to have won, it would have won by a tiny amount. But this is incompatible with (33), which claims that were our team not to have won, then were it to have won, it *might* have won by a tiny amount.

But this account of the situation also relies on David Lewis's semantics, and again does so in a context in which Lewis's semantics *fail*. For, by this reasoning, if we are in a world where our team has not won, then we should say that were it to have won, it would have won by exactly one point. But this need not be true. Perhaps were it to get ahead by a point at some point in the game, then the other team would have become disheartened and lost by more. We can even more clearly see the problem in the Lewisian reasoning if we substitute a game very much like soccer except that its scores can take on any real value: perhaps instead of a flat one point for a goal, one gets a real-valued additive score depending on how close to the middle of a goal one hits. Then, by this reasoning, were our team not to have won, it would be true that were it to have won, it would have won by no more than 1/10 of a point. Worlds where one wins by no more than 1/10 of a point are closer than worlds where one wins by more than that. But this reasoning is perfectly general, and the "1/10" can be replaced by any positive number, no matter how tiny. But this is absurd. It is absurd to suppose that were our team not to have won, it would be true that were it to have won, it would have won by no more than $10^{-1,000}$ points.¹⁹

3.3. An objection: causing the causing

While the van Inwagen objection applies specifically to the PSR, there is a related objection to CPs. Suppose that our CP applies to all wholly contingent, concrete, contingent states of affairs. Suppose that state of affairs C causes state of affairs E . It may be that C is necessary (e.g. C might be God's existing and having such and such values), and thus one cannot get a regress by asking for C 's cause, but there is a different move available. We can form a third state of affairs, the state of affairs C_1 of C 's causing E , and then can ask what causes C_1 . It seems plausible that if C is wholly contingent, then so is C_1 . The object in asking this question is to generate a vicious regress. Once we gave the cause of C_1 , we would form the state of affairs of that cause's causing C and so on.

19. This is similar to the coat thief example cited in Edgington (1995).

This problem looks formidable, but the real difficulty lies in choosing from the abundance of possible solutions. For instance, the literature contains Koons's solution (Koons 1997) on which C_1 is not a *further* state of affairs. Instead, C_1 consists simply in the mereological sum of C and E . Or one might argue that C_1 is only partly contingent, since in the case that interests us C is necessary and in some way enters into C_1 , and its wholly contingent part C_1^* might not actually be distinct from E . Further, the scholastics apparently liked to say that the actuality of the cause *qua* cause is the effect. A good translation of "the actuality of the cause *qua* cause" may be "the cause's causing," so if they are right, then the cause's causing may not actually be distinct from E .

A different solution is to allow for a regress but to claim that it is not vicious. Not all regresses are vicious. If p is true, then it is true that p is true and so on. There does not appear to be any philosophical consensus on which regresses are vicious. A plausible suggestion is that regresses are vicious if they involve a dependence, whether explanatory, causal, or grounding. Thus, we should reject a theory of truth on which what makes a proposition true is that it is true that it is true, for then the truth regress would be a grounding regress. But as long one does not take such a theory of truth, the truth regress is not vicious.

Now, one might think that the causal regress here is an objectionable causal or explanatory dependence regress. Why did E happen? Because of C . But why did C cause E ? Because of C_1 . But why did C_1 cause C 's causing E ? Because of C_2 . And so on. But it is mistaken to think that viciousness always occurs. The following seems a coherent account. *Ex hypothesi*, the cause of E is C . In particular, the cause of E is not C 's causing E , at least not if that is an event distinct from C (and if it is not an event distinct from C , the problem disappears). In the process from C to E , C 's causing E transpires. It is not the case that C 's causing E is more ultimate causally than C . In fact, one might reasonably say that C causes C 's causing E , and C causes C 's causing C 's causing E , and so on. These epi-events are not a part of the causal explanation of E , however.

Granted, we may sometimes explain that E happened not just because of C but because of C 's causing E . We might, however, question whether this is always a perspicacious expression of the explanation – recall the fact that we do not say that "He died because he died of being stabbed" (see Section 2.3.2.2, above).

Given that the question is most interesting in the ultimate case of causation by a First Cause, and that the cosmological arguer thinks the ultimate case of causation is a case of agency, it might be good to consider how, on one interpretation of the phrase " C 's causing E ," this looks for agency. Suppose that x does A on account of reason R . Then the cause of A is x -who-appreciates- R , or maybe x 's appreciating R . We can now ask why it is that R was what moved x , or why it is that x who appreciates R does A . After all, x may well also appreciate other reasons, reasons in favor of some other action, say B . We might ask why R is the reason that in fact moves x , and one way of putting this question is to ask for the cause of x -who-appreciates- R 's causing A .

This is at least sometimes a substantive question to which a substantive answer is possible. Jones joined the Antarctic expedition because he appreciated the value of scientific discovery. But why was it that his appreciation of the value of scientific discovery moved him to join the Antarctic expedition, when instead his appreciation of the value of a congenial climate might have moved him to stay in Kansas? Perhaps it was because of a higher-order reason. Maybe he judges warmth to be a private bodily good and scientific discovery to be a nonprivate intellectual good, and he appreciates the value of sacrificing private bodily goods to nonprivate intellectual ones. In that case, there is a substantive

answer to the substantive question: the value of scientific discovery moved him to join the Antarctic expedition because of his appreciation of sacrificing primitive bodily goods to nonprivate intellectual ones. And, of course, there might be a further question about why the second-order reason moved Jones.

But it is clear that such explanation will have to come to an end. We do not in fact have an infinite chain of reasons for every action, finite creatures as we are. And in fact, it is plausible that sometimes x does A for R , and there is no further answer to the question why it was that x did A for R , why it was that x -who-appreciated- R did A , beyond the fact that x appreciated R . Kant's ideal of being moved by the respect for duty provides an example of this. Why it is that Kenya, a perfect Kantian agent, kept her promise? It is because of her respect for duty. And why was Kenya moved by her respect for duty? Because her respect for duty requires of her that she be moved by reasons of duty. This sounds circular, but in fact we can see it as a case where respect for duty not only moves the Kantian to keep her promise but also to keep her promise out of respect for duty. This would be a case like the one I suggested earlier, where C causes C 's causing E : respect for duty causes her respect of duty to be the cause of her action. No vicious regress ensues here, since there is nothing further to explain about why respect for duty moved Kenya beyond her respect for duty. Respect for duty causes her respect of duty to be the cause of her concretely action, precisely in and through respect causing that action in the absence of other causes.

Nor is Kantianism the only kind of case where this happens. Suppose George does something out of love for his child. It seems quite plausible that not only is love moving him to benefit the child but also love is moving him to benefit the child out of love. So cases of agency can, plausibly, have a structure that evades the Regress Problem.

There is a final answer to our Regress Problem, which I think is the best, and it is to identify C 's causing E with C 's causal activity. Now, in some cases, we can ask for a further cause of C 's causal activity – we can ask what moved the mover. But what if C 's causal activity is a necessary being? Then C 's causal activity does not itself fall within the purview of the CP. Those who accept divine simplicity, with its contention that God is God's activity, will likely accept this; if God is simple and a necessary being, his activity must itself be necessary, being identical with him.

One might say that this cannot be right, at least not in the case closest to the cosmological arguer's heart, the case of God's causing this universe. For if God's causal activity is necessary, then God's causing this universe is necessary, and hence this universe is a necessary being, which is absurd, besides being contrary to the assumptions of typical cosmological arguments. But this objection commits a *de re/de dicto* fallacy. Consider the argument written out:

- (45) C_1 is God's causal activity and is a necessary being. (Premise)
- (46) C_1 is God's causing E . (Premise)
- (47) Therefore, God's causing E is a necessary being.
- (48) Therefore, God necessarily causes E .

The fallacy is in the last step, which has essentially the following logical form:

- (49) The F is a necessary being.
- (50) Therefore, necessarily, the F exists.

But this is fallacious if “the F ” in (50) is read *de dicto*, as it must be in the case where F is “God’s causing E .” The number of eyes of the tallest person is, let us suppose, the number two and, let us also suppose, that the number two is a necessary Platonic being. But it is incorrect to conclude that, necessarily, the number of eyes of the tallest person exists, since that would entail the falsehood that necessarily there is a tallest person (it could be that no person has a height or that there is a tie).

The inference from (49) to (50) requires that F ness be an essential property of the F . It is not an essential property of the number two that it be the number of eyes of the tallest person, and hence the inference fails in that case. Similarly, in the God case, the argument is only going to work if it is an essential property of God’s activity that it be the same as God’s causing E . But this proponents of divine simplicity should deny. They should instead insist that the same activity would count as God’s causing E in those worlds where God causes E and as God’s causing F in those worlds where God causes F . A very imperfect analogy for this is that the same act of writing down a sequence of numbers is, in some worlds, the filling out of a winning lottery ticket and in others the filling out of a losing lottery ticket. The reason for the imperfection in the example is that in the lottery case, it does not depend on the act which lottery tickets are drawn, but everything, in some way, depends on God’s act. But it should be no surprise if there is no close analogue to doctrines coming from divine simplicity.

The reason for the multitude of responses to the objection is that it is just not clear what “ C ’s causing E ” picks out. It might pick out the mereological sum of C and E . It might pick out E . It might pick out something causally posterior to C . To some extent the details will depend on fine details in the theory of causation, and to some extent it may simply be a matter of choosing what one means by the ambiguous phrase “ C ’s causing E .”

4. Toward a First Cause

4.1. *Overcoming the Regress and Taxicab Problems*

I have defended three principles, each of which is sufficient to overcome the Glendower Problem: a PSR for contingent propositions, a CP for wholly contingent positive states of affairs, and a CP not just for individual events/substances but for chains of events/substances. Now the question is whether they are sufficient to overcome the Regress and Taxicab problems and yield the existence of a First Cause. I shall argue in the positive, thereby giving versions of three cosmological arguments by Leibniz, Koons (1997), and Meyer (1987), respectively. I shall follow these three by giving a fourth argument, due to White (1979), and another argument based on the ideas in Pruss (2004a), both of which arguments are based on restrictions of the PSR.

4.1.1. The PSR

4.1.1.1. *The basic argument*

Consider once again the BCCF p , which was the conjunction of all contingently true propositions, perhaps with truth-functional redundancy removed. By the PSR, p has an

explanation, call it q . What is q like? There are two general options. Either q is necessary or q is contingent. If q is contingent, then it is contained in p , and since q explains p , it follows that q is self-explanatory. Thus, q must be necessary or else contingent and self-explanatory. (Here we are, of course, retracing a part of the van Inwagen argument from Section 2.3.2.1.)

Next, observe that it is plausible that contingent existential propositions ultimately can only be explained causally. Since p includes many contingent existential propositions, q must state the existence of one or more causes. If these causes are all contingent substances or events, then the existence of these causes will be among the contingent existential propositions in p that are to be explained. But given a set of contingent entities, these entities can neither collectively nor individually causally explain their own existence. Nothing can be a cause of itself, *pace* Descartes. The existence of a cause is explanatorily prior to the existence of the effect, but nothing can be explanatorily prior to itself.

So the cause must be something necessary, presumably either a necessarily existing substance or a necessarily occurring event. Plausibly, there can be no events without substances – events are what happens to substances. A necessarily existing event happens to a necessarily existing substance. So we do get to a necessarily existing substance.

Moreover, as far as we can tell, there are three ways that something can be explained. First, one can have a conceptual explanation that explains one fact in terms of a conceptually connected fact entailing it in an explanatorily relevant way, as when we say Pat's action was wrong because it was the breaking of a promise, or that a knife is hot because its molecules have high kinetic energy. A conceptual explanation of a contingent proposition will itself involve a contingent proposition, and a nonconceptual explanation will be needed. We can say that the Queen of England exists because Elizabeth of Windsor is the Queen of England, a conceptual explanation, but the explanation of the existence of Elizabeth of Windsor will involve the gametes of the Duke and Duchess of York, and this will not be a conceptual explanation. Since q is the ultimate explanation of all contingent propositions, it will not be a conceptual explanation, except perhaps in part.

Second, one can explain things scientifically by citing laws of nature and initial conditions. Now, on some accounts of laws of nature, the laws of nature are contingent and non-self-explanatory. They will thus have to enter into the explanandum p , but not the explanans q . Moreover, the most plausible account of laws of nature that makes them necessary grounds them in the essences of natural objects. But natural objects are contingent. Hence, even though the laws of nature will be necessary, which laws are *applicable* to a given situation will depend on the contingent question of which contingently existing natural objects are involved in the situation. The ultimate explanation q cannot involve laws grounded in the essences of contingently existing natural objects, since q explains the existence of contingently existing natural objects. Moreover, the initial conditions cited in scientific explanations are contingent and non-self-explanatory. But q is either necessary or contingent and self-explanatory. So q cannot be a scientific explanation.

The last kind of explanation we know of involves the causal activity of an agent or, more generally, a substance. The substance will have to be a necessary being, or else it will absurdly be a *causa sui*, something that causally explains its own existence (since it explains all of the BCCF). Hence, the ultimate explanation involves one or more causally efficacious necessary beings, whom we may call First Causes. Were it not for the Gap Problem, we could now say *et hoc dicimus deum*.

4.1.1.2. *Objection 1: explanations in terms of a principle*

The main objection here is that perhaps there is some *further* mode of explanation, one that is not conceptual, scientific, or agentive. The main actually proposed alternative is explanation in terms of a metaphysical principle. This principle would have to be different in kind from run-of-the-mill metaphysical principles such as the Principle of Identity of Indiscernibles or the Principle of Impossibility of Circular Causation: it must be a principle capable of explaining the existence of the apparently contingent denizens of our world.

The best candidate principle is due to John Leslie (2001) and Nicholas Rescher (2000). I will discuss Rescher's formulation here. The idea is to explain the BCCF in terms of the Principle of Optimality: of metaphysical necessity, the best narrowly logically possible world is actual. However, Rescher's suggestion is one that the defender of the cosmological argument need not worry about too much, since it is plausible that the best narrowly logically possible world is a world that contains God, considered as a maximally great being. Rescher himself thinks this. And so, in any case, we get the existence of God, albeit in a somewhat more roundabout way. Leslie does not agree – he opts instead for an infinity of divine knowers. I suspect Rescher is right in supposing a single deity – any particular number of deities other than one would seem *ad hoc* vis-à-vis optimality, while a world with a single deity has an elegant and valuable unity to it. This is true whether the number is finite, say 117, or infinite, say \aleph_{117} . And if the number is Cantorian “absolute infinity,” then it does not seem as though one can make any sense of it.

Moreover, it is plausible (although Rescher denies it) that principles need to be made true by something, and this something must have being. A principle cannot by itself pull beings into existence out of a metaphysical magic hat, since a principle itself must be true of something and true in virtue of something.

4.1.1.3. *Objection 2: can we form the BCCF?*

Forming the BCCF may present set-theoretic concerns. Not every conjunction of propositions makes sense, as was shown by Davey and Clifton (2001). Modifying their construction slightly, let p be conjunction of all true propositions that do not contain themselves as proper subformulas. Then p is true. Let q be the proposition *that p is true*. Now, either q is a proper subformula of itself or not, and in either case a contradiction ensues. For if q is a proper subformula of itself, then it is not a conjunct of p . But the only way q could be a proper subformula of itself is if it is a subformula of p , since all proper subformulas of q are subformulas of p . Since q is not a conjunction, the only way it can be a subformula of p is if it is a subformula of one of the conjuncts of p . Now, q is not a conjunct of p , as we said. Hence, q must be a proper subformula of one of the conjuncts of p , say of p_1 . But p is a proper subformula of q , and p_1 is a subformula of p , so it follows that p_1 is a proper subformula of itself, and hence not a conjunct of p , contrary to the assumption. Suppose on the contrary that q is not a proper subformula of itself. Then it is a conjunct of p , and hence a proper subformula of itself, and absurdity ensues again.

This argument is a challenge: if some conjunctions do not make sense, how do we know that the BCCF makes sense? One way to meet the challenge is to try to shift the burden of proof. A conjunction of propositions should be assumed to make sense unless it is proved not to.

Alternately, one might use the strategy of Gale and Pruss (2002) here. Replace the BCCF by the BCCF*, which is the conjunction of the following:

- (a) all true contingent atomic propositions,
- (b) a “that’s all clause” that says that any true contingent atomic proposition is one of these ones (this clause will involve an infinite disjunction such as in: “for all p , if p is a true contingent atomic proposition, then p is a_1 or a_2 or . . .”),
- (c) all true propositions appearing in the explananda of contingent atomic propositions or of conjunctions thereof,
- (d) all true basic propositions reporting causal relations,
- (e) a “that’s all clause” that says that all the actual explanatory relations supervene on the facts reported in the above conjuncts. (Gale & Pruss 2002, p. 95)

Here, “basic propositions” might be taken to be ones that are not true in virtue of some more basic propositions’ being true, in the way in which the proposition that George is human or rhino is true in virtue of the more basic proposition that he is human. Plausibly, the truth of the BCCF supervenes on that of the BCCF*, and we could probably run our cosmological argument with the BCCF* in place of the BCCF.

Also, some PSR-based cosmological arguments are not subject to this objection. For instance, we could ask why there are any contingent beings. It is highly plausible that *that there are contingent beings* is itself a contingent proposition. For if it were a necessary proposition that there are contingent beings, then we would have odd necessary truths such as that, necessarily, if there are no contingent nonunicorns, then there are contingent unicorns. Moreover, the explanation of why there are contingent beings cannot involve the causal efficacy of contingent beings. But, plausibly, an existential proposition can only be explained by citing the causal efficacy of something, and hence of a necessary being.

Note that if there is no way of forming something relevantly like the BCCF, then the van Inwagen objection, which is the main objection to the PSR, fails. So it could be a service to a defender of the PSR if nothing like the BCCF could be formed, although harder work would be needed then for running the cosmological argument, perhaps along the lines of the Gale and Pruss (2002) strategy.

4.1.1.4. *Objection 3: the Hume–Edwards–Campbell principle*

Hume states:

Did I show you the particular causes of each individual in a collection of twenty particles of matter, I should think it very unreasonable, should you afterwards ask me, what was the cause of the whole twenty. This is sufficiently explained in explaining the cause of the parts. (Hume 1907, p. 120)

Paul Edwards (1959) illustrates this with the case of five Inuit on a street corner in New York. If for each Inuit we gave an explanation of why he or she is there, we would thereby have explained why they are all there.

We can generalize this to the Hume–Edwards Principle (HEP):

- (51) (HEP) In explaining every conjunct of a proposition, one has explained the whole proposition.

If the HEP is true, the PSR-based cosmological argument can be blocked. The objector can simply suggest that one contingent proposition is explained by a second, and the second by a third, and so on *ad infinitum*, and thereby the whole BCCF is explained.

But the HEP is false. The first objection to HEP is that it does not take into account the fact that there can be more to explaining the conjunction than explaining the conjuncts. If there were a hundred Inuit on a street corner in New York, individual explanations of each one's presence would miss the point of explaining why there are a *hundred* Inuit all there. There is a coincidence to be explained. This kind of objection has been pressed, for example, by Gale (1999, p. 254).

This response is, however, insufficient. First of all, while *sometimes* explaining the conjuncts is unsatisfactory as an explanation of the whole, sometimes it is quite satisfactory. If there are two Inuit there on the corner, then to say that one is there to give a paper on the cosmological argument at a conference in the hotel at that corner and the other is there because New York City winter is preferable to Iqaluit winter is likely to be a fine explanation of the conjunction. The cosmological arguer might then be required to show that the BCCF is in fact a case where an explanation of the conjuncts is insufficient for an explanation of the whole. There are issues as to onus of proof here, but it is better for cosmological arguer to sidestep them if possible.

The second problem with this response is that while it provides a counterexample to the HEP, there is a weaker version of the HEP due to Campbell (1996) that these sorts of examples do not address. Campbell agrees that sometimes we need to do more than explain the parts to explain the whole. Indeed, there may be a further story, an Inuit conspiracy, say. However, it *might* be that the individual explanations of the parts are the whole story and are the explanation of the whole. If there are a hundred Inuit on the street corner, then it seems *likely* that there is a further explanation beyond individual ones. But there *might not be*. It could be that it is just a coincidence, and the whole is correctly explained solely in terms of the individual parts.

This suggests the following Hume–Edwards–Campbell Principle (HECP):

(HECP) For any proposition p such that one has explained every conjunct of a proposition, one might have thereby explained the whole.

We can take the “might” as epistemic possibility absent evidence of a further explanation. The HECP is sufficient for blocking the PSR-based cosmological argument. For the defender of the HECP can say we *might* have explained the whole of the BCCF by explaining one proposition in terms of another *ad infinitum*, and then the onus would be on the cosmological arguer to provide evidence that there is a further explanation of the BCCF. But if one can provide such evidence, one probably does not need the PSR-based cosmological argument.

However, the HECP is also subject to counterexample, and any counterexample to the HECP will automatically be a counterexample to the stronger HEP. Perhaps the simplest counterexample is the following (Pruss 1998). At noon, a cannonball is not in motion, and then it starts to fly. The cannonball flies a long way, landing at 12:01 p.m. Thus, the cannonball is in flight between 12:00 noon and 12:01 p.m., in both cases noninclusive.

Let p_t be a proposition reporting the state of the cannonball (linear and angular moment, orientation, position, etc.) at time t . Let p be a conjunction of p_t over the range

12:00 < t < 12:01. I now claim that p has not been explained unless we say what caused the whole flight of the cannonball, for example, by citing a cannon being fired. This seems clear. If Hume is right and it is possible for causeless things to happen, then it could be that there is no cause of the whole flight. But that is just a case where p has not been explained. To claim that there was no cause of the flight of the cannonball but we have explained the flight anyway would be sophistry.

But if the HECF is true, then there might be an explanation of p without reference to any cause of the flight of the cannonball. For take any conjunct p_i of p . Since 12:00 < t , we can choose a time t^* such that 12:00 < t^* < t . Then, p_i is explained by p_{t^*} together with the laws of nature and the relevant environmental conditions, not including any cause of the whole flight itself. By the HECF we *might* have explained all of p by giving these explanations. Hence, by the HECF we might have explained the flight of a cannonball without giving a cause to it. But that is absurd.²⁰

Perhaps the defender of the HECF might say that it is relevant here that there was a time before the times described by the p_i , namely *noon*, and the existence of this time is what provides us with evidence that there is a further explanation. But that is mistaken, for if there were no noon – if time started with an open interval, open at noon – that would not make the given explanation of p in terms of its conjuncts, the laws, and the environmental conditions any better an explanation.

The HECF is, thus, false. But there is some truth to it. In the cannonball case, we had an infinite regress where each explanation involved another conjunct of the same proposition. Such a situation is bound to involve a vicious regress. The HEP and HECF fail in the case where the individual explanations combine in a viciously regressive manner. They also fail in cases where the individual explanations combine in a circular manner. If it should ever make sense to say that Bob is at the party because Jenny is and that Jenny is at the party because Bob is, that would not explain why it is that Bob and Jenny are at the party. A further explanation is called for (e.g. in terms of the party being in honor of George's birthday, and George being a friend of Bob), and if, contrary to the PSR, it is lacking, then that Bob and Jenny are at the party is unexplained. Likewise, if we could explain that Martha constructed a time machine because of instructions given by her future self, and that Martha's future self gave her instructions because she had the time machine to base the instructions on and travel back in time with, that would not explain the whole causal loop. How an explanation of the whole loop could be given is not clear – maybe God could will it into existence – but without an explanation going beyond the two circular causes, the loop as a whole is unexplained.

Thus, not only is the HECF false in general, but it is false precisely in the kind of cases to which Hume, Edwards, and Campbell want to apply it: it is false in the case of an infinite regress of explanations, each in terms of the next.

Interestingly, the HEP *may* be true in finite cases where there is no circularity. Indeed, if we can explain each of the 20 particles, and if there is no circularity, we will at some point in our explanation have cited at least one cause beyond the 20 particles, and have given a fine explanation of all 20. Granted, there may be worries about coincidence, about

20. This counterexample to HECF makes use of the density of time. Many of the best reasons for doubting the density of time involve the sorts of considerations about traversing infinity that are at the heart of conceptual versions of the *kalam* argument (the Grim Reaper Paradox is the main exception to this). Thus, some of the people who object to the density of time assumption here will therefore have to worry about the *kalam* argument, which is discussed in Chapter 3 of this book.

whether the *whole* has been explained, but these worries can, I think, be skirted. For if I have explained individually why each of a hundred Inuit is at the street corner, and done so in a noncircular manner, then I have explained why there are a hundred Inuit there. Now my explanation may not be the best possible. It may not tell me everything there is to know about how the hundred Inuit got there, for example, through the agency of some clever manipulator who wanted to get Inuit extras for a film, but it is *an* explanation. But that is because the explanation goes beyond the contingent hundred Inuit, if circularity is avoided.

At this point it is also worth noting that anybody who accepts the possibility that an infinite regress of contingent propositions be explained by the regressive explanatory relations between these propositions is also committed to allowing, absurdly, that we can have two distinct propositions r and q such that r has the resources to explain q , q has the resources to explain r , and where the explanatory relations here also have the resources to explain the conjunction $r \& q$. For suppose that we have a regress of explanations: p_1 explained by p_2 , p_2 by p_3 , and so on. Let r be the conjunction of all the odd-numbered p_i , and let q be the conjunction of all the even-numbered p_i . Since the conjunction of all p_i has been explained in terms of the individual explanatory relations, by the same token we can say that r is explained by means of the explanatory resources in q since each conjunct of r is explained by a conjunct of q (p_7 being explained by p_8 and so on), and q is explained by means of the resources in r since each conjunct of q is explained by a conjunct of r (p_8 being explained by p_9 and so on). And these relations, in virtue of which r has the resources to explain q and q to explain r , also suffice to explain $r \& q$.

Quentin Smith has argued that the universe can cause itself to exist, either via an infinite regress or a circle of causes. However, while he has claimed that such a causal claim would provide an answer to the question “why does the universe exist?” (Smith 1999, p. 136), he appears to have provided no compelling argument for that conclusion. Hence, it is possible to grant Smith that the universe can cause itself to exist via an infinite regress or a circle of causes while denying that this can provide an *explanation* of why the universe exists. Granted, in normal cases of causation, to cite the cause is to give an explanation. However, when we say that the universe “caused itself to exist” via a regress or a circle of causes, we are surely using the verb “to cause” in a derivative sense. The only *real* causation going on is the causes between the items making up the regress or the circle. If we have such a story, maybe we can say that the universe caused itself to exist in a derivative sense, perhaps making use of some principle that if each part of A is caused by a part of B , then in virtue of this we can say that A is caused by B , but this derivative kind of causation is not the sort of causation which must give rise to an explanation. For, in fact, regresses and circles do not explain the whole.²¹

21. It is worth noting that Smith’s arguments for circular causation are weak. One argument relies on quantum entanglement. Smith seems to think that simultaneous but spatially distant measurements of entangled states could involve mutual causation between the states. But that is far from clear; circular causation is only one possible interpretation of the data, and how good an interpretation it is depends in large part precisely on the question whether circular causation is possible. Smith’s more classical example in terms of Newtonian gravitation simply fails: “There is an instantaneous gravitational attraction between two moving bodies at the instant t . Each body’s infinitesimal state of motion at the instant t is an effect of an instantaneous gravitational force exerted by the other body at the instant t . In this case, the infinitesimal motion of the first body is an effect of an instantaneous gravitational force exerted by the second body, and the infinitesimal motion of the second body is an effect of an instantaneous gravitational force of the first body. This is a case of the existence of a state $S1$ being caused by another state $S2$, with the existence of $S2$ being simultaneously caused by $S1$ ” (Smith 1999, pp. 579–80). This is

4.1.1.5. *Objection 4: can there be a necessarily existent, causally efficacious being?*

Perhaps a necessary being is impossible. Abstracta such as propositions and numbers, however, furnish a quick counterexample to this for many philosophers. However, one might argue further that there cannot be a *causally efficacious* necessary being, whereas the unproblematic abstracta such as propositions and numbers are causally inefficacious.

A radical response to this is to question the dogma that propositions and numbers are causally inefficacious. Why should they be? Plato's Form of the Good looks much like one of the abstracta, but we see it in the middle dialogues as explanatorily efficacious, with the *Republic* analogizing its role to that of the sun in producing life. It might seem like a category mistake to talk of a proposition or a number as causing anything, but why should it be? Admittedly, propositions and numbers are often taken not to be spatiotemporal. But whence comes the notion that to be a cause one must be spatiotemporal? If we agree with Newton against Leibniz that action at a distance is at least a metaphysical possibility, although present physics may not support it as an actuality, the pressure to see spatiality or even spatiotemporality as such as essential to causality is apt to dissipate – the restriction requiring spatiotemporal relatedness between causal relata is just as unwarranted as the restriction requiring physical contact.

Admittedly, a Humean account of causation on which causation is nothing but constant conjunction only works for things in time, since the Humean distinguishes the cause from the effect by temporal priority. But unless we are dogmatically beholden to this Humean account, to an extent that makes us dogmatically *a priori* deny the existence of deities and other nonspatiotemporal causally efficacious beings, this should not worry us.

Moreover, there is actually some reason to suppose that propositions and numbers enter into causal relations. The primary problem in the epistemology of mathematics is of how we can get to know something like a number, given that a number cannot be a cause of any sensation or belief in us. It is plausible that our belief that some item x exists can only constitute knowledge if either x itself has a causal role in our formation of this belief or if some cause of x has such a causal role. The former case occurs when we know from the smoke that there was a fire, and the latter when we know from the sound of a match struck that there will be a fire. But if something does not enter into any causal relations, then it seems that our belief about it is in no way affected by it or by anything connected with it, and hence our belief, if it coincides with the reality, does so only coincidentally, and hence not as knowledge. Of course, there are attempts to solve the conundrum on the books. But

simply not really a case of circular causation. The infinitesimal motion of each body at t does not cause the infinitesimal motion of the other body at t . Nor does the infinitesimal motion of either body at t cause the gravitational force of either body. Nor does the gravitational force of either body cause the gravitational force of the other. Rather, the gravitational force of each body partly causes the infinitesimal motion of the other body. Now, if we throw into the state of each body *both* its infinitesimal motion *and* its gravitational force, then we have a case where a part of the state of each body causes a part of the state of the other body. But that is not really circular causation, except maybe in a manner of speaking, just as it would not be circular causation to say that if I punch you in the shoulder while you punch me in the face, a part of my total state (the movement of my arm) causes a part of your total state (pain in your shoulder) and a part of your total state (the movement of your arm) causes a part of my total state (pain in my face). For there are four different parts of the total state involved, whereas in circular causation there would be only two.

the puzzle gives us some reason to rethink the dogma that numbers can neither cause or be caused.

But even if abstracta such as numbers and propositions are causally inefficacious, why should we think that there cannot be a nonabstract necessary being that is causally efficacious? One answer was already alluded to: some will insist that only spatiotemporal entities can be causally efficacious and it is implausible that a necessary being be spatiotemporal. But it was difficult to see why exactly spatiotemporality is required for causal connections. (See also Chapter 5 in this volume on the argument from consciousness.)

A different answer might be given in terms of a puzzlement about how there could be a nonabstract necessary being. The traditional way of expressing this puzzlement is that given by Findlay (1948), although Findlay may have since backpedaled on his claims. A necessary being would be such that it would be an analytic truth that this being exists. But it is never analytic that something exists. If $\exists x(Fx)$ is coherent, so is $\sim\exists x(Fx)$. Basically, the worry is caused by the Humean principle that anything that can be thought to exist can also be thought not to exist. But it is by no means obvious why this principle should be restricted, without thereby doing something *ad hoc*, to nonabstract beings.

Indeed, why should abstract beings alone be allowed as necessary? Why should its necessarily being true that $\exists x(x \text{ is a deity})$ be more absurd than its necessarily being true that $\exists x(x \text{ is a number})$? Perhaps the answer is that we can *prove* the existence of a number. In fact, mathematicians prove the existence of numbers all the time. Already in ancient times, it was shown that there exist infinitely many primes.

However, these proofs presuppose axioms. The proof that there are infinitely many primes presupposes a number system, say, with the Peano axioms or set theory. But a statement of the Peano axioms will state that there *is* a number labeled “0” and there *is* a successor function s such that for any number n , sn is also a number. Likewise, an axiomatization of set theory will include an axiom stating the existence of some set, for example, the empty set. If our mathematical conclusions are existential, at least one of the axioms will be so as well. The mathematical theory $\sim\exists x(x = x)$ is perfectly consistent as a mathematical theory if we do not have an existential axiom. Thus, if we are realists about numbers, we are admitting something which exists necessarily and that does not do so simply in virtue of a proof from nonexistential axioms.

Of course, one might not be a realist about abstracta. One might think that we do not need to believe that there necessarily exist propositions, properties, or numbers to be able to talk about necessarily true propositions or necessarily true relations between numbers or properties. But if the critic of the PSR had to go so far as to make this questionable move, the argument against the PSR would not be very plausible.

In any case, not all necessity is provability. We have already seen that the work of Gödel questions the thesis that all necessity is provability or analyticity (cf. Section 2.2.6.2, above). Kripke (1980), too, has questioned the same thesis on different grounds. That horses are mammals is a proposition we discover empirically and not one we can prove *a priori*. But it is nonetheless a *necessary* thesis. So is the proposition that every dog at some point in its life contained a carbon atom.

Now, it is admittedly true that the Kripkean necessities are not necessities of the *existence* of a thing. But they provide us with an example of necessarily true but not analytic propositions. Another such example might be truths of a correct metaphysics, such as that it is

impossible that a trope exists, or that it is necessary that a trope exists in any world containing at least two material objects that are alike in some way, or that there are properties, or that there are no properties. But, likewise, it could be that the true system of ontology entails the existence of God.

Another option for the defender of the epistemic possibility of a necessarily existing deity is provided by the ontological argument. The ontological argument attempts to show that from the concept of God one can derive the necessary existence of God. A necessary being could then be one for which there was a successful ontological argument, although perhaps one beyond our logical abilities. While the extant ontological arguments might fail (although see Chapter 10 in this volume), the best ones are valid and the main criticism against them is that they are question-begging. Thinking about such arguments gives us a picture of what it would be like for something to have a successful ontological argument for the existence of something. It could, thus, be that in fact God exists necessarily in virtue of an ontological argument that is beyond our ken, or perhaps the non-question-begging justification of whose premises is beyond our ken, while we know that he necessarily exists by means of a cosmological argument within our ken. We do not, after all, at present have any good in principle objection to the possibility that a sound ontological argument might one day be found (cf. Oppy 2006, chap. 2).

4.1.1.6. *Objection 5: the Taxicab Problem*

Since the First Cause that this cosmological argument arrives at is a necessary being, while the PSR as defended applies only to contingent states of affairs, the problem of applying the PSR to the existence of the necessary being does not arise. And even if one defended a PSR that also applies to necessary beings, one could simply suppose that the being's existence is explained by the necessity of its existence, or that there is a sound ontological argument that we simply have not been smart enough to find yet.

However, a different way to construe the Taxicab Problem is to ask about what happens when we apply the PSR to the proposition allegedly explaining the BCCF. But this issue has already been discussed when we discussed the van Inwagen objection to the PSR. There are two live options at this point. The proposition explaining the BCCF might be a contingent but self-explanatory proposition. For instance, it might be that the proposition *that a necessarily existing agent freely chose to do A for reason R* is self-explanatory in the sense that once you understand the proposition, you understand that everything about it has been explained: the choice is explained by the reason and the fact that the choice is free, and the necessity of the agent's existence is, perhaps, self-explanatory, or perhaps explanation is understood modulo necessary propositions. The other, I think preferable, option is that the BCCF is explained by a necessary proposition of the form: *a necessarily existing God freely chose what to create while impressed by reason R*.

One might continue to press a variant of the Taxicab Objection on this second option, using an argument of Ross (1969). Granted, it is necessary that God freely chose what to create while he was impressed by *R*. Let *q* be this necessary proposition. Let *p* be the BCCF. Then, even though *q* is necessary, it is contingent that *q* explains *p* (or even that *q* explains anything – for if God created something else, that likely would not be explained by his being impressed by *R*, but by his being impressed by some other reason). And so we can ask why *q* explains *p*.

But at least one possible answer here is not particularly difficult. The question comes down to the question of why God acted on R to make p hold. But God acted on R to make p hold because he was impressed by R . And God's acting on R to make p hold is explained by his making a decision while impressed by R , with its being a necessary truth that God's action is explained by R and by every other good reason. So, ultimately, q not only explains p , but also explains why q explains p . Had God acted on some other reason S that he is also impressed by to make not- p hold, then we could say that this was because he was impressed by S . (See Section 2.3.2.3, above.)

And so the PSR-based argument circumvents the Regress and Taxicab problems.

4.1.2. The CP for wholly contingent states

Following Koons (1997), suppose that each wholly contingent state of affairs has a cause. Form the maximal, wholly contingent state of affairs M as the mereological sum of all wholly contingent states of affairs. Then, M has a cause C by the CP. Moreover, plausibly, C is wholly disjoint from M . For suppose that C and M overlapped in some state J . Then, J would be caused by C . But J is a part of C , and although a substance can cause a part of itself, for example, by growing a new limb, it is absurd to suppose that a state of affairs as a whole can cause a part of itself. One difference here lies in the fact that states of affairs are nothing but the sums of their constituent states, while substances can cause parts of themselves by being more than the sum of their parts. So, C and M are wholly disjoint.

But if C and M are wholly disjoint, then C cannot be contingent. For if it were contingent, it would have a nonempty, wholly contingent part (see the argument in Koons 1997), and that part would then be a part of M .

Thus, the cause of M is a necessary state of affairs. As in the PSR case, plausibly only an existential state of affairs can cause an existential state of affairs. Hence, C involves the existence of something. Moreover, it had better involve the existence of a *necessary* being. The alternative is that C involves a quantificational state of affairs that says that some being satisfying a description D exists, where it is a necessary truth that some being satisfies D , but there is no being b such that it is a necessary truth that b satisfies D . However, such quantificational states of affairs are unlikely to be genuine causes, just as disjunctive states of affairs are not causes. Nor is this scenario, with its being necessary that some contingent being or other exist, plausible.

Hence, once again, we get to a necessary being. If, furthermore, we think that causes can only function through the causal efficacy of a being, then we get a causally efficacious necessary being, a First Cause.

The Taxicab Problem here takes the form of asking what causes C to cause M , and this question has already been discussed in Section 3.3.

We can also modify the argument by only assuming a CP for *positive* wholly contingent states of affairs. Assuming the cause of a positive state of affairs is found in a positive state of affairs, we can again get to a necessary being.

4.2. The CP for chains

We assume in this argument that we are dealing with a kind of causation such that the *causes* relation is transitive (if x causes y and y causes z , then x causes z) and irreflexive (x does not cause x). A *causal chain* of items is a set S of items totally ordered under the *causes*

relation. In other words, if x and y are members of S , then either $x = y$, or x causes y , or y causes x . The relata of the causal relation can be concrete things or events or concrete states of affairs – I shall call these relata “items.” The relevant CP now states that if S is a causal chain of contingent items, then there is a cause of every item in the chain. If S is a finite chain, this follows from a CP for individual items – we just take the first item in the chain and ask for its cause, which cause then is the cause of every item in the chain.

THEOREM 1. Assume the given CP, assume that causation is transitive and irreflexive, and assume the Axiom of Choice (AC). Suppose there is a set U of all items. Then, for any contingent item e , there is a necessary being G such that G causes e .

This Theorem was shown by Robert K. Meyer (1987) based on a suggestion of Putnam. The AC is a technical assumption needed for the proof. The AC states that if B is a nonempty set of disjoint nonempty sets, then there is at least one set C that has exactly one member from each of the members of B . The AC is obvious in finite cases. If $B = \{\{1,2\}, \{3,4\}, \{0,7,8,9\}\}$, we can just let $C = \{0,2,3\}$. Another way to put the AC is to say that the Cartesian product of nonempty sets is always nonempty. The AC is also equivalent to the claim that every set can be well ordered, that is, given a total ordering such that every subset has a least element. One direction of the proof of this equivalence is easy and illustrative of how the AC works: given a set B of disjoint nonempty sets, let B^* be the union of all the members of B ; suppose there is a well ordering on B^* ; let C be the set each of whose members is the *least* member of one of the members of B .

The AC is an assumption most working mathematicians are willing to grant as intuitively obvious, although mathematicians being who they are, if they can prove something without using the AC, they would rather do so. The AC is independent of the other axioms of Zermelo–Fraenkel set theory, so we cannot expect a proof of it. The difficulty in the AC is that no procedure for constructing C out of B is given – all we get is the fact that there is a C satisfying the requirements.

Theorem 1 follows quickly from Zorn’s Lemma, which is equivalent to the AC. Zorn’s Lemma says that if every nonempty chain (i.e. totally ordered subset) S in a nonempty partially ordered set V has an upper bound (i.e. an element y of V such that $x \leq y$ for all x in S), then V has a maximal element. To see that Theorem 1 follows, write “ $x \leq y$ ” provided that either y causes x or $y = x$. Let $U(e)$ be the set of all items in U that cause e . Suppose for a *reductio* that all the members of $U(e)$ are contingent. By the CP, every nonempty chain S of items in $U(e)$ has a cause y in U . But every item in $U(e)$ is a cause of e , and hence by transitivity this y will be a cause of e , and thus a member of $U(e)$. Hence, every nonempty chain in $U(e)$ has an upper bound. Thus, $U(e)$ has a maximal element. Call that element c . Since every member of $U(e)$ is contingent, c is contingent. Thus, by the CP, c has a cause, say b . Then by transitivity, b is also a cause of e , and hence a member of $U(e)$. But then $c < b$, and hence c is not maximal, contrary to the assumption. Thus, the claim that all members of $U(e)$ are contingent must be rejected. Thus, e has a noncontingent cause.

As before, if the Taxicab Problem amounts to asking for the cause of G , the problem cannot get off the ground, since we only assumed that contingent things had causes. Now in the case where the items are events, if G ’s causing e counts as an event, we can still ask for the cause of G ’s causing e . For responses to this question, we can refer back to Section 3.3.

4.3. *The Principle of Only Explanation (POE)*

White (1979) has proposed a principle that if only one putative explanation can be given of a phenomenon, that putative explanation is correct. As already mentioned, this is what Sherlock Holmes means by “when you have eliminated the impossible, whatever remains, *however improbable*, must be the truth” (Doyle 1890, p. 93; italics in the original). But only a necessary being’s causal efficacy can explain global explananda like the BCCF. Hence, a necessarily existing First Cause exists.

The difficulty with this principle is that it is not clear how one *counts* putative explanations. Suppose a coroner sees a woman with a wound and can rule out all explanations except a stabbing with a knife. Does the POE apply? After all, one might say that there is still more than one explanation available. Maybe the woman was stabbed with a knife for profit or maybe she was stabbed out of a revenge. If the latter two count as alternative explanations and make the POE not applicable, then POE does not apply to the cosmological case, since more than one set of motives could be assigned to the First Cause.

Suppose instead that POE still applies in the murder case. Then POE must be understood as follows. If at some specific level of generality only one putative explanation can be given, then that one explanation must be correct. At one level of generality, we have a stabbing with a knife. All alternatives to that have been ruled out. Hence, we need to accept that a stabbing with a knife happened.

However, now POE becomes much more controversial, and it is not clear that it is a gain over versions of the CP. In fact, for a wide class of items, it implies a CP. For at a very high level of generality that “a cause caused *E*” seems to be an explanation, and for a large class of items *E* it seems plausible to suppose that no other explanation would be possible. This last claim requires ruling out conceptual explanations. To do that we would have to work with the ontologically most basic items, where further conceptual explanations are impossible so only causal ones are available, and then say that if the most basic items have causes, so do the less basic ones.

Perhaps more specificity is required than just “a cause caused *E*.” Maybe “a necessary being caused *E*” has that kind of specificity, or maybe we can prove that only *God*’s activity could explain the BCCF. But significant amounts of work would be needed here.

And, besides, it is not clear why someone who accepts the POE would deny the PSR. Supposing that the coroner rules out all explanations other than being poisoned or having a heart attack, the inference to the claim that the person was either poisoned or had a heart attack would be just as good as the inference to the claim that she was poisoned if only that option remained.

4.4. *The R-PSR*

Recall the R-PSR, which claimed that every proposition that *can* have an explanation *does* have an explanation (see Section 3.2.2). The R-PSR enables a cosmological argument for one or more necessary beings whose existence explains why there is something contingent. It certainly does so if, as is claimed in Section 3.2.2, the R-PSR entails the PSR. But it also does so if that entailment argument fails.

I do, however, need a technical assumption:

- (52) There is a set Q of kinds such that: (a) for no x does x 's being of K , where K is in Q , depend on anything essentially originated, and (b) every contingent object x is a member of at least one kind from K .

If the cosmos consists of finitely many contingent objects, as seems fairly plausible empirically, then (52) is trivially true. Here, "kind" is a technical term. It does not simply mean a set of objects, but a classification falling under which is explanatorily prior to all the exercises of the entity's causal power. Natural kinds like *animal* and *electron* are paradigm instances of this. Whether kind membership is essential is a question we can stay neutral on.

Now, consider the proposition p that at least one of the kinds in Q has at least one contingently existing member. I claim that there could be an explanation of p . All we need to do to see this is to imagine a possible world that has a being that is not a member of any of the kinds in Q and is not caused into existence by any member or members of any of the kinds in Q , but which being causes at least one of the kinds in Q to have a member. The being might be a contingent being, as long as it is not a contingent being that is a member of any of the kinds exemplified in our world or dependent on any members of the kinds exemplified in our world. For instance, that being might be an angel or a witch that brings an electron into existence. Or it might be a necessary being, such as God.

Thus, possibly, p is explained. Thus, by the R-PSR there is an explanation for p . Assuming further that existential propositions about substances can only be explained by the causal activity of one or more substances, we conclude that there is a set U of substances whose causal activity explains p . If all of the members of U are contingent, then they are each a member of at least one kind from Q . Moreover, their kind membership is explanatorily prior to their causal activity. Thus, no one of these members of U can explain why *it* is a member of the kind (or kinds) it is a member of or why the kind (or kinds) that it is a member of has (have) a member, and hence no one can explain p . Neither can they collectively explain p , for collective causal powers derive from individual causal powers, and I have assumed that being the kind of being one is is explanatorily prior to having of the causal powers one does – one is able to thermoregulate *because* one is a mammal.

Therefore, at least one member of U is not contingent. But in fact, no contingent member of U should enter into an explanation of why some kind from Q has a member, since to have a contingent member of U exercising its causal powers, that contingent member must already have existed, and thus been a member of some kind from Q . So all the members of U are necessary beings. Therefore, in fact, one or more necessarily existing substances explain p through their causal activity.

5. The Gap Problem

5.1. Introduction

The last 50 years have seen significant progress in clarifying the philosophical issues involved in the Glendower, Regress, and Taxicab problems. Indeed, several rigorous versions of the cosmological argument are available to overcome these. The Gap Problem has yet to see as

much progress. Perhaps the reason is merely sociological. The typical philosophical atheist or agnostic not only does not believe in God but also does not believe in a necessarily existing First Cause. The typical philosopher who accepts a necessarily existing First Cause is also a theist. Thus, there is not much of an audience for arguments that the necessarily existing First Cause is God. Moreover, it makes sense to proceed in order – first, get clear on the argument for a necessarily existing First Cause, and only then on the argument that this is God.

Probably the most important part of the Gap Problem is the question whether the First Cause is an agent. After all, if the First Causes would have to be nonagentive necessarily existing substances that randomly spit out island universes, then the conclusion of the cosmological argument would be incompatible with theism.

In addition to the problem of personhood, there is the question of the other attributes that God has traditionally been believed to have: uniqueness, simplicity, omniscience, omnipotence, transcendence, and, crucially, perfect goodness. At the same time, it is quite reasonable for a defender of the cosmological argument to stop deriving attributes of the First Cause at some point, and say that the other attributes are to be accepted by a combination of faith and data from other arguments for the existence of God. In any case, rare is the Christian cosmological arguer who claims to be able to show that the First Cause is a Trinity, and indeed Christian theologians may say this is good, since that God is a Trinity is a matter of faith. Nor does the inability to show by reasoned arguments that the First Cause has some attribute provide much of an argument *against* the claim that the First Cause has that attribute.

There are two general approaches for bridging the gap between the First Cause and God: inductive and metaphysical. Inductive arguments may claim that supposing that the First Cause exemplifies some attribute is the best explanation of some feature of the First Cause's effects, and in doing so the arguments may reprise the considerations of design arguments. Typical metaphysical arguments, on the other hand, argue that a First Cause must have some special metaphysical feature, such as being simple or being pure actuality, from which feature a number of other attributes follow.

Considerations of space do not, however, allow a full discussion of these arguments, and of objections to them, so I shall confine the discussion to the barest sketches.

5.2. Agency

One might argue for *agency* in the causal activity of the First Cause in several ways. In order to not beg the question as to the number of First Causes, simply stipulatively define “*the First Cause*” as the aggregate of all First Causes – it might be a committee or a heap, but that is fine at this point.

If we got to the First Cause by means of the PSR, then the First Cause's activity must in some way explain everything contingent. If one accepts that all explanations of contingent states of affairs are either scientific, agential, or conceptual – at least these are all the kinds of explanations we know of, and since the concept of explanation is a concept of ours, we have some insight into what can and cannot yield an explanation – then one can argue that the First Cause is an agent. For the First Cause's activity does not provide a *scientific* explanation. As far as we can tell, science explains things in terms of contingent causes. Nor does the First Cause's activity *conceptually* explain everything contingent. In contingent reality we find substances, and the existence of a substance is not conceptually explained by the

activity of something other than that substance – substances are *self-standing*. At worst, the existence of a substance is conceptually explained by the existence of constituent parts, but if so, then these constituent parts will themselves be substantial. In the end we shall have to give a nonconceptual explanation, or else to find parts that are necessary. But it is false that goats and people are made up of necessarily existing parts. So, in the end, a nonconceptual explanation must be given. Hence, the explanation cannot be entirely conceptual. Since the explanation is not scientific either, it follows that it is at least in part agential, and hence the First Cause either is or contains a necessarily existing agent, unless there is some fourth relevant kind of explanation.

Alternately, one might argue that the only way to resolve the van Inwagen problem is to posit agency in the explanation of the BCCF. Perhaps only agential explanations in terms of a necessary being combine the two crucial features: contingency of effect and the impossibility of asking for a further explanation of some further contingent fact. However, if one thinks that nonagential statistical explanations can also have this feature, then this argument will not impress.

Finally, one can bring to bear the full panoply of design arguments available. The First Cause is an entity that has produced a universe apparently fine-tuned for life, containing beauty and creatures attuned to beauty, containing moral obligations and creatures aware of them; a universe containing conscious beings with free will; and a universe some of whose contents have objective functions (eyes are for seeing and so on – these kinds of functional attributions arguably cannot be reduced to evolutionary claims, although there is a large literature on this controversial claim). We have shown, let us suppose, that there is a First Cause. The further supposition that the First Cause is a highly intelligent and very powerful person acting purposively is highly plausible given all this data.

Finally, one might argue for agency on metaphysical grounds. If the metaphysical arguments show that the First Cause has every positive property, then the First Cause will in particular have knowledge and will, and hence be an agent.

5.3. Goodness

Whether we can argue on inductive grounds that the First Cause is good is a particularly difficult question in light of all the evil in the world. If the First Cause is an agent, we have three options to choose from: he is a good agent, an evil agent, or an agent morally in the middle. I will argue that at least we can dismiss the worst of these options on inductive grounds.

Here is one set of considerations. We might see evil as ontologically inferior to the good. For instance, we might see evil as a *privation* of the good. Or we might see evil as a *twisting* of the good: the good can stand on its own axiologically, but evil is metaphysically something parasitic. Seen from that point of view, evil can never be seen to be the victor. Whatever power evil has is a good power twisted to bad ends. Human cruelty is only an evil because human nature has a power of transcending cruelty. Evil can only mock the good but can never win.

Suppose we do indeed see things this way. Then evil only makes sense against a background of goodness. And hence, the cause that the universe originates in, since that cause is the ultimate background, cannot but be perfectly good. If, further, perfect good is *stable*, then we might think that this cause *still* is perfectly good. This will be a metaphysical argument.

Moreover, if we see evil as metaphysically inferior to the good, then the idea that the First Cause is an evil person makes the First Cause be rather stupid, and so we have an inductive argument against the worst of the three options under consideration. For whatever gets created, there will be more good than evil. Behind the twisting of human nature in a serial killer, there is the good of human nature – if it were not good, and if it were not in some way metaphysically superior to the evil so as to provide a standard against which that evil is to be measured, then the twisting would not be an evil. So by creating, the First Cause makes more good than evil come into existence, and if the First Cause is evil, then to do that is, well, stupid. But the fine-tuning of the universe suggests that the First Cause is highly intelligent.

Furthermore, I think it *is* fair to say that there is much more good than evil in the human world. Consider the constant opportunities available for malice, opportunities that would result in no punishment at all. We can assume, with almost total certainty, that if we ask strangers for the time, they will not look at the time and subtract 10 minutes just to make sure we are late for whatever appointment we are rushing. Is it not wondrous that I regularly find myself around many omnivorous animals armed with teeth and guns (I am in Texas!), but I have not yet suffered serious harm from them? At least on the assumption that these omnivorous animals were created by an evil being, there would be some cause for surprise. When the rules of morality are transgressed, rarely are they transgressed wantonly. Granted, there have been genocides of massive proportions. But it is noteworthy that even there, there tends to be a background that makes the cruelty not be *entirely* wanton: a destructive ideology or a vengeful, and often mistaken, justice. The victims are demonized. This demonization is itself an evil, but it is an evil that underscores the fact that the victims *need* to be seen as demonic before most of us will be induced to be cruel to them. The hypothesis that the First Cause is evil is not a very plausible one, then.

Whether the hypothesis that the First Cause is good is any more plausible will depend on how we evaluate the arguments of various theodicies. Some of the aforementioned considerations might possibly be the start of a theodicy, but that is not what I intended them for: I intended them merely as data against the hypothesis of an evil First Cause. On the theodicy front, on the other hand, we might see in freely chosen virtue a goodness outweighing the evils of vice, and that might lead us to suppose the First Cause is good.

5.4. *Simplicity and beyond*

It is at least plausible that if something has parts, then it makes sense to ask why these parts are united. If so, then the existence of a being with parts cannot be self-explanatory. The same is true of what one might call “metaphysical parts,” like distinct powers, tropes, and so on. If we suppose that the First Cause’s existence is self-explanatory, rather than explained in terms of some further metaphysical principles, then we might well conclude that there cannot be any composition in the First Cause. Taking this seriously leads to the well-known difficulties concerning divine simplicity (see Pruss 2008; Brower, 2008), but might also make possible Aquinas’ solution to the Gap Problem as given in the *Prima Pars* of the *Summa Theologiae*. (Note that this approach requires that we got to the First Cause through a PSR strong enough to allow us to ask for an explanation of the First Cause’s existence and of the composition of any elements in the First Cause.)

Such a Thomistic approach would start by noting that a strong doctrine of divine simplicity entails that there is no potentiality in the First Cause. Potentiality entails the possession of modally accidental intrinsic properties, that is, intrinsic properties that one might not have. But if the First Cause had any modally accidental intrinsic properties, then there would be the aspects of the First Cause that make true its having its particular contingent intrinsic properties and the aspects of the First Cause that make true its having its essential properties, and these aspects would have to be different because of the modal difference here. However, such a distinction would be contrary to a sufficiently strong doctrine of divine simplicity.

Moreover, Aquinas argues that one of the forms of simplicity that the First Cause has is a lack of a distinction between it, its essence and its existence, which is that by which it exists. This assures a kind of *aseity*: whereas our existence is at least dependent on our essence and conversely, in God there is no such dependency.

From lack of potentiality, Aquinas derives perfection in the sense of *completeness*. If there were something lacking in the First Cause, then the First Cause would have a potentiality for filling in that lack. But this is perhaps a kind of perfection that only metaphysicians will get excited about. If there were a particle that always had exactly the same intrinsic properties, and could have no others, it would count as perfect in this sense.

Aquinas' next step is to argue that "the perfections of all things" are found in God (Aquinas, forthcoming, I.4.2). Here we start to get something that the ordinary believer cares about. Aquinas offers two arguments. One of them depends on Aquinas' ontological system. Aquinas thinks each thing has existence, which gives it reality, and an essence that delimits the existence by specifying the kinds of reality that the object has. Thus, our essence specifies that we exist in respect of an ability to think and choose, as well as in respect of various physical abilities. In the First Cause, by divine simplicity, there is no essence distinct from existence to limit that existence, and so existence is found in an unlimited way: every "perfection of being" is found in the First Cause. Note that this argument not only yields the claim that the perfection of every actual being is found in the First Cause but also that the perfection of every possible being is found there. A full evaluation of this argument would require an evaluation of Thomistic ontology, and that is beyond the scope of this essay.

Aquinas' other argument relies on the scholastic axiom that:

the same perfection that is found in an effect must be found in the cause, either (a) according to the same nature, as when a man generates a man, or (b) in a more eminent mode. . . .
(Aquinas, forthcoming I.4.2)

This axiom is a staple of classic discussions of the existence of God, reappearing in Descartes' argument from our idea of God, and used by Samuel Clarke for the same purpose as in Aquinas. The idea is that a cause cannot produce something with a completely new kind of positive feature. A cause can produce combinations of positive features it has, as well as derivative forms of these. This would be a good axiom for cosmological arguers. But is it true?

Emergentist theories of mind are predicated precisely on a rejection of this axiom, but it will not do to use them as counterexamples to the axiom, since emergentism is controversial precisely because it allows for nonphysical properties to arise from physical ones in contravention of our axiom. One might try to find a counterexample to the axiom in evolutionary theory: beings that fly, see, think, walk, produce webs, and so on all come from

unicellular beings that can do none of these things. But here we should separate out the mental and the physical properties. It might be argued that there is no qualitative difference between flying, walking, and making webs, on the one hand, and doing the kinds of things that unicellular organisms do, on the other. Indeed, perhaps, we can argue that biology shows that these behaviors just *are* a matter of lots of unicellular organisms going through their individual behavioral repertoire, since higher organisms are composed of cells. On the other hand, whether mental properties can arise from things without them is controversial, and some accounts on which they can do so manage this feat simply by supposing that mental properties reduce to physical ones. Accounts that do not allow such reduction make the arising mysterious, and our being mystified here is a testimony to the plausibility of the axiom.

A different kind of objection to the axiom is an *ad hominem* one: the axiom is incompatible with theism because the peculiar perfections of material objects can only be found in material objects. Thus, the cause of material objects, God, must either lack some of the perfections of material objects, in which case the axiom is false, or else God is material, contrary to theistic orthodoxy. However, the axiom as Aquinas understands it allows that a more eminent version of a perfection could be found in the cause than in the effect. It could be that omnipresence is helpful here. Thus, God's omnipresence could be a more eminent version both of perfections of shape and movement. Thus, the earth is spherical, and God is not spherical (*pace* Xenophanes), but God by his omnipresence is also everywhere where the earth is, and so he has a more eminent version of sphericity. The cheetah certainly can run fast, whereas God cannot run, but God is always already where the cheetah's run ends, while also being at the starting line.

Unfortunately, these are only defensive maneuvers. It may be that the axiom is self-evident, but simply asserting its self-evidence will not help those who do not see it as such. And there has been very, very little attempt in contemporary philosophy to give a good argument for the axiom. Historically, Samuel Clarke (1823) had tried. His argument was that if a perfection comes from something that does not have it, then the perfection comes from nothing, and it is absurd for something to come from nothing. But that argument misunderstands what opponents of the axiom think: they do not, presumably, think that the perfection comes from nothing, but from something different in kind from itself.

If we do accept this argument for the First Cause's having all the perfections of created things, we can proceed to argue further as follows. The First Cause either is or is not the First Cause in every nonempty possible world. If it is the First Cause in every possible world, and these arguments are sound and work in all possible worlds, then in every world, it is true that the First Cause has all the perfections of the things in that world. Assuming perfections are intrinsic properties, it follows that what perfections the First Cause has cannot differ between worlds, since there is no contingency in the First Cause, by simplicity. Therefore, any perfection the First Cause has in one world, it has in all worlds. Consequently, the First Cause not only has all the perfections of the things that exist in our world but also all the perfections of the things that exist in any possible world.

Can different worlds have different First Causes? One way to settle this is stipulatively. Just let the First Cause be the aggregate of all the necessary beings. Any First Cause is a necessary being, and now our First Cause is indeed the First Cause of every world, and hence has all the perfections of things that exist in any possible world.

This seems to imply that the First Cause of the Leibnizian cosmological argument is the same being who is found in the conclusion of ontological arguments for a being

with all perfections. In particular, if we allow that personhood is a perfection, it follows that the First Cause either is a person, or has some quality that is even greater than personhood.

At the same time, the aggregation move that we had made raises the possibility that the First Cause is a polytheistic committee, having all perfections collectively but with no one deity having them all individually. If Aquinas is right that in any First Cause there must be identity between the thing and its existence, and that having all perfections follows from this identity, then the worry does not arise. If this is not a satisfactory solution, we may need to employ some other argument for the unity of the First Cause, such as that if there were multiple necessary beings, we would not expect to see a nomically unified world (Aquinas, forthcoming, I.11.3).

From being perfect and having all perfections of things, of course, the sailing is fairly smooth, just as it is after one has come to the conclusion of an ontological argument. Aquinas, thus, proceeds to goodness, infinity, omnipresence, immutability, eternity, unity, knowability, omniscience, and omnipotence.

5.5. Gellman's argument for oneness and omnipotence

Jerome Gellman (2000) has offered a clever argument from the claim that in every possible world there is a necessarily existing cause that explains all contingent truths (perhaps a different one in different worlds) to the claim that there is a necessarily existing cause that is *omnipotent* and that explains all contingent truths in every world. The argument is intricate, and here I shall give a variant that I think is in some ways superior.

If N is a necessary being that explains all the contingent truths of a world w , I shall call N "a creator in w ." I shall assume the Iterativeness Postulate (IP):

- (IP) If x has the power to gain the power to do A , then x already has the power to do A , although x might have to take two steps to do A (first acquire a power to directly do A , and then exercise the power).

It follows from IP that if N is a creator in w , then the powers of N are necessary properties of N . To see this, for a *reductio* suppose that w is actual and N contingently has the power to do A . Then N 's causal activity explains why N has the power to do A , since N 's causal activity explains all contingent truths. But then explanatorily prior to N 's causal activity, N had the power to bring it about that it had the power to do A . But by IP, N had the power to do A explanatorily prior to N 's causal activity, which contradicts the claim that this causal activity explains the power.

Next, we show that a creator N_1 in w_1 and a creator N_2 in w_2 must be the same individual. Suppose first that w_1 and w_2 are distinct worlds. Let p be some contingent proposition true in w_1 but not true in w_2 . Beings N_1 and N_2 exist necessarily, and hence both exist in w_1 . Let q be the proposition that N_2 's causal activity does not explain not- p . This proposition is true at w_1 since not- p is false at w_1 and only true propositions have explanations; on the other hand, q is false at w_2 . Since N_1 is a creator in w_1 , N_1 's causal activity explains q . Therefore, N_1 in w_1 has the power to make q true, a power it exercises. By what we have already shown, N_1 *essentially* has the power to make q true, and hence it also has this power in w_2 . Call this power P_1 . We can now ask why it is the case at w_2 that N_1 fails to exercise this

power. Since N_2 is a creator in w_2 , we must be able to explain N_1 's contingent failure to exercise P_1 in terms of N_2 's causal activity. Therefore, N_2 at w_2 has the power to prevent N_1 from exercising P_1 . Call this power P_2 . By what has already been shown, N_2 has P_2 essentially.

Moreover, N_2 does not exercise P_2 at w_1 , since at w_1 N_1 does exercise P_1 . Why does N_2 fail to exercise P_2 at w_1 ? This must be explained in terms of N_1 's causal activity, just like all other contingent facts about w_1 . Hence, at w_1 N_1 has the power, P_3 , of preventing N_2 from exercising P_2 . Hence, N_1 has that power essentially and is prevented at w_2 from exercising it by N_2 . Therefore, arguing as before, N_2 essentially has the power, P_4 , of preventing N_1 from exercising P_3 . And so on.

This regress seems clearly vicious, and so we conclude that N_1 cannot be distinct from N_2 (if $N_1 = N_2$, we can say that what explains N_2 's not bringing it about that p at w_1 is simply that N_2 brings it about that not- p at w_1 , and then we can reference our previous discussions of libertarian explanations in Section 2.3.2.3, above). But perhaps we can make the argument work even without going through with the regress. What explains at w_1 , we may ask, why it is that N_2 exercised *none* of its powers to prevent N_1 from engaging in the kind of activity it engages in in w_1 ? It must be that the explanation lies in the exercise of some power P by N_1 in w_1 . But then N_1 also had this power in w_2 and did not exercise it, and its failure to exercise it must be explained by N_2 's exercise of some preventative power Q . But Q is one of the powers whose exercise in w_1 is prevented by N_1 's exercise of P . Repeating the argument with the two entities and worlds swapped, we conclude that each of N_1 and N_2 has the power to prevent the other from its preventing the other. But that is, surely, absurd! (It might not be absurd if $N_1 = N_2$ since in having the power to do A , I have the power to prevent myself from not doing non- A , but that is likely just because my doing A is identical with my refraining from doing non- A .)

So, if N_1 is a creator in w_1 and N_2 is a creator in w_2 , then $N_1 = N_2$. It also follows that each world has only one creator. For if N_1 and N_2 were each a creator in w_1 , then we could choose any second world w_2 , let N_3 be a creator in w_2 , and use the said argument to show that $N_1 = N_3$ and $N_2 = N_3$, so that it would also follow that $N_1 = N_2$.

Thus, there is a unique being that essentially has the power to explain every contingent truth in every world via its causal activity. But, surely, having the power to explain every possible contingent truth via one's causal activity implies omnipotence. (We can stipulate this if need be, and the stipulation will not be far away from ordinary usage.)

There are two difficulties in this line of argument. The first is that it requires that each world have one being that *by itself* explains all contingent truth. What if one takes the cosmological argument only to establish the weaker claim that there is at least one necessary being and the necessary beings collectively explain all contingent truths? In that case, the said argument can still be applied, with "a creator" being allowed to designate a collective and not just an individual. The conclusion would be that the very same omnipotent collective explains contingent truth in every world. Can there be an omnipotent collective? It is tempting to quip that there is a conceptual impossibility in a committee's being omnipotent, since committees always suffer from impotence, say, due to interaction issues within the committee. There may be something to this quip. How, after all, could a collective collectively be omnipotent? How would the powers of the individuals interact with one another? Would some individuals have the power to prevent the functioning of others? These are difficult questions. It seems simpler to posit a single being.

The second difficulty is that on this argument, the creator's causal activity explains *all* contingent activity, including, presumably, any free choices by creatures. This problem infects other Leibnizian cosmological arguments. Probably, the way to handle it is to give a subtler and more careful definition of what it is to be a creator in *w*. Maybe the First Cause's activity does not have to explain all free choices made by everybody; it may simply have to explain both the *prerequisites* for all free choices made by any contingent beings, and everything that does not depend on the free choices of contingent beings? This is probably all we need for the crucial uniqueness argument.

6. Conclusions and Further Research

The cosmological argument faces the Glendower, Regress, Taxicab, and Gap problems. Cosmological arguments using a sufficiently comprehensive CP or an appropriate PSR are able to overcome the Regress and Taxicab Objections. The Glendower Problem of justifying the explanatory principle is an important one. However, in recent years, a number of arguments for such explanatory principles, as well as weaker versions of these principles still sufficient for the purposes of the cosmological argument, have been produced. There is, of course, still much room for research here: for examining arguments for or against the relevant explanatory principles, and for trying to produce cosmological arguments using yet weaker principles.

What contemporary analytic philosophers have not sufficiently worked on – and what is perhaps the most promising avenue for future research – is the Gap Problem. There are both inductive and deductive approaches here. The deductive ones that are currently known proceed through exciting metaphysical territory of independent interest. The metaphysics of existence/essence composition involved in Aquinas' bridging of the gap is fascinating, and the axiom that the perfections of the effect must be found in the cause is one that needs further exploration, both in connection with the cosmological argument, as well as in connection with emergentist theories of mind.

References

- Adams, R. M. (1974) Theories of actuality. *Noûs* 8, 211–31.
- Aquinas, T. (Forthcoming) *Summa Theologiae*. Trans. A. Freddoso. Also available online at <http://www.nd.edu/~afreddos/summa-translation/TOC.htm>.
- Black, M. (1952) The identity of indiscernibles. *Mind* 61, 153–64.
- Brower, J. E. (2008) Making sense of divine simplicity. *Faith and Philosophy* 25, 3–30.
- Campbell, J. K. (1996). Hume's refutation of the cosmological argument. *International Journal for Philosophy of Religion* 40: 159–73.
- Clarke, S. (1823) *A Discourse concerning the Being and Attributes of God, the Obligations of Natural Religion, and the Truth and Certainty of the Christian Revelation*. Glasgow: Griffin and Co.
- Davey, K. and Clifton, R. (2001) Insufficient reason in the 'new cosmological argument'. *Religious Studies* 37, 485–90.
- Doyle, A. C. (1890) *The Sign of Four*. London: Spencer Blackett.
- Dretske, F. (1972) Contrastive facts. *Philosophical Review* 81, 411–37.
- Dudley, U. (1987) *A Budget of Trisections*. New York: Springer.
- Edgington, D. (1995). On conditionals. *Mind* 104, 235–329.

- Edwards, P. (1959) The cosmological argument. *The Rationalist Annual for the Year 1959*. London: Pemberton. Reprinted in D. R. Burrill (ed.), *The Cosmological Argument*, New York: Doubleday, 1967.
- Elga, A. (2001) Statistical mechanics and the asymmetry of counterfactual dependence. *Philosophy of Science* 68, S313–24.
- Empiricus, S. (1993) *Outlines of Pyrrhonism*. Trans. R. G. Bury. Cambridge, MA: Harvard University Press.
- Findlay, J. N. (1948) Can God's Existence. Be Disproved? *Mind* 57, 176–183.
- van Fraassen, B. (1980) *The Scientific Image*. Oxford: Oxford University Press.
- Francken, P. and Geirsson, H. (1999) Regresses, sufficient reasons, and cosmological arguments. *Journal of Philosophical Research* 24, 285–304.
- Gale, R. M. (1999) *On the Nature and Existence of God*. Cambridge: Cambridge University Press.
- Gale, R. M. and Pruss, A. R. (2002) A response to Oppy and to Davey and Clifton. *Religious Studies* 38, 89–99.
- Gellman, J. (2000) Prospects for a sound stage 3 of cosmological arguments. *Religious Studies* 36, 195–201.
- Hempel, C. G. (1962) Deductive-nomological vs. statistical explanation. In H. Feigl and G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science*, vol. III, 98–169. Minneapolis: University of Minnesota Press.
- Hempel, C. G. and Oppenheim, P. (1948) Studies in the logic of explanation. *Philosophy of Science* 15, 135–75.
- Hume, D. (1907) *Dialogues concerning Natural Religion*. Edinburgh: Blackwood and Sons.
- van Inwagen, P. (1983) *An Essay on Free Will*. Oxford: Oxford University Press.
- van Inwagen, P. (1986) Two concepts of possible worlds. *Midwest Studies in Philosophy* 11, 185–213.
- Johnson, D. (Forthcoming) The sense of deity and begging the question with ontological and cosmological arguments. *Faith and Philosophy*.
- Koons, R. C. (1997) A new look at the cosmological argument. *American Philosophical Quarterly* 34, 193–212.
- Kripke, S. (1980) *Naming and Necessity*. Cambridge, MA: Harvard.
- Leslie, J. (2001) *Infinite Minds*. Oxford: Oxford University Press.
- Lewis, D. (1986) *On the Plurality of Worlds*. Malden, MA: Blackwell.
- Lukasiewicz, J. (1961) O determinizmie. In J. Ślupiecki (ed.), *Z zagadnień logiki i filozofii*. Warsaw. Trans. Z. Jordan in S. McCall (ed.), *Polish Logic*, Oxford: Clarendon, 1967.
- Meyer, R. K. (1987) God exists! *Notus* 21, 345–61.
- Norton, J. D. (2003) Causation as folk science. *Philosopher's Imprint* 3, 1–22.
- O'Connor, T. (2008) *Theism and Ultimate Explanation*. Malden, MA: Blackwell.
- Oppy, G. (2006) *Arguing about Gods*. Cambridge: Cambridge University Press.
- Plantinga, A. (1974) *The Nature of Necessity*. Oxford: Clarendon.
- Pruss, A. R. (1998) The Hume-Edwards principle and the cosmological argument. *International Journal for Philosophy of Religion* 43, 149–65.
- Pruss, A. R. (2003) A new free will defense. *Religious Studies* 39, 211–33.
- Pruss, A. R. (2004a) A restricted principle of sufficient reason and the cosmological argument. *Religious Studies* 40, 165–79.
- Pruss, A. R. (2004b) David Lewis's counterfactual arrow of time. *Notus* 37, 606–37.
- Pruss, A. R. (2006). *The Principle of Sufficient Reason: A Reassessment*. Cambridge: Cambridge University Press.
- Pruss, A. R. (2007). Conjunctions, disjunctions and Lewisian semantics for counterfactuals. *Synthese* 156, 33–52.
- Pruss, A. R. (2008) On two problems of divine simplicity. In J. Kvanvig (ed.), *Oxford Studies in Philosophy of Religion*, vol. 1, 150–167. Oxford: Oxford University Press.

- Rescher, N. (2000) *Nature and Understanding: The Metaphysics and Method of Science*. Oxford: Clarendon.
- Ross, J. F. (1969) *Philosophical Theology*. Indianapolis: Bobbs-Merrill.
- Rowe, W. L. (1975) *The Cosmological Argument*. Princeton, NJ: Princeton University Press.
- Rowe, W. L. (1984) Rationalistic theology and some principles of explanation. *Faith and Philosophy* 1, 357–69.
- Salmon, W. (1990) *Four Decades of Scientific Explanation*. Minneapolis: University of Minnesota Press.
- Shakespeare, W. (2000) *Henry IV: Part I*. Ed. Claire McEachern. New York: Penguin Bantam.
- Shapiro, L. S. (2001) 'The transition from sensibility to reason *in regressu*': indeterminism in Kant's *Reflexionen*. *Kant-Studien* 92, 3–12.
- Smart, J. J. C. and Haldane, J. J. (2003) *Atheism and Theism*, 2nd ed. Malden, MA: Blackwell.
- Smith, Q. (1999) The reason the universe exists is that it caused itself to exist. *Philosophy* 74, 579–86.
- Sullivan, T. D. (1994) On the alleged causeless beginning of the universe: a reply to Quentin Smith. *Dialogue* 33, 325–35.
- Taylor, R. C. (1974) *Metaphysics*, 2nd ed. Englewood Cliffs, NJ: Prentice Hall.
- White, D. E. (1979) An argument for God's existence. *International Journal for Philosophy of Religion* 10, 101–15.